Intelligent Structure and Vibration Control

Edited by
Shaobo Zhong, Yimin Cheng and Xilong Qu

TRANS TECH PUBLICATIONS

Intelligent Structure and Vibration Control

Edited by:
Shaobo Zhong
Yimin Cheng
Xilong Qu

Intelligent Structure and Vibration Control

Selected, peer reviewed papers from the
International Conference on
Intelligent Structure and Vibration Control (ISVC)2011,
January 14-16, 2011, Chongqing, China

Edited by

Shaobo Zhong,Yimin Cheng,Xilong Qu

TRANS TECH PUBLICATIONS LTD
Switzerland • UK • USA

Trans Tech Publications Ltd
Laubisrutistr. 24
CH-8712 Stafa-Zurich
Switzerland
http://www.ttp.net

Volumes 50-51 of
Applied Mechanics and Materials 2-vol.-set
ISSN 1660-9336

Full text available online at http://www.scientific.net

Distributed worldwide by

Trans Tech Publications Ltd
Laubisrutistr. 24
CH-8712 Stafa-Zuerich
Switzerland

Fax: +41 (44) 922 10 33
e-mail: sales@ttp.net

and in the Americas by

Trans Tech Publications Inc.
PO Box 699, May Street
Enfield, NH 03748
USA

Phone: +1 (603) 632-7377
Fax: +1 (603) 632-5611
e-mail: sales-usa@ttp.net

ISVC 2011

Committees

Conference Chairs

J.X. Chen, Gradiant ETSI Telecomunication, Spain

B.Gao, Centrum Wiskunde & Informatica, The Nethelands

Y.M.Cheng, Yanshan University, China

Program Committee Chairs

B.J. Zhang， Henan Institute of Science and Technology ,China

A.M.Yang, Hebei Polytechnic University, China

Y.P. Luo, Hunan Institute of Engineering, China

Organizing Chair

M.N.Gui, Texas State University, USA

L. Wang, Hebei Polytechnic University, China

X.L.Qu,Hunan Institute of Engineering, China

Contact Co-Chairs

G.CZhou Hebei Polytechnic University, China

J.G.Qu, Hebei Polytechnic University, China

Publication Chair

S.B.Zhong, Chongqing Normal University, China

Technical Program Committee

J.N. Jue, University of Texas at Dallas, USA
C.X. Zhang, Juniper Networks, USA
J.X. Chen, University of Vigo, Spain
L. Zhou, ENSTA-ParisTech, France
R.C. Tang, Ocean University of China, China
H.R. Yoshino, NTT, Japan
A.M. Yang, Hebei United University, China
M.N. Devetsikiotis, North Carolina State University, USA
C.J. Wang, Northwestern University, USA
M. Chen, NEC Laboratories America, USA
Z.X. Duan, Florida State University, USA
N.X. Ghani, University of New Mexico, USA
W.J. Kabacinski, Poznan University of Technology, Poland

ISVC 2011 Reviewers

T.R. Melia, Cisco Systems, Switzerland
W.M. Eddy, NASA/Verizon, USA
H.R. Haddadi, University College London, UK
J.H. Ott, Helsinki University of Technology, Finland
S.-Y. Hua, National Central University, Taiwan
K.J. Li, University of Georgia, USA
G.H.Gao，Henan Institute of Science and Technology ,China
M.J. Stiemerling, NEC Europe Labs, Germany
F. Kuo, University of Goettingen, Germany
J.G. Qu, Hebei United University, China
J.X. Chen, University of Vigo, Spain
R.C. Tang, Ocean University of China, China
L.X. Shi, IBM Research Lab, China
R. Zhu, South-Central University for Nationalities, China
G.X. Danilewicz, Poznan University of Technology, Poland
Y.Z. Xu, Polytechnic University, USA
K.W. Xu, Arizona State University, USA
H.J. Wang, College of William and Marry, USA
C.X. Wu, University of Toronto, Canada
S.H. Yan, Hebei United University, China
W.M. Zhang, North Dakota State University, USA
D.L. Liu, ShenYang Ligong University, China
Y.X. Tang, Illinois State University, USA
L. Zhou, ENSTA-ParisTech, France
A.M. Yang, Hebei United University, China
W.X. Wei, NEC Labs America, USA
Y.L. Zhao, Northwestern University, USA
G.F. Gui, CMC Corporation, China
A.D. Baiocchi, University of Roma, Italy
L.C. Feng, Hebei Polytechnic University, China
M.Z. Liao, Cisco China R&D Center, China
O.X. Altintas, Toyota InfoTechnology Center, Japan
S.J. Giordano, University of Pisa, Italy
Y. Lin, Norwegian University of Science and Technology, Norwegian
Y.W. Zou, Apple China Co., Ltd., China
X.G. Zhang, Nanyang Normal University, China
Z.G. Wei, Beijing Forestry University, China
Y.J. Li, Shanghai Jiao Tong University, China
H.B. Nguyen, The Aerospace Corporation, USA

PREFACE

Dear Distinguished Delegates and Guests,

International Conference on Intelligent Structure and Vibration Control 2011 (isvc2011) was held in Chongqing, China, from January 14-16, 2011, serving as a platform for expertise exchange. ISVC 2011 had drawn the attention of researchers from various disciplines: Advanced Intelligent Structure, Bio-inspired Smart Materials and Structures, Materials for energy storage, Active Materials, Mechanics and Behavior, Vibration and Control, Modeling, Simulation, Control and Applications, Fuzzy System and Fuzzy Control, Management Information Systems, etc.

Persons who attended the conference were engineers, scientists, managers of various companies and professors of the universities abroad and home. We have had record number of submission 683 this year. Only original and unpublished paper would be considered, and 200 papers have been accepted for presentation at the conference and will be published by TTP, in Applied Mechanics and Materials (AMM) Journal (ISSN: 1660-9336), which is online available in full text via the platform www.scientific.net. AMM should be indexed by EI according the previous TTP index results.

We express our special gratitude to all the members of the General Committee Chairs, Program Committee Chairs, Technical Program Committee and Steering Committee who worked so hard to prepare the conference and who supported the conference so professionally.

Our deep thanks also go to the sponsors: Shanghai Jiao Tong University, Nanyang Normal University, Hebei Polytechnic University, Henan Institute of Science and Technology, Chongqing University of Arts and Sciences, Hunan Institute of Engineering, their kind support in making ISVC 2011 possible.

Finally, we would like to thanks all the authors, speakers and participants of this conference for taking part in and contributing to the International Conference on Intelligent Structure and Vibration Control 2011.

We hope you have a unique, rewarding and enjoyable week at ISVC 2011 in Chongqing.

With our warmest regards,

ISVC2011 Organizing Committees
January 14-16, 2011
Chongqing, China

Table of Contents

Part 2

Part 1

doi:10.4028/www.scientific.net/AMM.50-51.3

CDMA Multiuser Detection Based on Improved Particle Swarm Optimization Algorithm

Nanping LIU[1,2,a], Fei ZHENG[1,3,b,†] and Kewen XIA[1,4,c]

[1]School of Information, Hebei University of Technology, Tianjin 300401, China

[2]Tianjin Normal University, Tianjin, 300387, China

[3]Erik Jonsson School of Engineering & Computer Science, University of Texas at Dallas, Richardson, TX 75080, USA

[4]Dept. of IESE, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[a]liunanping05@126.com, [b]fei.zheng@student.utdallas.edu, [c]kewenxia@illinois.edu

[†]corresponding author

Keywords: CDMA, Multiuser detection, Improved particle swarm optimization, Simulation analysis

Abstract. CDMA multiuser detection (MUD) is a crucial technique to mobile communication. We adopt improved particle swarm optimization (PSO) algorithm in MUD which incorporates factor w and utilizes *sigmoid* function to discrete PSO. Comparison of BER and near-far effect has verified its effectiveness on multi-access interference (MAI). The algorithm accelerates the convergent speed meanwhile it also displays feasibility and superiority in case simulation.

Introduction

Wideband CDMA is an important technology in 3G mobile communication. Multiuser detection (MUD) serves as a key technique in anti-interference in communication, which was first proposed by Schneider in 1979 and was comprehensively studied by Prof. S. Verdu afterwards [1].

Performance of conventional MUD is hampered mainly by the transmitting of continuous training data series, ensued by the waste of spectrum frequency. Moreover the detector is undermined where the channel varies greatly.

New optimization algorithms have been applied to multiuser detection to deal with the inherent deficiency of MUDs, such as artificial neural networks, genetic algorithm, swarm intelligence, which can be categorized as computational intelligence.

Particle Swarm Optimization (PSO) is a swarm intelligent algorithm that performs better than genetic algorithm in many aspects [2, 3, 4]. Generally most of engineering problems can be transformed into combinatorial optimization, thus we improve PSO as discrete PSO in applying to MUD. Specifically, the factor w is introduced and *sigmoid* function is deployed as constraint to discrete PSO.

In section two we introduce the measure of MUD performance. The third section talks about improved PSO algorithm and MUD model based on it. Comparison and analysis on this model's performance comes at the fourth section. Finally we arrive at further conclusion and outlook future work.

CDMA multiuser detectors

Multiuser Detection (MUD) analyses signal of single user out of all the users by detecting the information from multiple access interference. It has proved to enjoy the capability of anti-interference and solving near-far effect. MUD helps to lower the accuracy demand that system controls the power; therefore the system can effectively exploit the upper link resource and enlarge the capacity of system. Fig. 1 exhibits types of multiuser detectors.

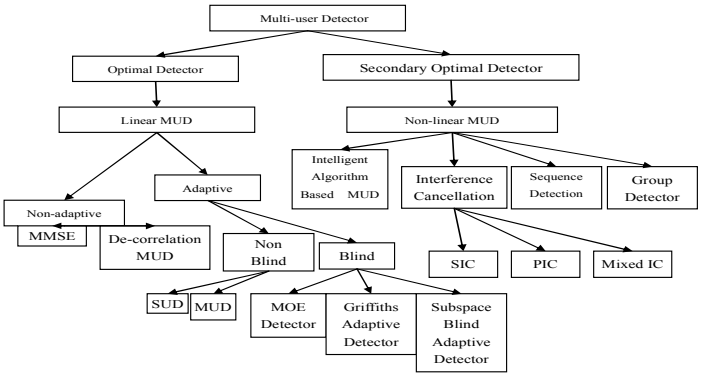


Figure 1. Types of multiuser detectors

Prof. Verdu has put forward an optimal multiuser detection (OMD) theory. It selects $b^* = [b_1^*, b_2^*, \ldots, b_K^*]^T$ that maximizes the likelihood function, according to the waveform of received signal $r(t)$ in every symbol interval $[jT_b, (j-1)T_b]$. When all the signal series are transmitted in equal probability, the OMD equation can be written

$$b^* = \arg\{\max_{b\in\{-1,+1\}^K}[-\frac{1}{2}b^T RWb + y^T b]\} \quad (1)$$

where $b^* = [b_1^*, b_2^*, \cdots, b_K^*]$ is the output vector of OMD; K is the total number of users; $b = [b_1, b_2, \cdots, b_k]^T$ is the transmitted data series from disparate users; $R = [r_{ij}]_{K\times K}$ is the PN code correlation matrix of every user which is symmetric, and $r_{ij} \neq 0$ e.g. MAI is eliminated in the condition that PN codes are incomplete orthogonal. W is the diagonal matrix of energy matrix whose diagonal elements represent the received energy of the i th user.

Bit Error Rate. In the situation of single user, there derives a Gaussian stochastic variable with mean of $\sqrt{w_k}$ and variance of σ^2 after match filtering the received signal. Bit error rate (BER) of single user is

$$P_e(k) = Q(\sqrt{w_k}/\sigma) \quad (2)$$

where

$$Q(x) = \frac{1}{2\pi}\int_x^\infty e^{-t^2/2}dt \quad (3)$$

Anti near-far effect. Lucas has proposed the degree of anti-near-far effect to calibrate the performance of MUDs, which represents the minimum tangent efficiency of detector when the energy of other users vary; namely

$$\bar{\eta}_k = \inf_{\substack{\sqrt{w_i}>0,\\ (i,j)\neq(0,k)}} \eta_k \quad (4)$$

Discrete particle swarm optimization algorithm

Particle swarm optimization algorithm mimics the foraging behavior of flocks of birds to solve optimization problems. As for combinatorial optimization problems, Kennedy and Eberhart propounded binary discrete PSO based the probability model to determine “yes” or “no” [5]. Suppose m particles consists of a swarm $x = \{x_1, x_2, \cdots, x_m\}$ in a D-dimension foraging space, the information

of particle i is represented by a D-dimension vector as $x_i = [x_{i1}, x_{i2}, \cdots, x_{iD}]^T$, $i = (1, 2, \cdots, m)$, where x_i is a vector composed of $\{1,0\}$ with its velocity being $v_i = [v_{i1}, v_{i2}, \cdots, v_{iD}]^T$.

To begin with, the m particles are randomly initialized, and then we have to trace individual optimum $p_{id} = [p_{i1}, p_{i2}, \cdots, p_{iD}]$ and global optimum $p_{gd} = [p_{g1}, p_{g2}, \cdots, p_{gD}]^T$, $i = (1, 2, \cdots, m)$. The velocity and location of particle i is updated by

$$v_i^{k+1} = v_i^k + c_2 r_1 (p_{id}^k - x_i^k) + c_2 r_2 (p_{gd}^k - x_i^k). \tag{5}$$

$$sigmoid(v_{id}^{k+1}) = \frac{1}{1 + \exp(-v_{id}^{k+1})} \ (d = 1, 2, \cdots, D). \tag{6}$$

$$x_{id}^{k+1} = \begin{cases} 1 & if \ \rho_{id}^{k+1} < sigmoid(v_{id}^{k+1}) \\ 0 & if \ \rho_{id}^{k+1} \geq sigmoid(v_{id}^{k+1}) \end{cases}. \tag{7}$$

where random number $\rho_{id}^{k+1} \in [0,1]$, k is the iteration; r_1 and r_2 are random variables between $[0,1]$; c_1 and c_2 are called the acceleration coefficient which adjust the maximum velocity step of global best and personal best respectively, practically we set $c_1 = c_2 = 2$. The interval of v_{id}^k is set confined by $[-4,+4]$ lest $sigmoid$ saturate and the particles may converge into local optimum early.

Improved MUD model based on DPSO

Standard PSO MUD is introduced and analyzed in [6].

Information of the target user compares to the food, information of any user acts as a bird. The distance amidst any bird and the food is quantified by a fitness function, herein we adopt OMD equation. The searched out minimum is regarded as optimal solution, thus finally we attain the target user information. Parameters are set as follows.

The system has K users which form a K dimension space; there is assumed to be N particles and the location of i th particle is $b_i = [b_{i1}, b_{i2}, \ldots, b_{iK}]^T$, and its velocity is $v_i = [v_{i1}, v_{i2}, \ldots, v_{iK}]^T$, $i = 1, 2 \ldots N$. Fitness of particle i at current location is

$$E_i = -F(b_i) = f(b_i). \tag{8}$$

where fitness function is

$$f(b) = 2b^T Ay - b^T ARAb. \tag{9}$$

which is the maximum probability in OMD.

Current optimal solution of particle i is $p_{ibest} = [p_{i1}, p_{i2}, \cdots, p_{iK}]^T$ marked by b_i^{pbest} as its corresponding location; optimal solution searched by the swarm is $g_{best} = [g_1, g_2, \cdots, g_K]^T$, denoted by b^{gbest} as its location.

Therefore the problem has been transformed from finding the optimal vector of MUD into searching the global optimum location in DPSO. Fig. 2 illustrates the model of DPSO MUD.

The DPSO based MUD follows the main procedure as

Step 1: Initialize the particles. The element b_i of location vector is randomly produced by -1 or +1.

Step 2: Define the fitness function as (9) and set p_{is} as personal best location p_{ibest}, p_g as initial global best location p_{gbest} among the swarm.

Step 3: As for all the particles, implement the subsequent operations

1) Update the location and velocity of all particles by (5), (6) and (7)

2) Calculate fitness of each particle. Location of p_{ibest} will be updated if the fitness of particle i is better than that of former p_{ibest}.

3) Update p_{gbest} if the fitness of all particles' p_{ibest} is superior to that of former p_{gbest}.

Step 4: Go to step 5 if termination condition is met; or else go to step 3.

Step 5: Output p_{gbest} and construe the outcome into solution.

Flow chart of the main procedure is expounded in Fig. 3

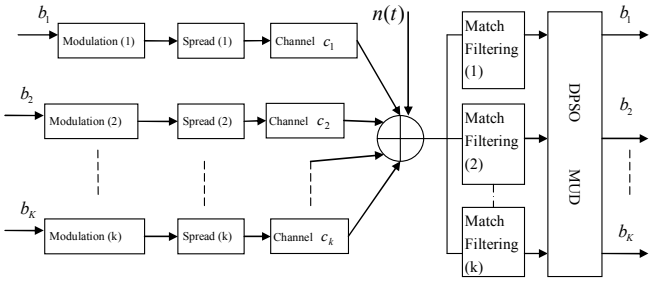


Figure 2. DPSO MUD model

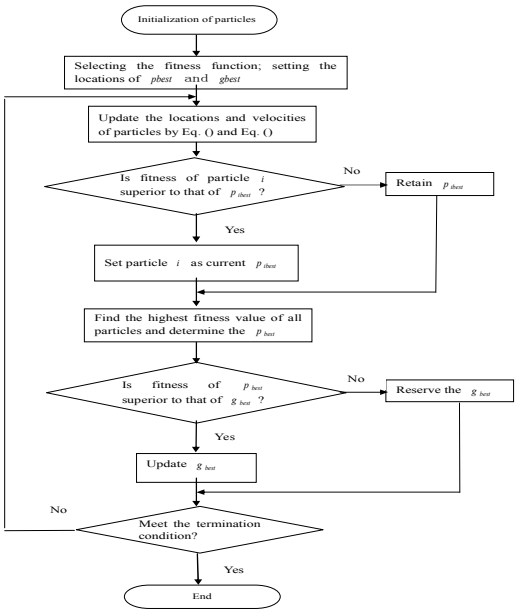


Figure 3. Flow chart of DPSO MUD main procedure

Simulation and analysis

Number of system user $K=10$, number of particles $N=12$, $V_{\max}=[-4,+4]$, $w_{\max}=0.9$, $w_{\min}=0.4$, $c_1=c_2=2$, termination epoch $T=32$. Here Walsh-Hadamard code is employed and we suppose AWGN noise is added in the channel.

We define the average BER in MUD at different SNR as

$$P_e=\frac{1}{10}\sum_{i=1}^{10}P_e^{(i)}. \tag{10}$$

where $P_e^{(i)}$ is the detection error rate of user *i*.

Performance comparison on BER versus SNR. With fixed number of users $K=10$, we let SNR vary from -20dB to 10dB. Fig. 4 and Fig. 5 display the performance of PSO detector and DPSO detector.

At $\mathrm{SNR}\in[-20,5]$, BER performance of DPSO MUD is a bit better than that of basic PSO MUD. Even though basic PSO MUD is roughly comparable with DPSO MUD, however the performance curve of DPSO is smooth and stable without precipitous humps. Moreover the operation of DPSO is conveniently accessible to empirical problems. BER improvement of DPSO system is conspicuously better than PSO MUD when SNR is greater than 5dB. The greater the SNR is, the lower the BER will be in DPSO MUD system. This phenomenon is ubiquitous among most MUDs individually, including basic PSO MUD as well as DPSO MUD. Specifically as for the two, the internal mechanism of PSO algorithm should be attributed to the difference between basic PSO and DPSO. When the SNR is small, as AWGN is a sort of continuous and stochastic noise added into signal, thence particles representing the signal in their location and velocity to search the optimal solution have ambiguous instruction or awareness of where to fly and how to communicate within itself and amidst the swarm, because of the intense interference. Therefore basic PSO and DPSO MUDs have

comparable performance. When Eb/No is great, DPSO imposes *sigmoid* function to convert continuous signal into binary bits which reflects original information. Since the energy of noise is much weaker than signal, errors of the convert by *sigmoid* function are rather statistically less than errors induced by PSO MUD whose particles wander the interval $[-1,+1]$.

BER performance versus number of users. If Eb/No=10dB is fixed, number of users is set to increase from 4 to 16, the BER performance is illustrated in Fig.6 and Fig. 7. Improved PSO MUD transcends basic PSO MUD in BER performance when users amount. Fig. 7 indicates BER performance and near-far effect have been improved remarkably by DPSO MUD, which plaudits its efficacy.

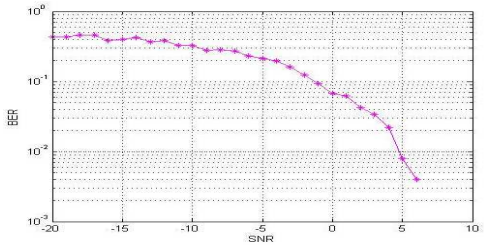


Fig 4. BER vs. SNR curve of Basic PSO MUD

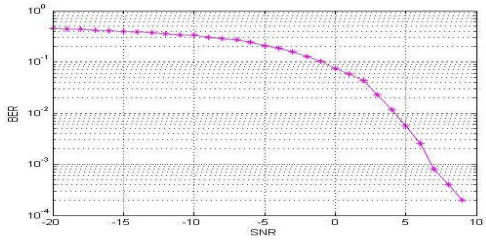


Fig. 5 BER vs. SNR curve of DPSO MUD

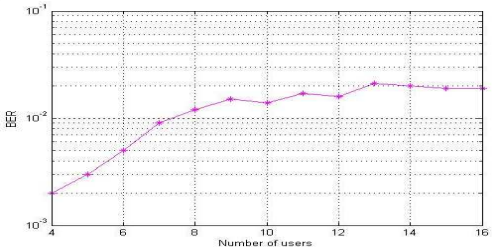


Fig 6. BER vs. user number curve of Basic PSO MUD

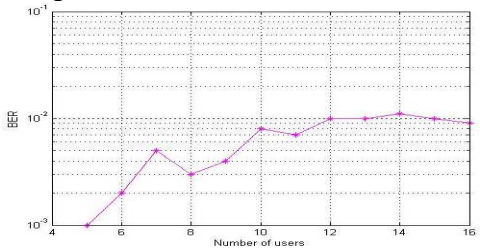


Fig. 7 BER vs. user number curve of DPSO MUD

Summary

The discrete PSO algorithm for multiuser detection has exhibited improved performance in convergent and BER compared with basic PSO MUD. DPSO MUD proves to be practically feasible in CDMA system but it requires much training information. Thus a more effective method needs to be studied so as to conquer intersymbol interference, near-far effect in mechanism, which enjoys a broader realm of application in wireless communication systems.

Acknowledgements

his work is supported by the National Science Foundation of China (No.60972106) and the Science Foundation of Hebei Education Department (No.2008315 and No.2009425).

References

[1] S. Verdu: Multiuser Detection. (Cambridge Univ. Press, Cambridge, U.K. 1998)

[2] L.D. Arya, L.S. Titare and D.P. Kothari: Applied Soft Computing Journal Vol. 9 No. 4 (2009), p. 1197-1207

[3] H. M. Elragal: International Journal of Neural Systems Vol. 19 No. 5 (2009), p 387-393

[4] H.W. Liu and J. Li: IEEE Signal Processing Letters Vol. 15 (2008), p. 29-32

[5] J. Kennedy and R. C. Eberhart: Proc. IEEE ICSMC (1997), p. 4104–4108

[6] S. Chen,W. Yao, H.R. Palally, and L. Hanzo, in: Computational Intel. in Expensive Opti. Prob., ALO 2, edtied by Y. Tenne and C.-K. Goh Springer-Verlag, Berlin Heidelberg (2010)

© (2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.8

Flutter Analysis of Compressor Blade Based on CFD/CSD

Jin Cai[1, a], Qingze Xu[2,b] , Zheng Wang[3,c]

1. Key Laboratory of Contemporary Design and integrated Manufacturing Technology, Ministry of Education, Northwestern Polytechnical University Xi'an 710072,China;

2. Shenyang Aeroengine Research Institute, Shenyang 110015,China;

3. Department of Aeroengines, Northwestern Polytechnical University, Xi'an 710072,China;

[a]lovecai@hotmail.com, [b]xqz780423@chinaren.com, [c]wzh2001@sina.com

Keywords: flutter characteristic; unsteady; aeroelastic responses; compressor stability

Abstract. Based on the CFD/CSD (Computational Fluid Dynamics / Computational Structural Dynamics) algorithm, flutter characteristic and fluid structure interaction (FSI) problems of turbomachinery blades were studied in present paper. The three-dimensional unsteady Navier-Stokes equations and three-dimensional structural model were solved by the finite volume method and the finite element method, respectively. High accuracy in calculation and data exchange were gained by using load transfer, deformation tracking and synchronization between two solvers. The procedure successfully simulated the aeroelastic responses of a high performance fan rotor, NASA Rotor 67, over a range of operational conditions, and the results were compared with the experiment. The results show that the flutter mechanics of the compressor blade could be illustrated based on mean pressure and the distribution of cycle work, which is helpful for the decision of compressor stability.

Introduction

Since last century 90s, with the compressors level loads of the advanced high-bypass pressure ratio turbofan engine were more and more higher, the working conditions of fans and compressors blades were very severe, so it often broke down. In the aeroengine area, vibration failures account for more than 60% of total aeroengine failures, and blade failures more than 70% of vibration failures. The flutter was the mainly reason on the blade failures, which belonged to an important aeroelastic stability problem. It adapt from the self-excited vibration induced by flow, the blade of elastic body produced aerodynamic coupling vibration[1].

The common methods of blade flutter analysis mainly consist of energy method and time-domain analysis method. The energy method evaluated flutter character by aerodynamic force work. The time-domain analysis method makes the fluid-structure-interaction (FSI) calculation for unsteady flow field, so it could define that the vibration displacement stabilized or diverged. The calculation method with FSI, which is coupling calculation of the unsteady aerodynamic force and structure instant deformation, could greatly increase ratio of precision in numerical simulation, especially in transonic domain where had the high nonlinear influence. The method could handle large deformation of the structure body, geometric nonlinear and material nonlinear complicated conditions [2].

With the rapid developments of CFD calculation technique and great improvements of computersperformance, the coupling calculation method of CFD/CSD developed promptly[3]. The paper used cell-centered discretization on multiblock meshes and dual-time stepping time integration with SST model , and the commercial CFD code CFX is coupled with a structure FEA solver, ANSYS, to perform fluid structure interaction (FSI) analysis for the typical turbomachinery blade – NASA Rotor 67, flutter character was researched on this basis, so provided the criterion for the analysis of compressor flutter .

Simulations and Results

NASA Rotor 67 has been recognized as a popular test case for three-dimensional viscous flow prediction procedures [4, 5] because of its detailed experimental data using a laser anemometer [6]. Thus the paper certificate s aeroelastic numerical method with using NASA Rotor 67. The Rotor 67 consists of 22 blades. The blade aspect ratio is 1.56. At the design point, the rotational speed of the rotor is 16,043 rpm, with a tip inlet relative Mach number of 1.38, a total pressure ratio of 1.63 and a mass flow of 33.25 Kg/s.

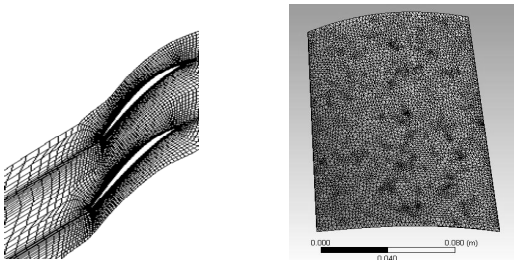


Fig.1 Flow and structure computational grids

The flow and structure computational grid used for the present study is shown in Figure 1 with the blade-to-blade view and one blade view, respectively. The flow mesh consisted of 152091nodal points and 142128 elements in total. The structure element was chosen as solid187， which element thickness matched the true thickness of the blade. And the blade root was fixed. With the computational results of steady fluid as initial condition, the analysis module Flexible Dynamic was solved the blade structure vibracation in the unsteady aerodynamic force. In modeling NASA Rotor 67, an imaginary material whose Young's modules, Poisson's ratio, and density are chosen to be 1.422e+11 Pa, 0.3 and 4539.5 kg/m3, respectively, such that the first natural frequency becomes around 400 Hz.

Flow Calculations

Before the unsteady aerodynamic calculation, it should use the results of the steady aerodynamic calculation as the initial conditions of the unsteady aerodynamic calculation. For simplifying the calculation model, the steady flow model did not consider blade tip clearance and the turbulence model used SST model. The initial conditions and the boundary conditions were given as follows:

- Inlet boundary conditions: the incoming flow was axial intake, the total temperature and the total pressure were given in inlet boundary;
- Outlet boundary conditions: the backpressure was given in outlet boundary;
- Wall boundary conditions: the no-slip wall boundary conditions were used, and the speed of flow mass point in the wall is equal to the speed of the wall. Otherwise, there was adiabatic in the wall boundary conditions;
- Phase contrast period boundary conditions: Phase contrast period boundary is located in the upstream and downstream stage cascade passage.

Figure 2 and Figure 3 are the comparison diagrams of characteristic calculation and experimental results [6]. The compressor prediction performances match well with experimental results in these diagrams, but there is some error in total pressure ratio, as Figure 3 shown. The error creates mainly by shock impact, viscous flow loss and flow separation.

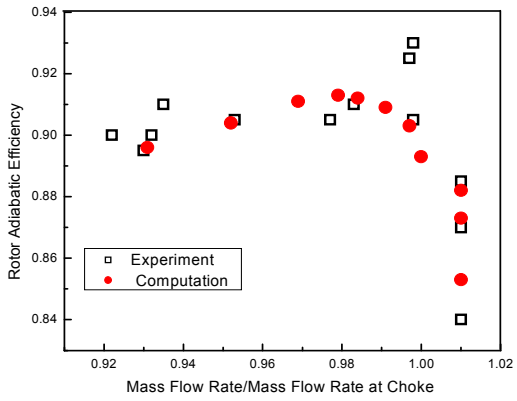


Fig.2 Rotor adiabatic efficiency at design speed

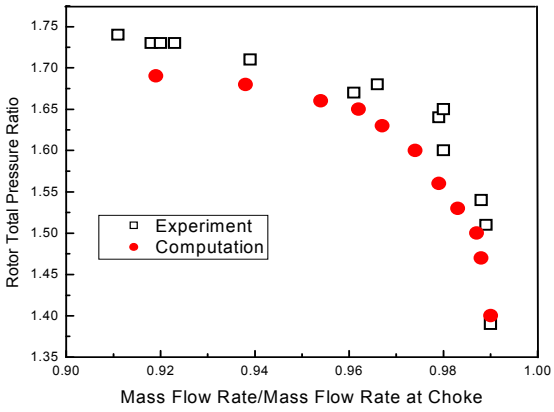


Fig.3 Rotor total pressure ratio at design speed

FSI Calculations

The FSI Calculations need combined with the unsteady aeroelastic results. The vibration effect of the blade is determined by the forced aeroelastic response with initial displacement. The unload blade structure calculates steady calculation results in the preamble. Before the aeroelastic response, the blade can move and power relaxes in the steady load, then the blade deforms in the first modal arrangement. Once the blade gains new convergence results with the initial distortion, it will move and deform response to the aeroelastic effect.

In the every physics time step, the displacement of the structure and the aeroelastic parameter of the flow field would solve with the coupling iteration. At first，the coupling control and the data changing process in the coupling face, and the aeroelastic parameter units are uniformed, then it calculates the aerodynamic force of the element mesh point on the structure surface, and solving Eq.(3) and Eq.(4) for get the deforming value of every point. Using the deforming value on the structure surface as the boundary conditions, the mesh solving process can gain the fluid mesh position of every point and the fluid mesh speed. Finally, the aerodynamic force equation is solved in the new mesh for obtain the update aerodynamic force field.

Stability Decision

The inter-blade phase angle (IBPA) σ is determined by the initial displacement and velocity of the displacement for multi-passage computations. Moreover, the IBPA is forced by the periodic boundary condition. For example, considering the case for σ=180 degree. The computation needs to be performed with two passages where one blade is given an initial displacement and zero initial velocity of the displacement and the other is given the initial displacement in the opposite direction with the same amplitude and zero initial velocity of the displacement as well. In this paper, cases for σ =0 (one passage) and σ =180° (two passages) degrees are examined in order to save computational costs. Four cycles of vibration, which is about 0.016 seconds, were found to be sufficient to distinguish whether the oscillation decays or grows. A total of 36 time steps per vibration period of the first oscillation mode were chosen to ensure the accuracy of the coupled system with the dual time stepping scheme and the Newmark scheme. In per unit physical time step would provide a reasonable time-accurate pressure distribution to the structure.

The time history of the displacements may assess the flutter stability. Figure 4 and Figure 5 show the variation of the circumferential direction displacements of the blade with time for σ = 0° and 180° at the near peak efficiency point.

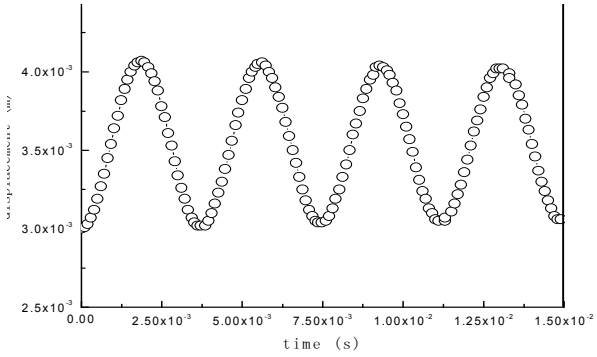


Fig. 4 Oscillation of the blade circumferential in $\sigma=0^{\circ}$

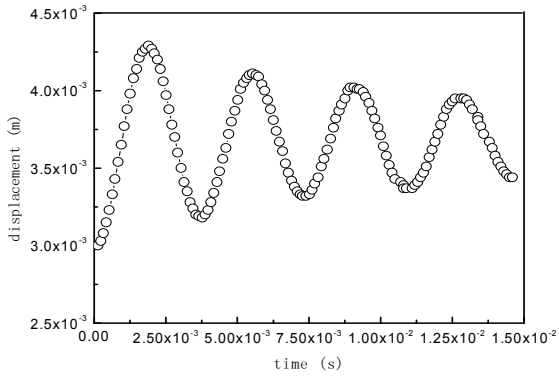


Fig. 5 Oscillation of the blade circumferential deflection in $\sigma=180^{\circ}$

Figure 4 shows almost a typical response with constant amplitude of oscillation. Figure 5 shows decreasing amplitude with each oscillation cycle, indicating decay in the oscillation and flutter free condition. So the aerodynamic damping for $\sigma=180^{\circ}$ is much greater than that for $\sigma=0^{\circ}$.

To examine the unsteady pressure influence on the stability of the blade motion, the work per cycle is a parameter often used to represent the contributions of unsteady pressure fields to the stability. The work per cycle W_c is the work done by the fluid on a given blade over one cycle of its motion. The work per cycle can be shown as follows,

$$W_C = \int_t^{t+T} \int_{\partial\Omega} p(x,y,z,t)\dot{q}(x,y,z,t) \cdot dSdt \tag{10}$$

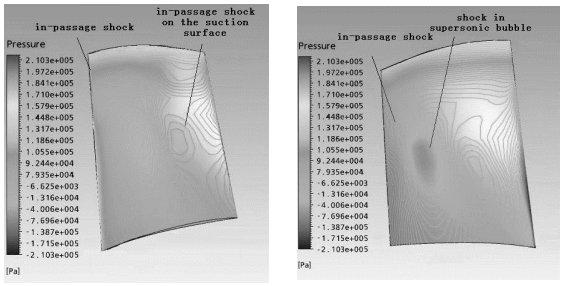


Fig. 6 Mean pressure distribution on the blade surfaces, $\sigma=0^{\circ}$

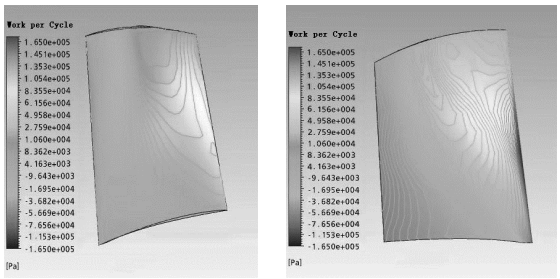


Fig. 7 Work per cycle distribution on the blade surfaces, $\sigma=0^{\circ}$

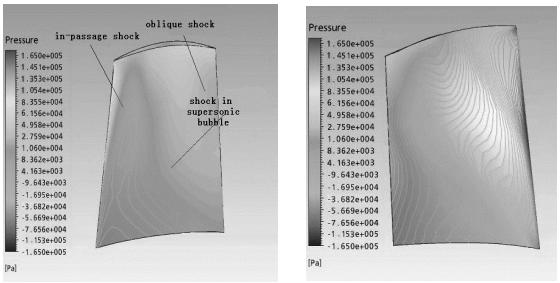


Fig. 8 Mean pressure distribution on the blade surfaces, $\sigma=180^{\circ}$

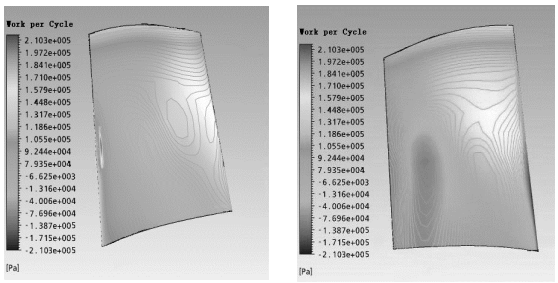


Fig.9 Work per cycle distribution on the blade surfaces, $\sigma=180^{\circ}$

From figure 6 to figure 9, the left figures are the pressure surface and the right figures are the suction surface. Figure 6 and figure 7 show the mean pressure and work per cycle distributions in $\sigma=0^{\circ}$. A peak in the work per cycle exists in the region where the shocks sit on the blade in figure 7. On the suction surface, shock appears continuous curve from the inlet in the span-wise direction.

However, a negative peak in the work per cycle distribution appears only at the location where the lower part of the shock sits on the suction surface. Between the casing and the 50% span, the passage shock is very strong and indicates that an adverse pressure gradient follows that may cause shock induced separation. On the pressure surface, there is a positive peak near the location of the in-passage shock. At the same time, the regions of exciting and damping forces on both surfaces are quite close.

Figure 8 and figure 9 show the mean pressure and work per cycle distributions for σ = 180°. For this case, the two kinds of shocks on the suction surface mentioned above do not form a continuous line. The forces generated by the in-passage shock on the suction surface works to dampen the oscillation, while the value along the oblique shock minishes. It shows obviously that the in-passage shock on the pressure surface produces the large damping forces toward the leading edge. Thus it can be clearly explained the physical mechanism for the decaying response of σ =180°.

Conclusion

A fluid/structure coupled aeroelastic solver for turbomachinery is developed for use in turbomachinery flutter simulations, the solver is based on a three-dimensional unsteady RANS solver, CFX, and the finite element structural analysis package, ANSYS. The high accuracy preserves with loading transform, mesh deformation transferring and synchronization methods in these parts. The solver resolves turbomachinery blade period oscillation in work condition and identifies the compressor stability based on the mean pressure and the related work per cycle. Results present for Rotor 67 reveal that the main contribution to the stability of a transonic fan is the unsteady forces generated by the shock motions. IBPA play the most important role for determining these shock motions. The coupling method of Fluid solver and structure solver such as presented here simulates the aeroelastic response for turbomachinery in work condition and the flutter model, the aeroelastic results match well with the experiment data, consequently, it validates the exactness and the effectivity of the coupling method.

References

[1] Daniela BARAN, Horia PAUN, Sebastian PETRISOR. Fluid-structure interaction, theoretical and practical aspects. SISOM 2006:85-90. (2006).

[2] S. Yigit, M. Schäfer, M. Heck. Grid movement techniques and their influence on laminar fluid – structure interaction computations. Journal of Fluids and Structures[J]. 24 819 – 832.(2008).

[3] Y.Zheng, Computational aeroelasticity with an unstructured grid method. Journal of Aerospace Power. Vol 24,p.2069. (2009)

[4] David A. Johnston, David A. Johnston, J. Mitch Wolff. An architecture for fluid/structure interaction analysis of turbomachinery blade. AIAA 2005-4013. (2005)

[5] X.G.Zheng, F.Liu, C.A.Zhang, Flutter computation of turbomachinery blades based on CFD. Aeronautical Computing Technique. Vol.39. p. 75. (2009).

[6] J.Zhao, Y.W.Liu, B.J.Liu, Frequency domian analysis of NASA67 unsteady flow field.Journal of Aerospace Power.Vol. 22.p.1371. (2007)

doi:10.4028/www.scientific.net/AMM.50-51.13

Code Reuse in Gene Expression Programming

Li Qu[1, 2 a], Yao Min[2,b] ,Wang Weihong[1,c] Du Yanye[1,d]

[1] College of Computer Science, Zhejiang University, China

[2]College of Computer Science, Zhejiang University of Technology, China

[a]liqu@zju.edu.cn, [b]myao@zju.edu.cn, [c]wwh@zjut.edu.cn, [d]du_yanye@163.com

Keywords: Gene expression programming, code reuse, symbolic regression

Abstract. Gene expression programming (GEP) is a kind of phenotype/genotype based evolutionary computation. Code reuse is an important issue in GEP. Various methods are used in current literature to achieve this task. In this paper, we compared six GEP based algorithms by experiments. We proved that although it's possible invent different kinds of code reuse strategies, current available strategies are powerful and efficient.

Introduction

Gene expression programming (GEP)[1] is a kind of phenotype/genotype based evolutionary computation. It uses the linear string as its genotype, on which its genetic operators are performed, and it also uses expression tree as its phenotype, on which the fitness of each individual is evaluated. GEP has been used in various fields, such as symbolic regression, function finding, classification rules mining[2,3,4]. It has gained wide applications to scientific and technical problems for modeling and prediction.

One problem in function finding is that some sub-structures show up repeatedly in the expressions of results, this kind of using gene blocks repeatedly is called code reuse. And this kind of problems is more worth noticing in complex problems. In [5, 6], Koza made the key point that the reuse of code is a critical ingredient to scalable automatic programming. The canonical strategy of code reuse in Genetic Programming is Automatically Defined Functions (ADFs). It has been proved to be useful and effective.

Code reuse is also an important issue in GEP [2]. Basic GEP uses single and multiple gene systems, in order to improve the efficiency and quality of solutions, various code reuse strategies have been proposed, some of the most commonly used are: Automatic Defined Functions (ADFs), Multi layer chromosome, overlapped gene technique and homeotic gene technique. These algorithms have been used to compare with single and multiple gene systems. However, these comparisons lack of sound theoretical foundation. And comparisons of these algorithms are also needed.

In this paper, we compared six GEP algorithms by experiment results. We proved that although we can still invent different kinds of code reuse strategies, current available strategies are powerful and efficient.

Introduction of Gene Expression Programming

Gene Expression Programming [1] uses fix length strings as its genotypes. The genotypes are later expressed as phenotypes, i.e., Expression Trees (ETs) with different sizes and shapes. In basic GEP, a chromosome is composed of predefined number of genes. Each gene is divided into head and tail parts. In the head region, nodes could be elements selected from a function set or a terminal set, while in the tail region, nodes could only be selected from terminal set. The size of the head (h) is determined by the user, and the size of the tail (t) is computed as:

$$t = h*(n-1)+1 \quad (1)$$

considering n the maximum arity in the function set for the particular problem. The gene is translated into an expression tree by following the level order traversal. And the expression tree is easy to be converted into a mathematical expression. The part in the gene not encoded in the expression tree is called non-coding area. When multiple genes in a chromosome are expressed, they are connected by linking functions such us arithmetical functions or Boolean functions. By this way, several ETs can combine to form a bigger ET.

The chromosomes in GEP are subjected to initialization, evaluation, selection, recombination and mutation operators, and the best solution is derived after generations of evolution, as other evolutionary computation techniques do. Usually five components are specified when using GEP to solve a practical problem: the function set, the terminal set that includes problem-specific variables and pre-selected constants, fitness function, control parameters, and stop condition [2].

Various GEP Encoding Methods

Basic Single/Multi Gene GEP and ADF. As GEP mimics the evolution of the nature, it uses genetic operator to introduce diversity into the system [1,2]. The genome or chromosome consists of a linear, symbolic string of fixed length composed of one or more genes. The simplest chromosome of GEP is single gene chromosome. The system which is a little bit more complex and efficient is system with multi-gene chromosome. In multi-gene systems, each gene encodes a sub-ET and the sub-ETs interact with one another to form a more complex entity. They can be efficiently used to evolve good solutions to virtually all kinds of problems.

However, multi-gene systems are not always the most efficient because good parts in the genome are often destructed by genetic operators such as crossover or mutation. And a part which may be very close to the one in the perfect solution or sub-task is discarded and never shows in the evolution again. This may lead to longer evolution or failure of the whole system.

A simple and natural way to solve this problem is a special class of genes - Homeotic genes. Homeotic genes have exactly the same kind of structural organization as conventional genes and they are built using an identical process. In genetic programming, this kind of homeotic gene is called Automatically Defined Functions (ADFs)[2]. In GEP, this name is also used with the same meaning

CHN-GEP. Although ADF is a powerful and simple method of code reuse, it is of course not the only way to achieve this purpose. Code reuse is also possible by just introducing some slight changes in the coding method. One of some effective ways is multi-layer chromosomes GEP (MGEP) or chromosome hierarchy network GEP (CHN-GEP)[7]. The characteristics of CHN-GEP algorithm are as follows:

1. Each individual in the population is composed of many chromosomes.
2. Each chromosome may have different character sets or length.
3. All chromosomes, from the top to the bottom can be divided into many levels, and the gene of the upper level chromosome can call the genes of lower level chromosome.

Multi-Gene Evolutionary Algorithm Based on Overlapped Expression (MEOE). Inspired by the overlap gene expression in biological study, Peng [8] proposed MEOE algorithms to simulate the natural biological evolutionary.

The translation process is as follows:

1. Scan each element of gene in order.
2. If the current symbol belongs to T, then let it be a leaf-node in ET.
3. If the current symbol belongs to F, then let it be a non-leaf node in ET, the number of its sub-trees equals the number of the function parameters. Let the element which is the directly succeeding of current symbol be the first root node of the sub-trees, the secondary element be the root node of secondary sub-trees, and the rest may be deduced by analogy. If it meets the end of the gene, then let the first element in T as sub-tree root-node.

Head+body+tail Based Multilayer GEP. In order to improve the ability of mathematical modeling, decrease the complexity of model, and introduce the mechanism of learning, Kang [9] proposed a new

gene structure. It is composed of three parts: the head with length h come from Function Set (F), the body with length b come from F and Terminal Set (T), and the tail with length t come from T, where h determines the lowest complexity of the program, and h+b determines the highest complexity of the program, and t is the function of h and b:

$$t= (n-1)(h+b)+1 \quad (2)$$

It is the extension of the expression (1). When h = 0, it becomes expression (1). Since its gene is composed of three parts, for the convenience of comparison, we call it M3GEP.

By this means, the multi-gene chromosome structure is extended too, so that the mechanism of learning can be introduced. Extends the terminal set of homeotic gene, its terminal set $T_g^*=\{T_g,T\}$ not only includes the terminal set T_g of the former homeotic gene but also includes the terminal set T of traditional gene, so the expression ability of main program is improved and then the searching ability of the algorithm is improved. Multi-level homeotic gene structure constructs multi-cellular chromosome. However, the terminal set of the root homeotic gene: $T_c^*=\{T_c,T_g\}$ which not only includes the elements set $T_c=\{c_1, c_2\}$ of the cell but also includes their terminal set T_g^*.

Experiment Results and Analysis

Experiment Settings. In order to compare the performance of the six algorithms, we tested them on two typical symbolic regression problems. The first is a relatively easy problem $f(a)=\frac{a^2}{2}+3a$, the second is a fairly high order function, i.e., the widely used sextic (order of 6) polynomial $f(a)=a^6-2a^4+a^2$. The two functions are widely used in evolutionary computation performance evaluation. In order to be fair, we use same parameters for all algorithms, and the total length of each algorithm is set to be equal to 51 so as to compare the code reuse and problem solving ability of all the algorithms. We choose a fairly small population size of 60 and a fairly small maximum generation of 1000 for all algorithms because these settings are big enough for at least one algorithm to find the target function. And the one which is able to find the target function can be used as the baseline of performance comparison. For genetic operators, we follow the settings of [1], we use the same settings for all algorithms except some operators are only designed for a specific algorithm. Detailed settings are listed in Table 1.

Table1. Experimental settings

	Single Gene	Multi Gene	ADF	CHN-GEP	MEOE	*M3GEP*
Function set	+ - * / Q	+ - * / Q	+ - * / Q	+-*/ Q	+ - * / Q	*+ - * / Q*
Number of chromosome	1	1	1	3	1	*3*
Linking function	+	+	NULL	+	+	+
Number of gene	1	3	3	(3, 3, 1)	3	*(3, 3, 1)*
Head length	25	8	5	(3, 3, 4)	-	*(3, 3, 4)*
Body length	-	-	-	-	17	-
Chromosome recomb. rate	-	-	-	0.1	-	*0.1*
Gene random recomb. rate	-	-	-	0.1	-	*0.1*
Precision	*0.001*	*0.001*	*0.001*	*0.001*	*0.001*	*0.001*

Experiment Results. The first experiment is a relatively easy problem. From the experiment results we can see that Multi Gene GEP gets the highest success rate and average runtime. While, as we look into the average generation of each success regression, ADF is the fastest to find the target function.

There may be two reasons for the results, on one hand, it is a function with two coefficients which need to be generate with complex replacement expressions in the system without RNCs. However, in

this situation, ADF still has the shortest average expression length, that means for the same problem, ADF needs the least gene to encode the solutions.

Although the problem seems to be simple and easy to find, the success rate of each algorithm is not satisfactory. This can also be explained in two ways, one is that this is not a typical problem for code reuse, the problem can be solved with multi-gene system. The other is that random numbers are necessary for problems with Coefficients.

Table 2. Experimental Results of Test Problem 1

	Single Gene	Multi Gene	ADF	CHN-GEP	MEOE	M3GEP
Avg. Exp. Length	31	26	20	25	36	24
Success Rate	0.37	0.67	0.28	0.52	0.24	0.49
Avg. Gen.	142.03	183.3	117.37	156.54	125.96	145.79
Max. Gen.	981	916	906	888	999	998
Min. Gen.	77	9	45	3	27	2
Avg. Runtime(s)	1.50063	1.34671	4.75063	3.96641	19.12907	5.54312

The second experiment is also a typical example of code reuse. Its simplification is $(a-1)^2 * a^2$, if there are gene block of a and $a-1$ which can be called repeatedly, then the search process would be much faster than just get the high order expression with multiplications.

Experiments show that except for MEOE, other algorithms are all able to solve the code reuse tasks. The success rates of function finding are all close to 100%. And on this problem, we can see that ADF has the least maximum success generation, minimum success generation and average runtime. On this problem, GEP is very close to GP. What's more, ADF find the most parsimonious solutions.

This problem enhances the importance of code reuse in high order function finding. And given limited length of gene expressions, the algorithm would not be able to find the solutions without code reuse, thus causing poor success rates. However, ADF, CHN-GEP and M3GEP are all able to solve this problem perfectly.

Table 3. Experimental Results of Test Problem 2

	Single Gene	Multi Gene	ADF	CHN-GEP	MEOE	M3GEP
Avg. Exp. Length	36	33	17	22	43	16
Success Rate	0.59	0.19	0.98	0.98	0.14	0.99
Avg. Gen.	185.98	99.74	73.66	108.15	68.81	134.74
Max. Gen.	983	989	553	856	812	879
Min. Gen.	9	101	1	6	113	3
Avg. Runtime(s)	2.2775	3.61313	0.78219	1.27063	12.27955	2.0386

From the experiments, we can see that in the four strategy of code reuse, ADF, CHN-GEP and M3GEP can get best success rate on different problems and no one algorithm outperforms others on runtime. And MEOE always gets the lowest success rate and the longest runtime.

On express ability, it is clear that ADF has the most powerful express ability, it gets minimum expression length on both problems, and especially on the third experiment its average length, along with M3GEP is far shorter than other algorithms.

Generally speaking, multi gene system is able to solve simple and medium complicate problems with some efficiency, its code reuse ability is limited. While for ADF, CHN-GEP and M3GEP, each strategy will show some specific trait. The three strategies are able to solve difficult, especially high order problems. However, MEOE is highly affected by its time cost of code reuse.

Conclusion

Code reuse is an important and useful tool for evolutionary computation techniques, especially gene expression programming to solve regression problems as well as other learning tasks. In this paper,

we've compared the available six algorithms with different strategies by mathematical proof and experiments. Performance and mathematical proof reveal that code reuse is useful in many regression problems.

It is clear that although multi gene GEP system is generally efficient and enough for simple problems, it is not good enough to solve hard problems with limited chromosome length and generations. ADF, CHN-GEP and M3GEP are three similar algorithms with hierarchical structure, they are easy to implement and effective for many regression problems. Maximum express space is reached only when all the nodes in the genotype are functions instead of terminals, and that case doesn't always happen. So it is not a convincing method to evaluate the expected performance of algorithms.

Future work should be more detailed mathematical proof of all the algorithms in the same frame and more experiments to verify the proof. Other kinds of code reuse methods such as Prefix GEP [3] also worth studying.

Acknowledgement

The work was supported by National Important Science & Technology Specific Projects under grant 2009ZX01038-001-06, National High Technology Research and Development Program of China (863 Program) under grant 2009AA011907 and 2008AA01Z132.

References

[1] C. Ferreira: Gene Expression Programming: A New Adaptive Algorithm for Solving Problems, Complex Systems, vol. 13, issue 2 (2001), p. 87-129.

[2] C. Ferreira: Automatically Defined Functions in Gene Expression Programming, In Genetic Systems Programming: Theory and Experiences , Studies in Computational Intelligence, Vol. 13, p. 21-56, Springer-Verlag (2006)

[3] L. Xin, C. Zhou, W. Xiao, P. C. Nelson: Prefix Gene Expression Programming. Proc. Genetic and Evolutionary Computation Conference, Washington, D.C., USA(2005)

[4] Zhou, C.; Xiao, W.; Nelson, P. C.; and Tirpak, T. M: Evolving Accurate and Compact Classification Rules with Gene Expression Programming. IEEE Transactions on Evolutionary Computation Vol.7(6): 519-531(2003)

[5] J. R. Koza: Genetic Programming. Cambridge, MA: MIT Press(1992)

[6] J. R. Koza: Genetic Programming II. Cambridge, MA: MIT Press(1994)

[7] Jing Peng, Chang-jie Tang, Chuan. Li, Jianjun Hu: A New Evolutionary Algorithm Based on Chromosome Hierarchy Network, In International Journal of Computers and Applications, Vol. 30, No. 3, p.183-191(2008)

[8] Jing Peng, Chang-jie Tang, Jing Zhang, and Chang-an Yuan: Evolutionary Algorithm Based on Overlapped Gene Expression, In Proc. ICNC 2005, LNCS 3612, p. 194-204(2005)

[9] Zhuo Kang, Yan Li, Li-shan Kang: Automatic Programming Methodology for Program Reuse, in Proc. International Conference on Computational Intelligence and Security, Guangzhou. pp. 208-214(2006)

© (2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.18

A derivative-free filled function method for non-linear inverse optimization of material parameters

LV Yinghui[1,a], Yanzhen YU[2,b]

[1,2]School of Civil and Architectural Engineering; University of Jinan;
Jinan 250022, China
[1]lvyinghui_2010@163.com, [2]cea_yuyz@ujn.edu.cn

Key words: non-linear analysis; filled function; derivative-free; dynamic;

Abstract: In solving non-linear inverse optimization problems, how to make sure one and only solution and cut down computational effort are two very representative challenges. As thus, the derivative-free filled function method is proposed, it couples the classical filled method theory and dynamic canonical descent method. The application results of non-linear inverse analysis of material parameters for a practical engineering show that the derivative-free filled function method can quickly solve non-linear inverse problems with one and only solution, and it will be robust one in non-linear inverse solution field.

Introduction

In recent years, it is very popular to solve nonlinear inverse analysis by optimization method, if we let n be the dimension of inverse analysis problem and f be a function from R^n to R, and the following optimization model(p) is often expressed as an inverse analysis problem:

$$\left\{\begin{array}{l}\min f(x)=\sum_{i=1}^{m}[u_i(x_1,x_2,\cdots,x_n)-u_i^*]^2\\ s.t.a_i\le x_i\le b_i, i=1,2,3,\cdots,n\end{array}\right\} \quad \text{(p)}$$

Where m is the number of monitoring points, u_i and u_i^* are respectively computable displacement and monitored displacement at one monitoring point, a_i and b_i express respectively lower and upper bound to x_i.

Considering nonlinear inverse optimization analysis, how to make sure one and only solution and cut down computational effort are two main challenges. So far, an expressive number of methods with global optimization capability have been tried to get the one and only solution, such as genetic algorithm (GA) [3], simulated annealing (SA) [4], particle swarm optimization (PSO) [8] and so on. However, these heuristic algorithms often have slow optimization speed even for not so complicated problems and limited capabilities of global search.

Because of that, a new optimization method is proposed to solve nonlinear inverse analysis problems, which integrates filled optimization method with dynamic canonical descent optimization method for global convergence and no differential coefficient.

The derivative-free filled function method

Firstly, the above optimization problem (P) will be translated into the following:

$$\left\{\begin{array}{l}\min U(x)\\ s.t.x\in X, X=\{x\mid a_i\le x_i\le b_i\}\end{array}\right\} \quad \text{(AP)}$$

Where $U(x)$ is a considered commonly filled function.

The derivative-free filled function method is applied to acquire solution of optimization problem (AP), the acquired solution is just the solution of optimization problem (P). The optimization proceedings of new proposed method can be expressed as the following:

Step0: Set an initial guess x^0 in bound-constrained district of model (P).

Step1: an ordinary minimization optimization algorithm is used to solve model (P) for a minimization point $x^* \in X$.

Step2: update x^0 on the ground of $x^0 = x^*$.

Step3: calculate $f(x^*)$.

Step4: if $|f(x^*)| \le \varepsilon$ (ε is a too much small positive number), go to Step11.

Step5: construct filled function $U(x)$ at x^*.

Step6: let updated x^0 be an initial guess, Call dynamic canonical descent method to solve model (AP) to get a minimization point x^2.

Step7: if $|f(x^2)| \le \varepsilon$, let $x^* = x^2$, then go to Step11

Step8: update x^0 on the ground of $x^0 = x^2$.

Step9: if $|f(x^2)| > |f(x^*)|$, go to Step6.

Step10: let updated x^0 be an initial guess, then go to Step1.

Step11: stop the proceedings, the solution of model (P) is x^*

In the above optimization proceedings, the dynamic canonical descent method, initially proposed by Bousson[7], is involved and is also described in detail in the article[1] and needs no repetition here. The argument in favor of the superiority of the dynamic canonical descent algorithm over other algorithms is that the minimization process is based on a high exploration search[1] with the good capability of global optimization[1].

Identification of parameters in elastic-plastic inverse analysis problem

In this section, the new proposed algorithm is applied to identify mechanical parameters about one side slope. Fig1 shows the structural sketch of side slope with the chosen measuring points, the density of rock slope is $2600\,Kg/m^3$. In the process of inverse optimization, the monitored displacements at every monitoring point are chosen as the foregone information. In addition, the software ABAQUS is brought into the inverse optimization for computable displacements at every measuring point with Mohr-Coulomb strength criterion. Fig2 shows the compartmentalized mesh of rock slope and the inversely analyzed parameters are the deformation modulus E in $[0.1, 20]\times 10^9\,Pa$, the Possion ration ν in $[0.15, 0.35]$, the cohesion c in $[0.1, 20]\times 10^3\,Pa$ and the

internal friction angle ϕ in$[10^0, 35^0]$. Finally, the proposed new algorithm gives the four parameters as such: $E = 3.0\text{GPa}$, $\nu = 0.25$,

$c = 2.0 \times 10^3 \text{Pa}$, $\phi = 25^0$.

Fig3 gives the deformation map about rock slope at the end of optimization and Fig4 shows the comparison between inverse analysis displacements and measured displacements at every measuring point. From Fig4, it can be seen that there only exists a small difference, so it can be concluded that the inversely analyzed parameters reasonably gives the mechanical properties of rock slope.

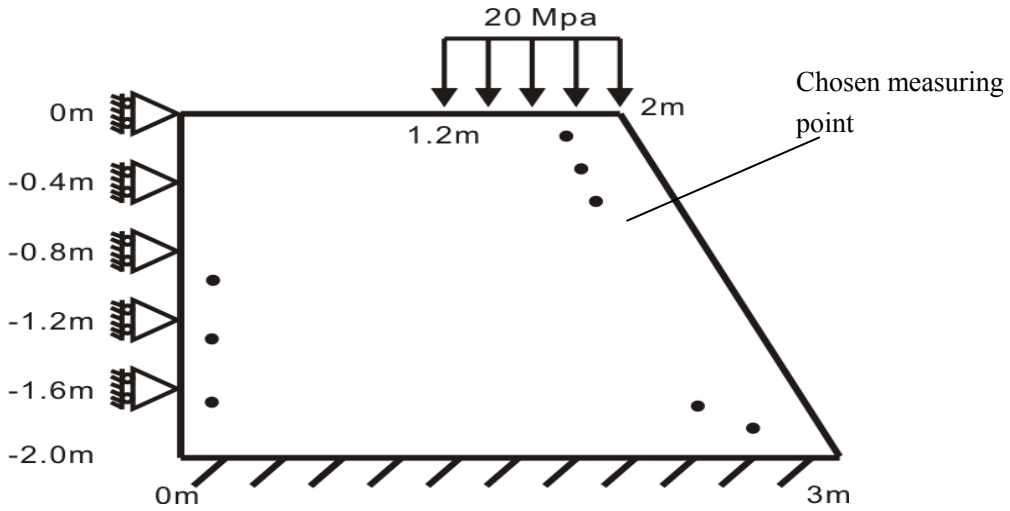


Fig1. The structural sketch of rock slope

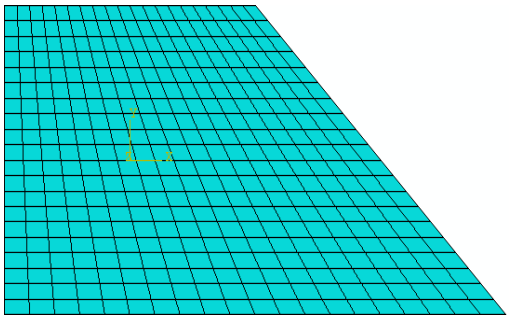


Fig2. The compartmentalized mesh of rock slope

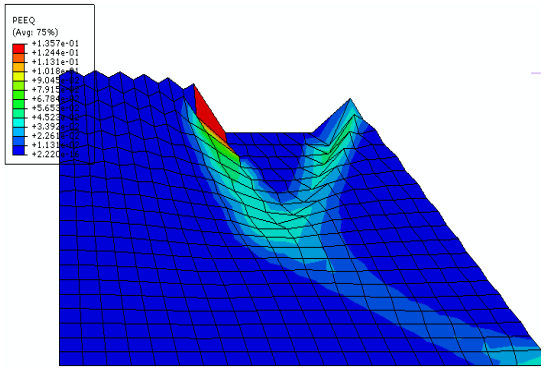


Fig3. The equivalent plastic deformation map

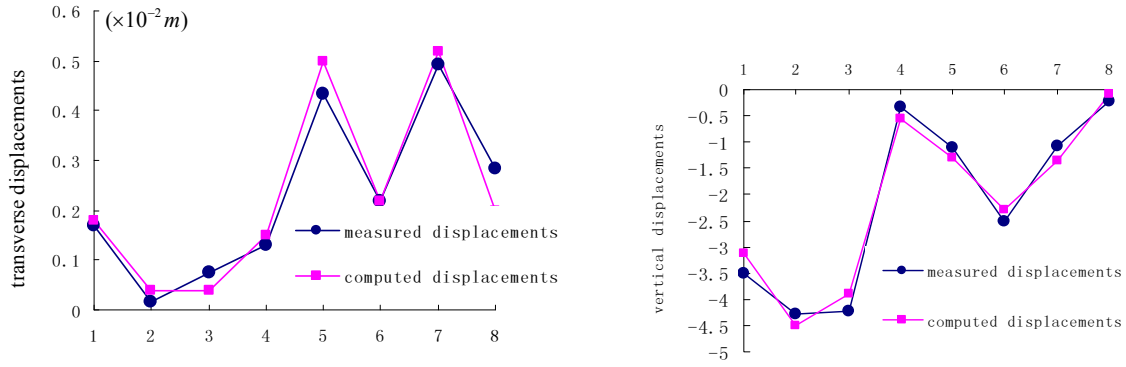


Fig.4. the comparison from inverse analysis displacements and measured displacements

Concluding remarks

This paper proposes to integrate the filled function method and dynamic canonical descent method for no differential coefficient in optimization. In the elastic-plastic inverse displacement analysis, the new proposed algorithm is introduced to give mechanical properties of one rock slope, the optimization results show that the new method can quickly converge to a reasonable point and gives exact inverse analysis parameters, so it can be concluded that the new proposed algorithm should be robust one in solving nonlinear inverse analysis problems

Acknowledgments

The research presented in the paper is supported by the Project of Natural Science Foundation for the Youth from People's Republic of China (NO.10902118).

Reference

[1] K.Bousson, S.D.Correia. Optimization algorithm based on densification and dynamic canonical descent. Journal of Computational and Applied mathematics, 191(2006), 269-279.

[2] K.Bousson. Efficient global optimization based on dynamic canonical descent. Systems Sci. 26(4) (2000), 61-78.

[3] D.E.Goldberg. Genetic Algorithm in Search, Optimization&Machine Learning, Addison-Wesley, 1989.

[4] S.Kirpatrick, C.D.Gelatt, M.P.Vecchi. Optimization by simulate annealing. Science,220 (1983),671-680.

[5] Zafer Bingul. Adaptive genetic algorithms applied to dynamic multiobjective problems. Applied Soft Computing.7 (2007), 791-799.

[6] Guang-yu He, Yong-li Gao: The aggregation of the common numerical algorithm in VisualFortran (Science Press, China 2002)

[7] K. Bousson, Efficient global optimization based on dynamic canonical descent. Systems Sci26(2000), 61-78

[8] J.Kennedy, R.C.Eberhart. Particale Swarm Optimization Proceedings of the 1995 IEEE International Conference on Neural Networks, Vol.4, IEEE Press, 1995, pp. 1942-1948.

[9] Nyssen C. An efficient and accurate iterative method allowing large incremental steps to solve elasto-plastic problems[J]. Comput.Struct., 1981,13(1-3):63-71.

doi:10.4028/www.scientific.net/AMM.50-51.22

On Stability of Neutral Systems with Time-Varying Delay

Meilan Tang[a], Xinge Liu[b]

School of Mathematical Science and Computing Technology, Central South University

Changsha, 410083, Hunan, China

[a]tangmeilang@163.com, [b]liuxgliuhua@163.com

Keywords: Neutral system; Time-varying discrete delay; Neutral delay; Linear matrix inequality

Abstract. This paper investigates the delay-dependent robust stability of neutral systems with time-varying discrete delays and time-varying structured uncertainties. New delay-dependent stability criteria are obtained and formulated in the form of a linear matrix inequality. Since the criteria take the sizes of the neutral delay, discrete delay and derivative of discrete delay into account, they are less conservative than previous methods. Numerical example is given to indicate significant improvements over some existing results.

Introduction

The problem of stability of neutral systems has received considerable attention in the last two decades[1-8]. This paper aims to present a new simple delay-dependent stability criterion for neutral systems with time-varying delay. Based on the relationship among all states and derivatives of all states and linear matrix inequality, a new both neutral-delay-dependent and discrete-delay-dependent criterion is established with less conservative than the previous methods for mixed delays. The criterion obtained is then extended to a neutral system with time-varying uncertainties. Finally, A numerical example illustrates the improvement the proposed method provides over previous methods.

Main result

Consider the following neutral system with a time-varying delay in the state:

$$\begin{cases} \dot{x}(t) - C\dot{x}(t-\tau_2) = (A+\Delta A)x(t) + (B+\Delta B)x(t-d(t)), \\ x(t) = \phi(t), \ t \in [-\tau, 0], \end{cases} \tag{1}$$

where $x(t) \in R^n$ is the state vector, and A, B are system matrices with compatible dimensions. The spectrum radius of the matrix $C, \rho(C)$, satisfies $\rho(C) < 1$.The time delay $d(t)$ is a time-varying continuous function satisfying

$$0 \le d(t) \le d, \quad d(t) \le \mu < 1, \tag{2}$$

where d and μ are constants, $\tau = \max(\tau_2, d), \phi(t)$ is a continuous vector-valued function of $t \in [-\tau, 0]$. The time-varying structured uncertainties are of the form

$$[\Delta A(t), \Delta B(t)] = LF(t)[E_a, E_b] \ , \tag{3}$$

where L, E_a, E_b are constant matices with appropriate dimensions. $F(t)$ is an unknown and possibly time-varying real matrix with Lebesgue measurable elements and its Euclidean norm satisfies

$$\| F(t) \| \le 1, \ \forall t \in R. \tag{4}$$

First, consider the nominal system of system (1)

$$\begin{aligned} &\dot{x}(t) - C\dot{x}(t-\tau_2) = Ax(t) + Bx(t-d(t)), \\ &x(t) = \phi(t), \quad t \in [-\tau, 0]. \end{aligned} \tag{5}$$

The operator G is defined to be $Gx_t = x(t) - Cx(t-\tau_2)$

Definition 1 The operator G is said to be stable if the zero solution of the homogeneous difference equation

$$Gx_t = 0,\ t \geq 0,\ x_0 = \psi \in \{\phi \in C([-\tau_2, 0]) : G\phi = 0\}$$

is uniformly asymptotically stable.

Theorem 1 For given scalars τ_2, d and μ, the neutral system (5) is asymptotically stable for any delay $d(t)$ satisfying Condition (2) if there exist positive definite matrices $P > 0, Q_1 > 0, Q_2 > 0$, $Z > 0, W > 0, R > 0$ and any appropriately dimensioned matrices $Y_1, Y_2, Y_3, Y_4, T_1, T_2, T_3, T_4$ such that the following linear matrix inequality (LMI) holds.

$$\bar{\Omega} < 0, \tag{6}$$

where

$$\bar{\Omega} = \begin{bmatrix} \bar{\Omega}_{11} & \bar{\Omega}_{12} & \bar{\Omega}_{13} & Y_4^T + T_4^T & -dY_1 & -\tau_2 T_1 & A^T S \\ * & \bar{\Omega}_{22} & \bar{\Omega}_{23} & -Y_4^T & -dY_2 & -\tau_2 T_2 & B^T S \\ * & * & \bar{\Omega}_{33} & -T_4^T & -dY_3 & -\tau_2 T_3 & 0 \\ * & * & * & -R & -dY_4 & -\tau_2 T_4 & C^T S \\ * & * & * & * & -dZ & 0 & 0 \\ * & * & * & * & * & -\tau_2 W & 0 \\ * & * & * & * & * & * & -S \end{bmatrix},$$

$$\bar{\Omega}_{11} = PA + A^T P + Y_1 + Y_1^T + Q_1 + Q_2 + T_1 + T_1^T, \quad \bar{\Omega}_{12} = -Y_1 + PB + Y_2^T + T_2^T, \quad \bar{\Omega}_{22} = -Y_2 - Y_2^T - (1-\mu)Q_1,$$
$$\bar{\Omega}_{13} = Y_3^T - A^T PC - T_1 + T_3^T, \quad \bar{\Omega}_{23} = -Y_3^T - B^T PC - T_2, \quad \bar{\Omega}_{33} = -Q_2 - T_3 - T_3^T.$$

Proof Choose the Lyapunov functional candidate for system (5) with time-varying delay as

$$V(x_t) = V_1(x_t) + V_2(x_t) + V_3(x_t) + V_4(x_t) + V_5(x_t) + V_6(x_t), \tag{7}$$

$$V_1(x_t) = \left(Gx_t\right)^T PGx_t, \tag{8}$$

$$V_2(x_t) = \int_{-d}^{0}\int_{t+s}^{t} \dot{x}^T(\alpha) Z \dot{x}(\alpha) d\alpha ds, \tag{9}$$

$$V_3(x_t) = \int_{t-d(t)}^{t} x^T(s) Q_1 x(s) ds, \tag{10}$$

$$V_4(x_t) = \int_{-\tau_2}^{0}\int_{t+s}^{t} \dot{x}^T(\alpha) W \dot{x}(\alpha) d\alpha ds, \tag{11}$$

$$V_5(x_t) = \int_{t-\tau_2}^{t} x^T(s) Q_2 x(s) ds, \tag{12}$$

$$V_6(x_t) = \int_{t-\tau_2}^{t} \dot{x}^T(s) R \dot{x}(s) ds, \tag{13}$$

where $x_t = x(t+\theta), -2\tau \leq \theta \leq 0$ and matrices $P > 0, Q_1 > 0, Q_2 > 0$ $Z > 0, W > 0, R > 0$.

Calculate the derivative of $V(x_t)$ along the solution of system (5). Then, by the Newton-Leibniz formula, we have

$$x(t-d(t)) = x(t) - \int_{t-d(t)}^{t} \dot{x}(s) ds.$$

$$\begin{aligned} \dot{V}_1(x_t) &= 2Gx_t^T(t)P[Ax(t) + Bx(t-d(t))] \\ &= x^T(t)[2PA + 2Y_1]x(t) + x^T(t)[-2Y_1 + 2PB + 2Y_2^T]x(t-d(t)) + x^T(t)[-2Y_3^T - 2A^T PC]x(t-\tau_2) \\ &\quad + x^T(t)2Y_4^T \dot{x}(t-\tau_2) - 2x^T(t-d(t))[-2Y_3^T - 2B^T PC]x(t-\tau_2) - 2x^T(t-d(t))Y_4^T \dot{x}(t-\tau_2) \\ &\quad - 2\int_{t-d(t)}^{t} \left\{x^T(s)Y_1 + x^T(t-d(t))Y_2 + x^T(t-\tau_2)Y_3 + \dot{x}^T(t-\tau_2)Y_4\right\} \dot{x}(s) ds. \end{aligned} \tag{14}$$

$$\dot{V}_2(x_t) + \dot{V}_3(x_t) = d\dot{x}^T(t) Z \dot{x}(t) - \int_{t-d}^{t} \dot{x}^T(s) Z \dot{x}(s) ds + x^T(t) Q_1 x(t) - (1-\dot{d}(t)) x^T(t-d(t)) Q_1 x(t-d(t))$$

$$\le x^T(t)Q_1x(t)-(1-\mu)x^T(t-d(t))Q_1x(t-d(t))+d\dot{x}^T(t)Z\dot{x}(t)-\int_{t-d(t)}^{t}\dot{x}^T(s)Z\dot{x}(s)ds$$
$$=x^T(t)[Q_1+dA^TZA]x(t)+x^T(t)[2dA^TZB]x(t-d(t))+x^T(t)[2dA^TZC]\dot{x}(t-\tau_2)$$
$$+x^T(t-d(t))[dB^TZB-(1-\mu)Q_1]x(t-d(t))+\dot{x}^T(t-d(t))[2dB^TZC]\dot{x}(t-\tau_2)$$
$$+\dot{x}^T(t-\tau_2)[dC^TZC]\dot{x}(t-\tau_2)-\int_{t-d(t)}^{t}\dot{x}^T(s)Z\dot{x}(s)ds. \quad (15)$$

$$\dot{V}_4(x_t)+\dot{V}_5(x_t)=\tau_2\dot{x}^T(t)W\dot{x}(t)-\int_{t-\tau_2}^{t}\dot{x}^T(s)W\dot{x}(s)ds+x^T(t)Q_2x(t)-x^T(t-\tau_2)Q_2x(t-\tau_2)$$
$$=x^T(t)[Q_2+\tau_2A^TWA]x(t)+x^T(t)[2\tau_2A^TWB]x(t-d(t))+x^T(t)[2\tau_2A^TWC]\dot{x}(t-\tau_2)$$
$$+x^T(t-d(t))[\tau_2B^TWB]x(t-d(t))+x^T(t-d(t))[2\tau_2B^TWC]\dot{x}(t-\tau_2)$$
$$+x^T(t-\tau_2)[-Q_2]x(t-\tau_2)+\dot{x}^T(t-\tau_2)[\tau_2C^TWC]\dot{x}(t-\tau_2)-\int_{t-\tau_2}^{t}\dot{x}^T(\theta)W\dot{x}(\theta)d\theta. \quad (16)$$

$$\dot{V}_6(x_t)=\dot{x}^T(t)R\dot{x}(t)-\dot{x}^T(t-\tau_2)R\dot{x}^T(t-\tau_2)$$
$$=x^T(t)A^TRAx(t)+2x^T(t)A^TRBx(t-d(t))+2x^T(t)A^TRC\dot{x}^T(t-\tau_2)+x^T(t-d(t))B^TRBx(t-d(t))$$
$$+2x^T(t-d(t))B^TRC\dot{x}(t-\tau_2)+\dot{x}^T(t-\tau_2)[C^TRC-R]\dot{x}^T(t-\tau_2). \quad (17)$$

$$2x^T(t)T_1x(t)+2x^T(t)T_2^Tx(t-d(t))+x^T(t)[-2T_1+2T_3^T]x(t-\tau_2)+2x^T(t)T_4^T\dot{x}(t-\tau_2)$$
$$-2x^T(t-d(t))T_2x(t-\tau_2)-2x^T(t-\tau_2)T_3x(t-\tau_2)-2x^T(t-\tau_2)T_4^T\dot{x}(t-\tau_2)$$
$$-\int_{t-\tau_2}^{t}\left\{2x^T(t)T_1\dot{x}(\theta)+2x^T(t-d(t))T_2\dot{x}(\theta)+2x^T(t-\tau_2)T_3\dot{x}(\theta)+2\dot{x}^T(t-\tau_2)T_4\dot{x}(\theta)\right\}d\theta=0. \quad (18)$$

Denote $S=R+dz+\tau_2W$, then

$$\dot{V}(x_t)\le\frac{1}{d(t)\tau_2}\int_{t-d(t)}^{t}ds\int_{t-\tau_2}^{t}\{x^T(t)[2PA+2Y_1+Q_1+Q_2+A^TSA+2T_1]x(t)$$
$$+x^T(t)[-2Y_1+2PB+2Y_2^T+2A^TSB+2T_2^T]x(t-d(t))$$
$$+x^T(t)[2Y_3^T-2A^TPC-2T_1+2T_3^T]x(t-\tau_2)+x^T(t)[2Y_4^T+2A^TSC+2T_4^T]\dot{x}(t-\tau_2)$$
$$+x^T(t-d(t))[-2Y_2-(1-\mu)Q_1+B^TSB]x(t-d(t))$$
$$+x^T(t-d(t))[-2Y_3^T-2B^TPC-2T_2]x(t-\tau_2)$$
$$+x^T(t-d(t))[-2Y_4^T+2B^TSC]\dot{x}(t-\tau_2)+x^T(t-\tau_2)[-Q_2-2T_3]x(t-\tau_2)$$
$$-2x^T(t-\tau_2)T_4^T\dot{x}(t-\tau_2)+\dot{x}^T(t-\tau_2)[C^TSC-R]\dot{x}(t-\tau_2)$$
$$-2x^T(t)Y_1d(t)\dot{x}(s)-2x^T(t-d(t))Y_2d(t)\dot{x}(s)-2x^T(t-\tau_2)Y_3d(t)\dot{x}(s)$$
$$-\frac{d(t)}{d}\dot{x}^T(s)Zd(t)\dot{x}(s)-2x^T(t)\tau_2T_1\dot{x}(\theta)-2x^T(t-d(t))\tau_2T_2\dot{x}(\theta)$$
$$-2x^T(t-\tau_2)\tau_2T_3\dot{x}(\theta)-2\dot{x}^T(t-\tau_2)\tau_2T_4\dot{x}(\theta)-\dot{x}^T(\theta)\tau_2W\dot{x}(\theta)\}d\theta$$
$$=\frac{1}{d(t)\tau_2}\int_{t-d(t)}^{t}ds\int_{t-\tau_2}^{t}\xi^T(t,s)\Omega\xi(t,s)ds, \quad (19)$$

where $\xi^T(t,s)=[x^T(t),x^T(t-d(t)),x^T(t-\tau_2),\dot{x}^T(t-\tau_2),d(t)\dot{x}^T(s),\dot{x}^T(\theta)]$,

$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} & Y_4^T + A^T SC + T_4^T & -Y_1 & -\tau_2 T_1 \\ * & \Omega_{22} & \Omega_{23} & -Y_4^T + B^T SC & -Y_2 & -\tau_2 T_2 \\ * & * & \Omega_{33} & -T_4^T & -Y_3 & -\tau_2 T_3 \\ * & * & * & C^T SC - R & -Y_4 & -\tau_2 T_4 \\ * & * & * & * & -\frac{1}{d} Z & 0 \\ * & * & * & * & * & -\tau_2 W \end{bmatrix}, \tag{20}$$

$\Omega_{11} = PA + A^T P + Y_1 + Y_1^T + Q_1 + Q_2 + + A^T SA + T_1 + T_1^T$, $\Omega_{12} = -Y_1 + PB + Y_2^T + A^T SB + T_2^T$,

$\Omega_{22} = -Y_2 - Y_2^T - (1-\mu)Q_1 + B^T SB$, $\Omega_{13} = Y_3^T - A^T PC - T_1 + T_3^T$,

$\Omega_{23} = -Y_3^T - B^T PC - T_2$, $\Omega_{33} = -Q_2 - T_3 - T_3^T$.

By Schur complement, $\Omega < 0$ is equivalent to

$$\begin{bmatrix} \bar{\Omega}_{11} & \bar{\Omega}_{12} & \bar{\Omega}_{13} & Y_4^T + T_4^T & -Y_1 & -\tau_2 T_1 & A^T S \\ * & \bar{\Omega}_{22} & \bar{\Omega}_{23} & -Y_4^T & -Y_2 & -\tau_2 T_2 & B^T S \\ * & * & \bar{\Omega}_{33} & -T_4^T & -Y_3 & -\tau_2 T_3 & 0 \\ * & * & * & -R & -Y_4 & -\tau_2 T_4 & C^T S \\ * & * & * & * & -\frac{1}{d} Z & 0 & 0 \\ * & * & * & * & * & -\tau_2 W & 0 \\ * & * & * & * & * & * & -S \end{bmatrix} < 0,$$

$\bar{\Omega}_{11} = PA + A^T P + Y_1 + Y_1^T + Q_1 + Q_2 + T_1 + T_1^T$, $\bar{\Omega}_{12} = -Y_1 + PB + Y_2^T + T_2^T$,

$\bar{\Omega}_{22} = -Y_2 - Y_2^T - (1-\mu)Q_1$, $\bar{\Omega}_{13} = Y_3^T - A^T PC - T_1 + T_3^T$,

$\bar{\Omega}_{23} = -Y_3^T - B^T PC - T_2$, $\bar{\Omega}_{33} = -Q_2 - T_3 - T_3^T$.

Multiplying the above linear matrix inequality by $diag(i, I, I, I, dI, I, I)$ gives linear matrix inequality (6) . This completes the proof.

From Theorem 1, we can obtain the following delay-dependent robust stability criterion for the system with time-varying structured uncertainties.

Theorem 2 For given scalars τ_2, d and μ, the neutral system (1) is robustly stable for any delay $d(t)$ satisfying Condition (2) if there exist positive definite matrices $P > 0, Q_1 > 0, Q_2 > 0$, $Z > 0, W > 0, R > 0$ and any appropriate dimensioned matrices $Y_1, Y_2, Y_3, Y_4, T_1, T_2, T_3, T_4$ such that the following linear matrix inequality (LMI) holds.

$$\begin{bmatrix} \hat{\Omega}_{11} & \hat{\Omega}_{12} & \hat{\Omega}_{13} & \hat{\Omega}_{14} & -dY_1 & -\tau_2 T_1 & A^T S & PL \\ * & \hat{\Omega}_{22} & \hat{\Omega}_{23} & -Y_4^T & -dY_2 & -\tau_2 T_2 & B^T S & 0 \\ * & * & \hat{\Omega}_{33} & -T_4^T & -dY_3 & -\tau_2 T_3 & 0 & \hat{\Omega}_{38} \\ * & * & * & -R & -dY_4 & -\tau_2 T_4 & C^T S & 0 \\ * & * & * & * & -dZ & 0 & 0 & 0 \\ * & * & * & * & * & -\tau_2 W & 0 & 0 \\ * & * & * & * & * & * & -S & S^T L \\ * & * & * & * & * & * & * & -I \end{bmatrix} < 0 \tag{21}$$

where

$\hat{\Omega}_{11} = PA + A^T P + Y_1 + Y_1^T + Q_1 + Q_2 + T_1 + T_1^T + E_a^T E_a$, $\hat{\Omega}_{12} = -Y_1 + PB + Y_2^T + T_2^T + E_a^T E_b$,

$$\hat{\Omega}_{22} = -Y_2 - Y_2^T - (1-\mu)Q_1 + E_b^T E_b, \qquad \hat{\Omega}_{13} = Y_3^T - A^T PC - T_1 + T_3^T,$$
$$\hat{\Omega}_{23} = -Y_3^T - B^T PC - T_2, \qquad \hat{\Omega}_{33} = -Q_2 - T_3 - T_3^T,$$
$$\hat{\Omega}_{14} = Y_4^T + T_4^T, \qquad \hat{\Omega}_{38} = -C^T PL.$$

Numerical example

.**Example 1** Consider the following uncertain neutral system with time-varying discrete delay:

$$\dot{x}(t) - \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \dot{x}(t-\tau_2) = \begin{pmatrix} -2+\delta_1 & 0 \\ 0 & -1+\delta_2 \end{pmatrix} x(t) + \begin{pmatrix} -1+\gamma_1 & 0 \\ -1 & -1+\gamma_2 \end{pmatrix} x(t-h(t)),$$

where $0 \le |c| < 1$ and $\delta_1, \delta_2, \gamma_1$ and λ_2 are unknown parameters satisfying

$$|\delta_1| \le 1.6, \ |\delta_2| \le 0.05, \ |\gamma_1| \le 0.1, \ |\gamma_2| \le 0.3.$$

For $\mu = 0.1$, applying the criterion in [8] and Theorem 2 in this paper when $\tau_2 = 10000$, Table1 gives the allowable maximum values d for stability of system for various parameter c. It is obvious that our method is less conservative than Han's method [8].Furthermore, as c increases, d decreases.

Table 1: Allowable time delay d for Example 1 with $\mu = 0.1$

|c|	0.0	0.2	0.3	0.4	0.5	0.6
d Han[8]	0.92	0.55	0.41	0.29	0.19	0.11
d Theorem 2	1.10	0.65	0.47	0.33	0.22	0.13

Conclusion

New delay-dependent stability criteria for neutral systems with time-varying discrete delay and time-varying structured uncertainties are obtained. These criteria are both neutral-and discrete-delay-dependent. Numerical example has shown that the results derived by criterion in this note are much less conservative than some existing ones in the literature.

Acknowledgment

This work is partly supported by ZNDXQYYJJH under grant No. 201021200019, NSFC under grant No. 61070190, NFSS under grant 10BJL020 and the Teaching and Research Award Program for Outstanding Young Teachers in Higher Education Institutions, Hunnan province.

References

[1] A. Bellen, N. Guglielmi and A.E. Ruehli: IEEE Trans. Circuits. System, Vol. 46(1999), p. 212

[2] Y.Y. Cao, Y.X. Sun and C.W. Cheng: IEEE Trans. Automat. Control, Vol. 43(1998), p.1608

[3] Q.L. Han: J. Appl. Math. Comput. Sci., Vol. 11(2001), p. 965

[4] Q.L. Han: Automatica, Vol. 38(2002), p. 719

[5] C.H. Lien, K.W. Yu and J.G. Hsieh: J. Math. Anal. Appl., Vol. 245(2000), p. 215

[6] S.I. Niculescu: Inter. J. Control, Vol. 74(2001), p. 609

[7] J.H. Park: J. Comput. Appl. Math. Vol. 136(2001), p. 177

[8] Q.L. Han: Automatica Vol. 40(2004), p.1087

doi:10.4028/www.scientific.net/AMM.50-51.27

Periodic Solutions for a Kind of Second Order Differential Equation with Multiple Delays

Meilan Tang[a], Xinge Liu[b]

School of Mathematical Science and Computing Technology, Central South University

Changsha, 410083, Hunan, China

[a]tangmeilang@163.com, [b]liuxgliuhua@163.com

Keywords: Multiple delays; Periodic solutions; Duffing equation; Mawhin's continuation theorem

Abstract. Based on Mawhin's continuation theorem and some analysis skill, some sufficient conditions for the existence of periodic solutions for a kind of second order differential equation with multiple deviating arguments are obtained. Results obtained in this paper extend the existing results.

Introduction

The problems of periodic solutions of many types of equations have been extensively studied due to their promising potential for applications in the areas of fluid mechanics, cellular neural networks and nonlinear elastic mechanical phenomena; see [1-10]. In this paper, we will consider the following second order differential equation with multiple delays.

$$x''(t-\tau)+\sum_{i=1}^{n} f_i(x(t-\tau_i(t)))=e(t) \tag{1}$$

where $\tau_i, e \in C(R,R)$ are continuous functions with period T, $\int_0^T e(t)dt=0$, $\tau_i'(t)<1$, $f_i \in C^1(R,R)$, and $\sum_{i=1}^{n} f_i(c) \neq e(t)$ for all $t, c \in R$. By applying Mawhin's continuation theorem and establishing new lemmas, we obtain some sufficient conditions for the existence of T-periodic solutions to Eq.1.

Let X and Y be real Banach spaces and $L: D(L) \subset X \to Y$ be a Fredholm operator with index zero; here $D(L)$ denotes the domain of L. This means that $\operatorname{Im} L$ is closed in Y and $\dim \ker L = \dim (Y / \operatorname{Im} L) < +\infty$. Consider X_1, Y_1, the supplementary subspaces of X, Y respectively, such that $X = \ker L \oplus X_1$ and $Y = \operatorname{Im} L \oplus Y_1$. Let $P: X \to \ker L$ and $Q: Y \to Y_1$ be natural projectors. Clearly, $\ker L \bigcap (D(L) \bigcap X_1) = \{0\}$; thus the restriction $L_P = L|_{D(L)\cap X_1}$ is invertible. Denote by K the inverse of L_P. Let Ω be an open bounded subset of X with $D(L) \bigcap \Omega \neq \phi$. A map $N: \bar{\Omega} \to Y$ is said to be L-compact in $\bar{\Omega}$ if $QN(\bar{\Omega})$ is bounded and the operator $K(I-Q)N: \bar{\Omega} \to X$ is compact.

Let $|x|_\infty = \max_{t\in[0,T]} |x(t)|$,

$$|x|_k = \left(\int_0^T |x(t)|^k \, dt\right)^{\frac{1}{k}}, C_T = \{x \mid x(\cdot) \in C(R,R) : x(t+T) = x(t), \forall t \in R\}$$

with the norm $\|x\|_1 = |x|_\infty = \max_{t\in[0,T]} |x(t)|$ and $C_T^1 = \{x \mid x(\cdot) \in C^1(R,R) : x(t+T) = x(t), \forall t \in R\}$ with the norm $\|x\|_2 = \max\{|x|_\infty, |x'|_\infty\}$. Let $X = C_T^1$ and $Y = C_T$. Then X and Y are two Banach spaces. Denote $D(L) = \{x \mid x \in X, x'' \in C(R,R)\}$. Define a linear operator $L: D(L) \subset X \to Y$ by

$$Lx = x''(t-\tau). \tag{2}$$

Define a nonlinear operator $N: X \to Y$ by

$$Nx = -\sum_{i=1}^{n} f_i(x(t-\tau_i(t))) + e(t). \tag{3}$$

It is easy to see that $\ker L = R$, and $\operatorname{Im} L = \{x \mid x \in Y, \int_0^T x(s)ds = 0\}$.

Thus, the operator L is a Fredholm operator with index zero.Define the continuous projectors $P: X \to \ker L$ and $Q: Y \to \operatorname{Im} Q$ by setting

$$P(x(t)) = x(0) \text{ , and } \quad Q(x(t)) = \frac{1}{T}\int_0^T x(s)ds.$$

Obviously, $\ker L = \operatorname{Im} P$ and $\ker Q = \operatorname{Im} L$. Denoting by $K: \operatorname{Im} L \to D(L) \bigcap \ker P$ the inverse of $L|_{D(L)\cap \ker P}$, We can see that N is L-compact on $\bar{\Omega}$, where Ω is an arbitrary open bounded subset of X.

Preliminaries

Corresponding to the operator equation $Lx = \lambda Nx$, $\lambda \in (0,1)$, we have

$$x''(t-\tau) + \lambda \sum_{i=1}^{n} f_i(x(t-\tau_i(t))) = \lambda e(t) \tag{4}$$

Lemma 1 [3] Suppose that $x(t) \in C_T^1$, and there is a point $t_0 \in R$ such that $x(t_0) = 0$, then

$$\int_0^T |x(t)|^2 dt \le \frac{T^2}{\pi^2}\int_0^T |x'(t)|^2 \, dt \tag{5}$$

Lemma 2[9] Suppose that X and Y are two Banach spaces, and that $L: D(L) \subset X \to Y$ is a Fredholm operator with index zero, furthermore, $\Omega \subset X$ is an open bounded set and $N: \bar{\Omega} \to Y$ is L-compact on $\bar{\Omega}$. If

(1) $Lx \neq \lambda Nx, \forall x \in \partial\Omega \bigcap D(L), \lambda \in (0,1)$;

(2) $Nx \notin \operatorname{Im} L, \forall x \in \partial\Omega \bigcap \ker L$

(3) the Brouwer degree $\deg(QN, \Omega \bigcap \ker L, 0) \neq 0$.

Then the equation $Lx = Nx$ has a solution in $\bar{\Omega} \bigcap D(L)$

Lemma 3 Assume that the following conditions hold

(H1) $(f_i(x_1) - f_i(x_2))(x_1 - x_2) > 0, x_1, x_2 \in R, x_1 \neq x_2, i = 1,2,\cdots n;$

(H2) There is a constant $d > 0$ such that

$$x\sum_{i=1}^{n} f_i(x) > 0, \quad |x| > d.$$

If $x(t)$ is a T-periodic solution of Eq. 4, then

$$|x|_\infty \le d + \sqrt{T}\,|x'|_2. \tag{6}$$

Proof. Let $x(t)$ be a T-periodic solution of Eq. 4. Integrating both sides of Eq. 4 from 0 to T gives

$$\int_0^T \sum_{i=1}^{n} f_i(x(t-\tau_i(t)))dt = 0 \tag{7}$$

which implies that there exists a $\xi \in [0,T]$ such that

$$\sum_{i=1}^{n} f_i(x(\xi - \tau_i(\xi))) = 0. \tag{8}$$

Now, we claim that there is a $t_0 \in R$ such that

$$|x(t_0)| \le d. \tag{9}$$

If there exist i and j such that $x(\xi - \tau_i(\xi))x(\xi - \tau_j(\xi)) \le 0$, there is a $t_0 \in R$ such that $x(t_0) = 0$, obviously, Eq. 9 holds.

If $x(\xi-\tau_i(\xi)), i=1,2,\cdots,n$ have the same sign, we will prove $|x(t_0)|\le d$ by way of contradiction. In fact, assume that Eq. 9 does not hold. Then

$$|x(t)|> d, \forall t \in R \tag{10}$$

There are two cases: (1) $x(\xi-\tau_i(\xi))>0, i=1,2,\cdots,n$ or(2) $x(\xi-\tau_i(\xi))<0, i=1,2,\cdots,n$.

If $x(\xi-\tau_i(\xi))>0, i=1,2,\cdots,n$, then $x(\xi-\tau_i(\xi))>d, i=1,2,\cdots,n$. Suppose that

$$x(\xi-\tau_k(\xi))=\min_{1\le i\le n} x(\xi-\tau_i(\xi)), \tag{11}$$

then $x(\xi-\tau_k(\xi))>d>0$. According to (H2)

$$\sum_{i=1}^{n} f_i(x(\xi-\tau_k(\xi)))>0. \tag{12}$$

Assumption (H1) implies that

$$\sum_{i=1}^{n} f_i(x(\xi-\tau_k(\xi)))<\sum_{i=1}^{n} f_i(x(\xi-\tau_i(\xi))) \tag{13}$$

Combining Eq.12 and Eq.13 gives

$$\sum_{i=1}^{n} f_i(x(\xi-\tau_i(\xi)))>0. \tag{14}$$

Obviously, Eq.14 contradicts Eq. 8. This contradiction implies that Eq.9 is true.

If $x(\xi-\tau_i(\xi))<0, i=1,2,\cdots,n$, then $x(\xi-\tau_i(\xi))<-d, i=1,2,\cdots,n$. Suppose that

$$x(\xi-\tau_m(\xi))=\max_{1\le i\le n} x(\xi-\tau_i(\xi)), \tag{15}$$

then $x(\xi-\tau_m(\xi))<-d<0$. According to (H2)

$$\sum_{i=1}^{n} f_i(x(\xi-\tau_m(\xi)))<0. \tag{16}$$

Assumption (H1) implies that

$$\sum_{i=1}^{n} f_i(x(\xi-\tau_m(\xi)))>\sum_{i=1}^{n} f_i(x(\xi-\tau_i(\xi))). \tag{17}$$

Combining Eq.12 and Eq.13 gives

$$\sum_{i=1}^{n} f_i(x(\xi-\tau_i(\xi)))<0. \tag{18}$$

Obviously, Eq.18 contradicts Eq. 8. This contradiction implies that Eq.9 is true.

Let $t_0=nT+t_1$, where n is an integer and $t_i\in[0,T]$. Since $x(t)$ is T-periodic, $x(t_0)=x(t_1)$. Then, $\forall t\in[0,T]$,

$$|x(t)|=\left|x(t_1)+\int_{t_i}^{t} x'(s)ds\right|\le d+\int_0^T|x'(s)|\,ds\le d+\sqrt{T}\left(\int_0^T|x'(s)|^2\,ds\right)^{\frac{1}{2}}. \tag{19}$$

Thus $|x|_\infty\le d+\sqrt{T}\,|x'|_2$.

From Lemma 2.4[12], we can obtain

Lemma 4 Let $\tau_i\in C_T^1, \tau_i'<1, i=1,2,\cdots,n$。 If $x(t)\in C_T$, Then

$$\int_0^T|x(t-\tau_i(t)|^2\,dt\le\gamma_i\int_0^T|x(t)|^2\,dt. \tag{20}$$

where $\gamma_i=\dfrac{1}{1-|\tau_i'|_\infty}$.

The existence of T-periodic solutions

Theorem 1 Suppose that (H1),(H2) and the following assumption hold:

(H3) There exist nonnegative constants k_i, c_i such that

$$|f_i(x)| \le k_i |x| + c_i, i = 1,2,\cdots,n, \quad |x| > d.$$

Then Eq. 1 has at least one T-periodic solution if

$$\frac{T^2}{\pi^2}\sum_{i=1}^{n} k_i\sqrt{\gamma_i} < 1. \tag{21}$$

where $\tau_i' < 1, \gamma_i = \dfrac{1}{1-|\tau_i'|_\infty}$.

Proof Let $x(t)$ be a T-periodic solution of Eq. 4. Multiplying both sides of Eq. 4 by $x(t-\tau)$ and integrating from 0 to T, we obtain

$$\int_0^T |x'(t-\tau)|^2\, dt = \lambda\int_0^T \sum_{i=1}^{n} f_i(x(t-\tau_i(t)))x(t-\tau)dt - \lambda\int_0^T e(t)x(t-\tau)dt$$

$$\le \int_0^T \sum_{i=1}^{n} |f_i(x(t-\tau_i(t)))x(t-\tau)|\, dt + \int_0^T |e(t)x(t-\tau)|\, dt \tag{22}$$

Let $E_i = \{t \in [0,T], |x(t-\tau_i(t))| \le d\}$, $S_i = \{t \in [0,T], |x(t-\tau_i(t))| > d\}$, $\tilde{f}_i = \max\limits_{|x|\le d} |f(x)|$

Then

$$\begin{aligned} |x'|_2^2 &= \int_0^T |x'(t-\tau)|^2\, dt \\ &\le \sum_{i=1}^{n}\int_0^T |f_i(x(t-\tau_i(t)))x(t-\tau)|\, dt + \int_0^T |e(t)x(t-\tau)|\, dt \\ &\le \sum_{i=1}^{n} \tilde{f}_i\int_0^T |x(t)|\, dt + \sum_{i=1}^{n} k_i\int_0^T |x(t-\tau_i(t))x(t-\tau)|\, dt \\ &\quad + \sum_{i=1}^{n} c_i\int_0^T |x(t)|\, dt + |e|_\infty \int_0^T |x(t)|\, dt \\ &\le \left[|e|_\infty + \sum_{i=1}^{n}(\tilde{f}_i + c_i)\right]\sqrt{T}\,|x|_2 + \sum_{i=1}^{n} k_i\sqrt{\gamma_i}\,|x|_2^2. \end{aligned} \tag{23}$$

From Eq.9, there exists a point $t_0 \in R$ such that $x(t_0) \le d$. Let $w(t) = x(t) - x(t_0)$, then $w(t+T) = w(t)$, $w'(t) = x'(t)$ and $w(t_0) = 0$. By Lemma 1

$$|w|_2 \le \frac{T}{\pi}|w'|_2 = \frac{T}{\pi}|x'|_2.$$

By using the Minkowski inequality, we have

$$|x|_2 = \left(\int_0^T |x(t)|^2\, dt\right)^{\frac{1}{2}} = \left(\int_0^T |w(t) + x(t_0)|^2\, dt\right)^{\frac{1}{2}} \le \left(\int_0^T |x(t_0)|^2\, dt\right)^{\frac{1}{2}} + \left(\int_0^T |w(t)|^2\, dt\right)^{\frac{1}{2}}.$$

So

$$|x|_2 \le \sqrt{T}d + |w|_2 \le \sqrt{T}d + \frac{T}{\pi}|x'|_2 \tag{24}$$

Combining Eq. 23and Eq. 24 gives

$$\begin{aligned} |x'|_2^2 &\le \left[|e|_\infty + \sum_{i=1}^{n}(\tilde{f}_i + c_i)\right]\sqrt{T}\,|x|_2 + \sum_{i=1}^{n} k_i\sqrt{\gamma_i}\,|x|_2^2 \\ &\le \left[|e|_\infty + \sum_{i=1}^{n}(\tilde{f}_i + c_i)\right]\left[Td + \frac{T\sqrt{T}}{\pi}|x'|_2\right] + \sum_{i=1}^{n} k_i\sqrt{\gamma_i}\left[Td + \frac{2T\sqrt{T}d}{\pi}|x'|_2 + \frac{T^2}{\pi^2}|x'|_2^2\right]. \end{aligned} \tag{25}$$

Since $\sum\limits_{i=1}^{n} k_i\sqrt{\gamma_i}\dfrac{T^2}{\pi^2} < 1$, there is a constant $M_0 > 0$ such that $|x'|_2 \le M_0$. So

$$|x|_\infty \le d + \sqrt{T}\,|x'|_2 \le d + \sqrt{T}M_0 := M_1.$$

Since $x(0) = x(T)$, there is a constant $\eta \in (0,T)$ such that $x'(\eta) = 0$.

Let $\hat{f}_i = \max_{|x|\le M_1} |f(x)|, M_2 = \sum_{i=1}^{n} \hat{f}_i T + |e|_\infty T$. Then, for all $t \in [0,T]$

$$|x'(t)| = \left| x'(\eta) + \int_\eta^t x''(t)dt \right| \le \sum_{i=1}^{n} \int_0^T |f_i(x(t-\tau_i(t)))|\,dt + \int_0^T |e(t)|\,dt \le M_2. \tag{26}$$

If $x \in \Omega_1 := \{x \mid x \in \ker L \cap X, \text{and } Nx \in \operatorname{Im} L\}$, then $x \in \ker L$, which means that there exists a constant M_3 such that $x(t) = M_3$, and $-\frac{1}{T}\int_0^T \sum_{i=1}^{n} f_i(M_3)dt = 0$.

So

$$\sum_{i=1}^{n} f_i(M_3) = 0. \tag{27}$$

Eq. 27 and (H2) imply that

$$|x(t)| = |M_3| \le d \quad \text{for all } x(t) \in \Omega_1.$$

Let $M = M_1 + M_2$, $\Omega = \{x \in X \mid |x|_\infty < M, |x'|_\infty < M\}$.

It is easy to see that N is L-compact on $\bar{\Omega}$, and from the above discussions that conditions (1) and (2) of Lemma 2 are satisfied. Furthermore, define

$$H(x,u) = ux + (1-u)\frac{1}{T}\int_0^T \sum_{i=1}^{n} f_i(x(t-\tau_i(t)))dt, \ \forall x \in \Omega, \ u \in [0,1].$$

Then

$$xH(x,u) = ux^2 + (1-u)x\sum_{i=1}^{n} f_i(x), \ \forall x \in \partial\Omega \cap \ker L, u \in [0,1].$$

Hence, using the homotopy invariance theorem, we have

$$\deg(QN, \Omega \cap \ker L, 0) = \deg(I, \Omega \cap \ker L, 0) \neq 0.$$

By Lemma 2, Eq. 1 has at least one T-periodic solution.

Acknowledgment

This work is partly supported by ZNDXQYYJJH under grant No. 201021200019, NSFC under grant No. 61070190, NFSS under grant 10BJL020 and the Teaching and Research Award Program for Outstanding Young Teachers in Higher Education Institutions, Hunnan province.

References

[1] P.J. Torres: J. Differential Equations Vol. 190(2003) , p.643-662.

[2] B.W. Liu:Nonlinear Anal. Vol. 69 (2008) , p. 724-729.

[3] X.P. Liu, M. Jia, R. Ren:Acta Math. Sci. Vol. 27 (2007), p. 379(in Chinese).

[4] M.R. Pournakia, A. Razanib: Appl. Math. Lett. Vol. 20 (2007) , p. 248.

[5] H.O. Walther: J. Differential Equations Vol. 244 (2008) , p.1910-1945.

[6] X.J. Li, S.P. Lu:Nonlinear Anal. Vol. 70 (2009) , p.1011-1022.

[7] S.P. Lu:Nonlinear Anal. Vol. 68 (2008) , p.1453.

[8] Z.H. Yang, J.D.Cao:Nonlinear Anal. Vol. 66 (2007) , p.1378.

[9] R.E. Gaines, J.L. Mawhin: Springer, Berlin, (1977).

[10] Z.G. Wang, L.X.Qian, S.P. Lu, and J. D. Cao: Nonlinear Analysis (2010), in press.

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.32

Research on Impedance Characterization of Triple-layer Piezoelectric Bender

Lijiao Gong[1,a] and Jiangquan Li[1,b]

[1]Machinery and Electrical Engineering College, Shihezi University, 280 Beisi Road, Shihezi Xinjiang, 832003, China

[a]emaill:glj_mac@shzu.edu.cn,[b]Author to whom correspondence should be addressed email: ljq_mac@shzu.edu.cn

Keywords: piezoelectric bender, lead zirconate titanate, electrical impedance, equivalent circuit

Abstract. The use of piezoelectric material as transducer is prevalent. Piezoelectric bending mode element can be used for vibration energy harvesting because it can convert mechanical energy into electrical energy. In this article, electrical impedance characterization and equivalent circuit of triple-layer piezoelectric bender are discussed. Triple-layer piezoelectric bending device is fabricated, measured and modeled. This paper is aimed to explore simple but practical equivalent circuit models, expressed using electrical parameters of triple-layer piezoelectric bender and investigate the applicability of the Van Dyke circuit model and the complex circuit model in modeling the specimen's equivalent circuit. The models produced impedance curves that closely matched the impedance measured for piezoelectric sample. The impedance characterization can provide a good understanding on the electrical behaviors of the triple layer piezoelectric bender when analyzing the performance of piezoelectric device.

Introduction

The piezoelectric materials, as the most widely used functional materials in smart structures, have many outstanding advantages for transducers. The advantages such as: little dimensions, simple structure, low noise factor, their excellent mechanical-electrical coupling characteristics and frequency response characteristics at operating provide wide applications in vibration generating, damping, converse of mechanical energy into electrical energy and oppositely. Some novel device based on piezoelectric materials have been developed in recent years for a wide range of electromechanical applications.[1-3] The theoretical analysis of the sensor effect of triple layer piezoelectric bender has been carried out. [2] The potential of the triple-layer piezoelectric benders has been investigated through experiments for utilization in energy harvesting applications.[4,5] Since electrical signals involved in most of the piezo-based systems, a good understanding of the electrical characteristics of piezoelectric materials is essential to design and analyze the systems. Therefore, circuit model for the piezoelectric material is desired.

Several researches have been done in this area. The Van Dyke's circuit model [6] is widely used for representing the equivalent circuit of a piezoelectric vibrator. It is recommended by the IEEE Standard on Piezoelectricity [7]. Sherrit et al. have used complex material constants to model the performance of impedance of the piezoelectric resonator [8]. Mingjie Guan et al. have studied the circuit model of piezoelectric ceramics [9].

Equivalent circuit models of the triple-layer piezoelectric bender will be discussed. The way to obtain the parameters that characterize the electrical model of the triple-layer piezoelectric bender is studied. In this paper, the fabrication and impedance properties of triple layer bender made of PZT are studied.

Methods & Materials

Sample Preparations. The structure of the triple-layer piezoelectric bender is illustrated as the basic metal-piezoelectric sandwich structures in Fig. 1. It is fabricated by a metal layer covered symmetrically by two transversely piezoelectric layers poled along the thickness direction. The piezoelectric material is using lead zirconate titanate (PZT) (P-51, Haiying Enterprise Group Co. Ltd), and the metal elastic layer is phosphor bronze because of its larger elastic modulus. The piezoelectric were bonded on the beryllium bronze using epoxy (Emerson &Cumming), and then cured at a temperature of 80°C for two hours. The overall design of insulation must be considered. As for the measured specimen, the geometries are given in Table 1 (length l and width w and the thickness t for a piezoelectric bender). Material key parameters are listed in Table 2.

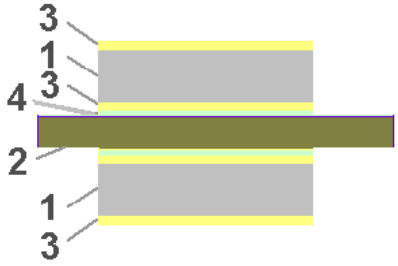


Fig. 1 Schematic drawing of the triple-layer piezoelectric bender. 1- piezoelectric material layer, 2- metal material, 3- thin film electrodes, 4- conductive adhesive layer.

Table 1. The physical sizes of the sample

Physical	l_p	w_p	t_p	l_m	w_m	t_m
Dimensions[mm]	72	14	0.4	112	14	0.3

Table 2. Properties of the piezoelectric and beryllium bronze

Property	Symbol	Value
Displacement Coefficient $[\times 10^{-12} C/N]$	d_{33}	400
Relative Dielectric Constant	ε^{T}_{33}	2100
Elastic compliances constant $[\times 10^{-12} m^2/N]$	S^{E}_{33}	19.0
Density of Piezo Material $[\times 10^{3} kg/m^3]$	ρ_p	7.45
Young's modulus of bronze $[\times 10^{11} N/m^2]$	E	1.30
Density of bronze$[\times 10^{3} kg/m^3]$	ρ_m	8.23

Impedance Measurements. An Impedance Analyzer (Agilent 4294A) will be used to record the impedance and phase angle. The sweep voltage level in the Impedance Analyzer is set at 500mV. Let Impedance Analyzer to perform electrical impedance measurements by producing sweeping frequency signal in the 40Hz-1 kHz range. The upper bound of the frequency range is chosen as 1 kHz. This is because these are usually the major concerns in structural application below the frequencies.

Results and Discussion

The response curves of the impedance value and phase of the specimen are illustrated as the solid line in Fig. 2. From the figure, we can see that the values of the impedance phase has its lowest point near 410Hz reaching -76°, besides that, the values of the impedance phase lie near -88.6°. This means that

give a current component -88.6° out of phase with an applied voltage if there is no resonance (resonances here refers to either the resonance or the anti-resonance frequency); the impedance characteristic of specimen is similar to capacitance. It is thus clear that the specimen behaves are like piezoelectric ceramics. Based on the above observation, we can divide the impedance curve into two parts: the resonant frequency part and non-resonant part. From the impedance phase response, we can see that the impedance phase curve within the non-resonant range is close to -90°, but not equally to -90°, which means that there is the little energy dissipation.

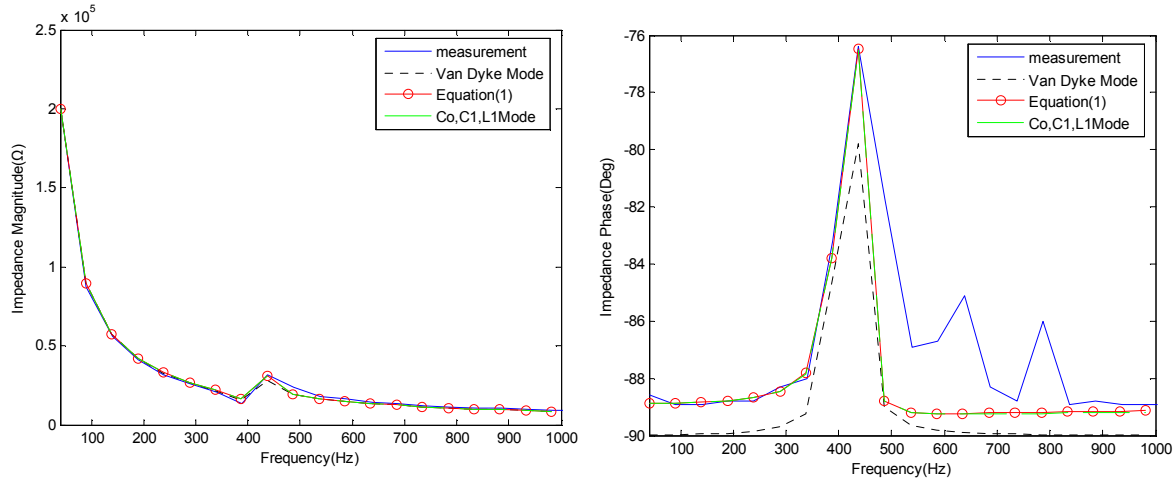


Fig. 2 Impedance and phase angle of the specimen vs. frequency.

The Impedance magnitude plotted on a linear scale versus the frequency (left) and the impedance phase versus the frequency (right) for specimen. The solid line is the response curves of the impedance magnitude and phase of the specimen. The dotted line is the Van Dyke mode fit to the experimental data. The data points shown as circles were generated using Eq. 1. The dashed line is the complex circuit model fit to the experimental data.

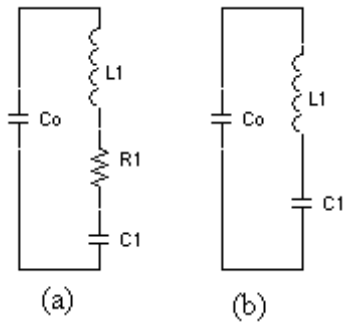


Fig. 3 (a) The Van Dyke circuit model: the values of the circuit elements are real. (b) The Complex Circuit with 3 complex parameters

In the analysis of piezoelectric materials there are two approaches to model the impedance of piezoelectric vibrators. One is lumped circuit model that is geometry independent and another is a circuit model of a piezoelectric material that is from the materials point of view. For transducer applications, it is more accessible to fit the impedance plots to lumped circuit models to fit the impedance curve in the vicinity of the resonant frequency and predict the electrical behavior of the resonator. We firstly consider the Van Dyke's circuit model as shown in Figure 3a. The Van Dyke model uses four real circuit parameters, C_0, C_1, L_1 and R_1, to represent the impedance of a free-standing piezoelectric resonator around resonance. Based on the model, we can fit the experimental impedance spectrum by choosing the values of the electronic components. When determining the values of the electronic components, the dominant component C_0 should be first considered according to the impedance magnitude. After the value of C_0 is chosen, the value C_1 is chosen according to the appearance of the impedance spectrum around the resonance. The component C_1 has significant effect on the total impedance in the vicinity of resonance frequency bandwidth. The value of the components L_1 is relevant to C_1 and is considered according to the resonant frequency.

The value of component R_1 is adjusted according to impedance phase in the vicinity of resonance frequency bandwidth. The results for the impedance of the circuit model are shown as the dotted line in Fig. 2. In addition, an alternative approach to fitting the experimental impedance spectrum is to use the complex circuit model as shown in Figure 3b. The model contains three circuit elements, $\mathbf{C_0}$, $\mathbf{C_1}$ and $\mathbf{L_1}$, which are defined as complex and takes into account the dielectric, elastic and piezoelectric losses of the piezoelectric material. Parameters selection method is similar to the previous, however, first choose the real value, and then choose the imaginary value. The circuit constants for this simulation are listed in Table 3. The results for the impedance of the complex circuit model are shown as the dashed line in Fig. 2.

Table3. Values of the components in the equivalent circuit model of the specimen. The various curves generated with these constants are shown in Fig. 2

Constant	Value(real and imaginary)
Van Dyke's model	
C_0 (F)	19.011×10^{-9}
C_1 (F)	0.86970×10^{-9}
L_1 (H)	171.9437
R_1 (Ω)	19.921×10^{3}
The complex circuit model	
$\mathbf{C_0}$ (F)	19.011×10^{-9}(1-0.0145i)
$\mathbf{C_1}$ (F)	0.86570×10^{-9} (1-0.1308i)
$\mathbf{L_1}$ (H)	167.0094 (1+0.0866i)
Equation (1) constant	
f_s (Hz)	415.92+9.0845i
f_p (Hz)	425.32+8.2126i
ε_{33}^S (Fm^{-1})	20.538×10^{-9}(1 – 0.01338i)
k_t	0.2312(1 - 0.0545i)

The IEEE Standard (IEEE Standards ANSI/IEEE 176-1987) [7] on the impedance equation for analyzing the thickness resonance mode is described as a function of the frequency ω by the equation

$$Z(\omega)=\frac{t}{iA\omega\varepsilon_{33}^S}\left(1-\frac{k_t^2\tan[\omega/(4f_p)]}{\omega/(4f_p)}\right) \quad (1)$$

Where t is the sample thickness, A is the electrode area; k_t is the thickness electromechanical coupling constant and ε_{33}^S is the clamped permittivity. The parameters k_t, ε_{33}^S and f_p were originally assumed to be real in the derivation of the resonator equation (1). S Sherrit et al. [8] have pointed out that the circuit model of a piezoelectric material can be characterized by its complex material constants. The complex circuit model proposed by S Sherrit is shown in Fig. 3b. When the complex circuit parameters are known, the materials constants can be calculated from circuit parameters. The material constants can be determined directly from the complex circuit constants by using Eq. (2) and (3) to determine the complex frequencies f_s and f_p.

$$f_s=\frac{1}{2\pi\sqrt{L_1C_1}} \quad (2)$$

$$f_p = \frac{1}{2\pi\sqrt{L_1 \frac{C_1 C_0}{C_1 + C_0}}} \tag{3}$$

The electromechanical coupling constant can then be calculated using

$$k_t^2 = \frac{\pi}{2}\frac{f_s}{f_p}\tan\left(\frac{\pi}{2}\frac{f_p - f_s}{f_p}\right) \tag{4}$$

The dielectric permittivity is

$$\varepsilon_{33}^S = (t/A)(1-k_t^2)(C_1 + C_0) \tag{5}$$

The complex material constants and the complex frequencies were calculated using the complex circuit constants and Eq. 2-5. The calculated materials constants are listed in Table 3. The predicted impedance spectra generated by the impedance function (Eq. 1) are shown as circles in Fig.2.

The measured impedance spectra together with the simulations obtained respectively by the Van Dyke model, the complex circuit model and the impedance function (Eq. 1) are shown in Fig. 2. It can be seen that the complex circuit model and the impedance function can fit the experimental data quite well from low frequencies to frequencies above the resonance. Although the Van Dyke model gives a good fit to the impedance magnitude, it can be seen to not reach -88.6° but asymptotically reach -90° at end of the frequency limits in the impedance phase. This is probably due to the known dielectric loss. Since the Van Dyke model has only one loss component (R_1), the impedance phase curve in the Van Dyke model will always reach -90°. The derived material constants, complex frequencies, complex circuit model constants and Van Dyke circuit constants used in the simulation are listed in Table 3.

Conclusions

The equivalent circuit models of the triple-layer piezoelectric bender are studied. The impedance information is evaluated and the relative material properties of the triple layer piezoelectric bender are calculated. The impedance curve spectrum are fitted with various models. The results of the fitting show that the impedance curves produced by complex circuit model are effectively matched the impedance measured for piezoelectric sample.

Acknowledgment

This work described in this paper was supported by a grant from the National Natural Science Foundation of China (Project No. 50907042)

References

[1] Steven R Anton and H. A Sodano: Smart Mater. Struct. **16** (2007) R1-R21
[2] Q.-M. Wang, X.-H. Du, Baomin XU, and L. Eric Cross:J. Appl. Phys., Vol.85,No.3(1999), p. 1702
[3] N. W. Hagood, W. H. Chung, A. Flowtow: J. INTEL. MAT. SYST. STR., vol. 1(1990), p. 327
[4] L. J. Gong, China patent, CN101262189. (2008).
[5] GONG Li-jiao, SHEN Xing: Adv. Tech. of Elec. Eng. & Energy, Vol.29, No.1(2010),p.54.
[6] Martin G: J.Acoust. Soc. Am. 26, (3) (1954), p.413
[7] *IEEE Standard on Piezoelectricity*, (IEEE Standard 176-1987), Institute of Electrical and Electronic Engineers (1987)
[8] S Sherrit et al: J. Phys. D: Appl. Phys.**30** (l997), p.2354
[9] Mingjie Guan, W.-H. Liao: IEEE Proceedings of International Conference on Information Acquisition (2004),p.26

doi:10.4028/www.scientific.net/AMM.50-51.37

Dynamic modelling and analysis of rectangle-shaped plastic-coated slideways in machine tools

Jianfu Zhang[1, a], Zhijun Wu[2, b], Pingfa Feng[3, c] and Dingwen Yu[4, d]

[1,2,3,4]Institute of Mechanical Engineering, Tsinghua University，Beijing, 100084, P.R. China

[a]zhjf, [b]wuzhijun, [c]fengpf, [d]yudw@tsinghua.edu.cn

Keywords: plastic-coated slideway; theoretical model; dynamic behavior; vibration frequency

Abstract. The plastic-coated slideways have been widely used for form-generating movement in machine tools. Its dynamic behavior plays an important role in the vibration properties of the whole machine. In this work, according to the situation that researches on this subject were rather insufficient, a theoretical research was analyzed concerning the stiffness and damping characteristics of rectangle-shaped plastic-coated slideways. The mathematical model was firstly suggested especially based on the assembly of the saddle and worktable. Both stiffness and damping characteristics on vertical and horizontal directions were theoretically determined. To derive the governing motion equation of the slideway system, the carriage and rail were considered as rigid bodies and connected with a series of spring and damping elements at the joint face. Moreover, through the Lagrange's approach, the frequencies of the carriage at vertical, pitching, yawing and rolling vibration mode were identified.

Introduction

Generally, the guideways can be classified into three types, depending upon what the interfacial layers are, i.e., the plastic-coated slideway, the hydrostatic guideway, and the rolling guideway. Of these guideways, the slideway is worth actively investigating, because it is a typical variant of flat joint [1]. The plastic-coated slideways have been widely used to transport machine parts through a linear path in machining centre, e.g., the milling head box and machine column, worktable and saddle, saddle and bed. The configuration style of the plastic-coated slideway is posted a layer of plastic or other chemical composition of soft plastic film on the friction surface of the moving carriage (such as the worktable). Some characteristics of the plastic-coated slideways are high load capacity, vibration absorption capacity of the plastics, wear resistance with adequate lubrication, simplicity of design and good manufacturing performance, simple process and low cost, close static and dynamic friction coefficient, and possibility of arising stick-slip. To accurately predict the dynamic characteristics, the micro and macro dynamic behavior of the plastic-coated slideways must be understood [2,3]. However, researches on this subject are rather insufficient and limited.

In this paper, theoretical researches of the dynamic behavior of the rectangle-shaped plastic-coated slideway were carried out based on the assembly of the worktable and saddle in the machining centre. A mathematical equivalent model was suggested concerning the stiffness and damping properties mounted on the joint faces of the worktable and saddle. The purpose of this paper was to derive the governing motion equation of the rectangle-shaped plastic-coated slideway. At last, the frequency and dynamic behavior of the carriage at different directions vibration mode are identified based on the Lagrange's approach.

Mathematical modeling of the plastic-coated slideways

Plastic-coated slideways are originally developed to be used for machine tools. This mechanical fit and transmission type has widely adopted to substitute the rolling guideway in some certain situations. They are well combined and configured using core parts, e.g., worktable (worktable slideway), saddle (saddle slideway), gib, and keep plate in the case of the machine tools[4]. A typical

application is the worktable slideway (worktable for carriage and saddle for rail) depicted in Fig.1. As can be seen from the section view of the assembly, a rectangle-shaped slideway consists of several flat joints. This slideway system of the bottom slide is composed of two horizontal flat guideways ‘a’, two vertical flat guideway with gib-embedded ‘b’ and two retainer keep plates for cover strips ‘c’. Because the main purpose of this paper is to develop an analysis concerning the stiffness and damping characteristics of the plastic-coated slideways mounted on the worktable and saddle, the dynamic characteristics of the hydrodynamic contact film are neglected firstly. According to the actual situation, each joint can be replaced with the mathematical model described by spring and damping elements. Moreover, if the interfaces of the carriage and rail are carefully considered as rigid joints, its stiffness and damping characteristics of the coated plastic will have a major influence on the carriage dynamic state.

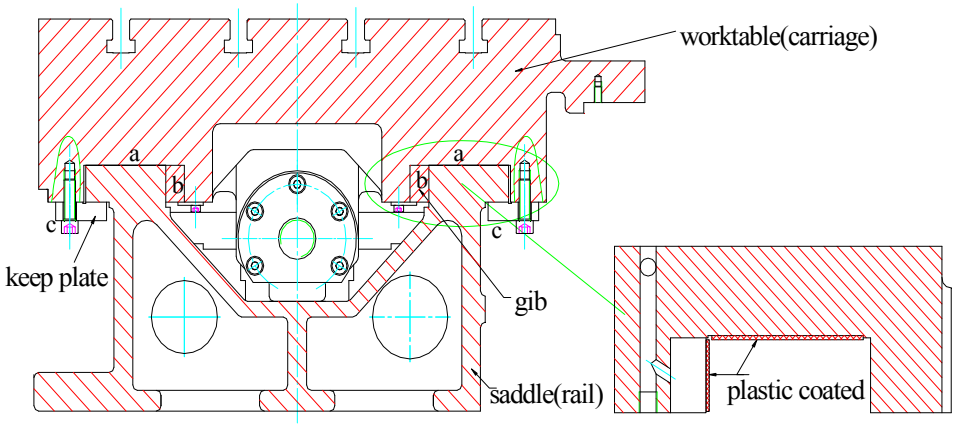


Fig.1 Section view of the worktable and saddle

To derive the motion equation of the plastic-coated slideway system mentioned above, the worktable and saddle are considered as rigid bodies and connected with a series of spring and damping elements at the joint face. In mathematical modeling for the contact structure, the following aspects are also considered[5-7]: 1) elastic deformation of the Hertzian contact (the contact between carriage and rail can be considered as a Hertzian contact model. 2) squeeze and damping effects in joint faces. 3) effects of hydrodynamic contact film are neglected. Based on the contact structure of the worktable and saddle, the coordinate system and equivalent model of the slideway system are illustrated in Fig.2.

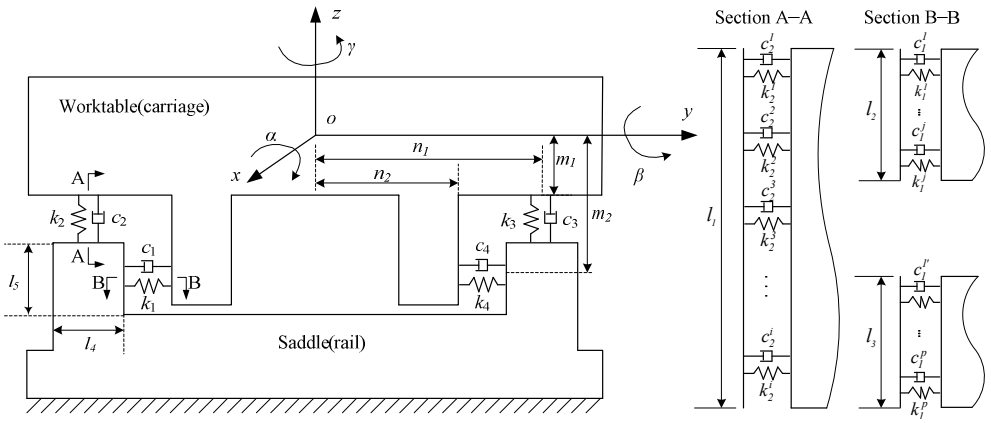


Fig.2 Equivalent model of the slideway system

In the equivalent model of the slideway, the origin of the coordinate is located at the mass center of the worktable, x, y and z are the displacements in the x-axis, y-axis and z-axis respectively. m_1, m_2 and n_1, n_2 are distances of the contact points between the carriage and rail to x-y plane and x-z plane. l_1 and l_4 are the length and width of the spring and damping elements at the horizontal joint face ‘a’ respectively. l_2 and l_3 (here, l_3 equals l_2, and it also equals the length of the two gibs) are the length of

the spring and damping elements at the vertical joint face 'b', and l_5 is the width of the spring and damping elements at the vertical joint face 'b'. α, β and γ are the angular displacements about x-axis, y-axis and z-axis respectively, which are termed rolling (α), pitching (β) and yawing (γ) motion respectively. The motion of the carriage along the y-axis is called the horizontal vibration motion, and the motion of the carriage along the z-axis is called the vertical vibration motion.

Dynamic characteristics of plastic-coated slideways

The equation of motion of the rectangle-shaped plastic-coated slideway system is obtained by using the Lagrange's approach [8, 9]. This method regards the system as a whole, from the energy point of view, using the scalars such as the kinetic energy and potential energy to describe the functions, and using the variational method to establish the system dynamic equation. Lagrange's equation gives a common, simple unified solution for the dynamics problems. According to the Lagrange's method, the typical vibration differential equation of the dynamic system can be denoted as follows

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right)-\frac{\partial T}{\partial q_i}+\frac{\partial V}{\partial q_i}=Q_i \quad (i=1,2,\cdots,n)\,. \tag{1}$$

Where, q_i and $\dot{q}_i$ are the generalized coordinates and generalized velocity, T, V and Q are the kinetic energy, potential energy and non-conservative force respectively.

Based on the proposed equivalent model of the slideway, and then using the Lagrange's approach, the kinetic energy T of the system in motion is given by

$$T=\frac{1}{2}M\dot{y}^2+\frac{1}{2}M\dot{z}^2+\frac{1}{2}J_x\dot{\alpha}^2+\frac{1}{2}J_y\dot{\beta}^2+\frac{1}{2}J_z\dot{\gamma}^2\,. \tag{2}$$

In Eqs. (2), M is the mass of the worktable, $\dot{y}$ and $\dot{z}$ are the velocities along the y-axis and z-axis respectively. J_x, J_y and J_z are the moment of inertia of the worktable about the x-axis, y-axis and z-axis. $\dot{\alpha}$, $\dot{\beta}$ and $\dot{\gamma}$ are the angular velocities along the x-axis, y-axis and z-axis respectively.

Because the worktable is supported by a series of distributed mormal springs, for simpler analysis, we assume the spring and damping elements are continuous and use k_v, c_v, and k_h, c_h to denote the equibvalent stiffnesses and equibvalent danmping per unit area of the vertical and horizontal directions. the potetial energy of the system contributed by normal spring elements can then be given by

$$\begin{aligned}V=&\frac{1}{2}\int_{-\frac{l_l}{2}}^{\frac{l_l}{2}}[k_v(z-n_l\alpha-x\beta)^2+k_v(z+n_l\alpha-x\beta)^2]l_4dx+\frac{1}{2}\int_{-\frac{l_l}{2}}^{-\frac{l_l}{2}+l_2}[k_h(y-m_2\alpha+x\gamma)^2+k_h(y+m_2\alpha-x\gamma)^2]l_5dx\\&+\frac{1}{2}\int_{l_3}^{\frac{l_l}{2}}[k_h(y-m_2\alpha+x\gamma)^2+k_h(y+m_2\alpha-x\gamma)^2]l_5dx\,.\end{aligned} \tag{3}$$

The non-conservative force of the system contributed by normal damping elements can then be given by

$$Q=Q_v+Q_h=2c_vl_4\int_{-\frac{l_l}{2}}^{\frac{l_l}{2}}(\dot{z}-x\dot{\beta})dx+2c_hl_5\left[\int_{-\frac{l_l}{2}}^{-\frac{l_l}{2}+l_2}(\dot{y}-m_2\dot{\alpha})dx+\int_{l_3}^{\frac{l_l}{2}}(\dot{y}-m_2\dot{\alpha})dx\right]. \tag{4}$$

Where, Q_v and Q_h denote the non-conservative force along the z-axis and y-axis respectively, and can be achieved as follows based on the presented equivalent model.

$$Q_v=\int_{-\frac{l_l}{2}}^{\frac{l_l}{2}}[c_v(\dot{z}-n_l\dot{\alpha}-x\dot{\beta})+c_v(\dot{z}+n_l\dot{\alpha}-x\dot{\beta})]l_4dx\,. \tag{5}$$

$$Q_h=\int_{-\frac{l_l}{2}}^{-\frac{l_l}{2}+l_2}[c_h(\dot{y}-m_2\dot{\alpha}+x\dot{\gamma})+c_h(\dot{y}-m_2\dot{\alpha}-x\dot{\gamma})]l_5dx+\int_{l_3}^{\frac{l_l}{2}}[c_h(\dot{y}-m_2\dot{\alpha}+x\dot{\gamma})+c_h(\dot{y}-m_2\dot{\alpha}-x\dot{\gamma})]l_5dx\,. \tag{6}$$

Substituting Eqs. (2)-(4) into equation (1), the following five equations of motion can be derived from the Lagrange's method:

$$M\ddot{y}+l_1l_5c_h\dot{y}+l_1l_5k_hy=0\,. \tag{7}$$

$$M\ddot{z}+2l_1l_4c_v\dot{z}+2l_1l_4k_vz=0\,. \tag{8}$$

$$J_x\ddot{\alpha}+l_1l_5c_hm_2\dot{\alpha}+(2l_1l_4k_vn_1^2+l_1l_5k_hm_2^2)\alpha=0\,. \tag{9}$$

$$J_y\ddot{\beta}+\frac{1}{6}l_1^3l_4k_v\beta=0\,. \tag{10}$$

$$J_z\ddot{\gamma}+(\frac{l_1^3}{2}-3l_1^2l_2+3l_1l_2^2)l_5k_h\gamma=0\,. \tag{11}$$

From Equ. (7) of the displacement y along the y-axis, Equ. (8) of the displacement z along the z-axis, Equ. (9) of the angular displacement α about x-axis, Equ. (10) of the angular displacement β about y-axis, Equ. (11) of the angular displacement γ about z-axis, the frequencies of the worktable at horizontal, vertical, rolling, pitching and yawing vibration mode can be derived as follows:

$$f_h=\frac{1}{2\pi}*\frac{\sqrt{4Ml_1l_5k_h-l_1^2l_5^2c_h^2}}{2M}\,. \tag{12}$$

$$f_v=\frac{1}{2\pi}*\frac{\sqrt{2Ml_1l_4k_v-l_1^2l_4^2c_v^2}}{M}\,. \tag{13}$$

$$f_r=\frac{1}{2\pi}*\frac{\sqrt{4J_x(2l_1l_4k_vn_1^2+l_1l_5k_hm_2^2)-l_1^2l_5^2c_h^2m_2^2}}{2J_x}\,. \tag{14}$$

$$f_p=\frac{1}{2\pi}\sqrt{\frac{l_1^3l_4k_v}{6J_y}}\,. \tag{15}$$

$$f_y=\frac{1}{2\pi}\sqrt{\frac{(l_1^3-6l_1^2l_2+6l_1l_2^2)l_5k_h}{2J_z}}\,. \tag{16}$$

Conclusions

The dynamic properties and modelling techniques of the plastic-coated retangle-shaped slideway were investtigated theoretically. The purpose was to develop an analytical model concerning the stiffness and damping characteristics of plastic-coated slideways mounted on the worktable and saddle. An equivalent model of the plastic-coated slideway system was suggested based on the assembly of the worktable and saddle, and both stiffness and damping characteristics on vertical and horizontal directions were theoretically determined considering the worktable and the saddle as rigid bodies and connected with a series of spring and damping elements. To derive the governing motion equation of the rectangle-shaped plastic-coated slideway system, the Lagrange approach was used, and then the dynamic behavior of the carriage at vertical, pitching, yawing and rolling vibration modes were identified.

Acknowledgements

The authors are grateful to the Key National Science and Technology Projects of China, for support of this research work under grant no. 2009ZX04014-035.

Refereces

[1] Y. Ito: *Modular Design for Machine Tools*. New York: McGraw-Hill, 2008.

[2] T.W. Black. *Machine Tool Way Bearings: A Brief Guide to Theor Selection*. Machinery, 1966, p. 106-113.

[3] Y.S. Yi, Y.Y. Kim, J.S. Choi, *etl*. *Dynamic Analysis of A Linear Motion Guide Having Rolling Elements for Precision Positioning Devices*. Journal of Mechanical Science and Technology, Vol. 22 (2008), p. 50-60.

[4] G. Szwengier, T. Godunski, S. Berczynski. *Identification of Physical Parameters in Contact Joints Models of Machines Supporting Systems*. Advances in Engineering Software,Vol. 31 (2000), p. 149–155.

[5] G.D. ttagiu, M.D. Gafitanu. *Dynamic Characteristics of High Speed Angular Contact Ball Bearings*. Wear, Vol. 211 (1997), p. 22-29.

[6] K.P. Nair, V.P.S. Nair, N.H. Jayadas. *Static and Dynamic Analysis of Elastohydrodynamic Elliptical Journal Bearing with Micropolar Lubricant*. Tribology International, Vol. 40 (2007), p. 297-305.

[7] J.S. Choi, J.Yoo, Y.S. Yi, Y.Y. Kim, etl, *Vibration Analysis and Its Applicationof A Linear Motion Guide Supported by Rolling Ball Bearings*. Transactions of KSME A, Vol. 29, (2005), p. 955-963.

[8] Z.D. Xu, L.W. Ma. *Structure Dynamics*. Beijing: Science Press, 2007.

[9] R.A. de Callafon. *Lagrange's Method-application to The Vibration Analysis of A Flexible Structure*. University of California. http://maecourses.ucsd.edu/labcourse/handouts/Lagrange_handout.pdf.

doi:10.4028/www.scientific.net/AMM.50-51.42

Acoustic Source Localization via Advanced TDoA for Self-organization Wireless Sensor Networks

Xuanya Li[1, a], Linlin Ci[1, 2, b], Genyan Ge[2], Dawei Li[2]

[1] School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China

[2] Beijing Institute of Information Technology, Beijing, China

[a]xuanya0214@sina.com, [b]cilinlin@263.net

Keywords: WSNs; acoustic source localization; advanced TDoA;

Abstract. Wireless Sensor Networks (WSNs) enable susceptible sensing of the environment, offering unprecedented opportunities for observing the physical world. Acoustic source localization is an interesting topic with many possible application areas, such as intruder detection, sniper localization, automatic tracking of speakers and so on. Many existing algorithms are on the premise that the exact coordinates of sensor nodes are already known. In this paper, we propose a vector-based advanced TDoA algorithm that would calibrate the coordinate of the acoustic source by the non-prepositioned nodes. In the meanwhile, the portable sensor nodes would adjust themselves through the feedback of the estimated positions as well. Finally, we show that the proposed mechanism has high accuracy through experiments.

Introduction

Advances in technology have made it possible to build Wireless Sensor Networks (WSNs) using inexpensive nodes consisting of a low power processor, a modest amount of memory, a wireless network transceiver and a sensor board [1]. Large numbers of cheap and smart sensors, networked through wireless links, provide exciting applications in multifarious areas including object tracking, habitat monitoring, mineral reconnaissance, intruder detection.

In all these applications, sensor node locations are critical not only to the application's goals but also to the operations of WSNs. Self-organized localization capability is a highly desirable characteristic of sensor networks in respect that the measurement data are meaningless without knowing the location where the data are obtained. There exist well organized surveys on sensor localization in the [2,3,4]. All these works report early results on sensor localization. Jing [5] categorize the localization problems into three groups: proximity-based, range-based and angle-based. One common character of these methods is that the measurements used in the algorithms all depend on the mutual relationships of the nodes, not on the environment.

On the other hand, acoustic source localization is a hot research topic with many possible surveillance applications in which the target should be detected immediately. Energy-based ideas present novel approaches. M. Z. Rahman et al. [6] proposes a method for source localization using power spectral analysis and decision fusion in wireless sensor networks. X. Sheng et al. [7] uses acoustic signal energy measurements taken at individual sensors to estimate the locations of multiple acoustic sources. Furthermore, T. Ajdler et al. [9] use a Maximum Likelihood (ML) estimation method via the measurements of Time Difference of Arrival (TDoA). This propagation time based algorithm is proved to be dramatically more accurate than energy-only methods.

In this paper, for the purpose of acoustic source localization, we propose a vector-based advanced TDoA algorithm that would calibrate the coordinate of the acoustic source by the non-prepositioned nodes. The portable sensor nodes would adjust themselves through the feedback of the estimated positions as well. In other words, the sensor node localization is, to a certain degree, integrated into the acoustic source localization which could obviously reduce the complexity of the overall

localization system and decrease the energy consumption of the sensor nodes. We show that the proposed mechanism has high accuracy through experiments.

The remainder of this paper is organized as follows. In the next section, some system design essentials are presented. The following section describes the method which splits out into three steps. Experimental results are drawn then and the last section concludes the paper.

Design Essentials

Anchor node

In most applications of WSNs, the coordinates of points can be global, meaning they get several prior knowledge with regard to the localization, or relative, meaning that they are an arbitrary "rigid transformation" [3] (rotation, reflection, translation) away from the global coordinate system. When concerning the exact position of the target event such as the sound, several anchor nodes are indispensable. Anchor nodes are plebeian nodes that know their global coordinates priori.

One considerable issue is that we should plan the anchor node layout in advance. Some groups proved that localization accuracy improves if anchors are placed in a convex hull around the network. In our experiment, three non-collinear anchor nodes that place on the edge of the experimental district define a global coordinate system in two dimensions.

Time synchronization

Traditional solution to sound detection is to use microphone arrays which are usually controlled by the same console so that the local time can be the same after initialization. Conversely, time synchronization which aims at equalizing the local times for all nodes in the network addresses novel techniques. A common timescale for local clocks of wireless devices should be provided so that the sensor nodes coordinate their operations and collaborate to achieve a complex sensing task.

In [10], Reference Broadcast Synchronization (RBS) was introduced, where a receiver with receiver, instead of sender with receiver, syschronization is used. Here, we put forward an advanced RBS method to implement the syschronization, which, however, is not the point in this paper.

The Proposed Algorithm

In order to solve our problem of source localization, we use a 3 steps technique. The first step estimates the coordinate of acoustic second through the TDoAs of different anchor pairs. The second step finds the portable nodes positions through a series of estimation in Step 1. Finally, this system could be in use in the third step in which the source coordinate is calculated through both anchor nodes and portable nodes.

Algorithm based on Time Difference of Arrival

In the beginning, we could only compute the acoustic source by the anchor nodes for the portable nodes coordinates are not prior known. Once the TDoA of a sensor pair is obtained, we expect the source to be located on a hyperboloid whose foci are the two sensors positions [8].

Define d as the distance difference which derives from the TDoAs, ξ as the experimental errors. The 3 groups of hyperboloids can be written as follows:

$$\begin{cases} \sqrt{(x-x_a)^2+(y-y_a)^2}-\sqrt{(x-x_b)^2+(y-y_b)^2}=d_{ab}+\xi_{ab} \\ \sqrt{(x-x_b)^2+(y-y_b)^2}-\sqrt{(x-x_c)^2+(y-y_c)^2}=d_{bc}+\xi_{bc} \\ \sqrt{(x-x_c)^2+(y-y_c)^2}-\sqrt{(x-x_a)^2+(y-y_a)^2}=d_{ca}+\xi_{ca} \end{cases}$$

With the knowledge of each 2 groups of hyperboloids above a unique solution can be found. Therefore from the equations set there are 3 answers which should be identical ideally. In a practical situation, with ambient noise and reverberation the three estimation points will never intersect in a single place. We will then need to find the point that has the maximum likelihood to be the

intersection point and the source position. Simple barycenter measurement or least square method could be used.

Portable node calibration

The objective of the portable nodes beyond the anchors is to calibrate the localization with more measurement data, whereas their measurement data in step 1 are meaningless due to the unknown location where the data are obtained. Accordingly, we propose a backcalculation to get the portable position via a series of test estimation in step 1.

As in Figure 1(a), we get the range measurement d_{as} between the anchor node and the estimated source after every estimation. Meanwhile, d_{ap} derived from the TDoA stands for the distance difference. Therefore, the range d_{ps} between the portable node and the estimated source is $d_{as}-d_{ap}$.

Consequently, the portable node locates on the circle with center at the estimated source and radius equal to d_{ps}. At least three times repetitive executions of step 1 ascertain the portable by reason that theoretically, one unique point can be fixed through 3 circles.

In this practical situation, 3 circles would meet not in a single point but in a region (see Fig.1(b)) considering that it's unfaithful to use estimated value to execute an estimation process. The solution in this paper is to run step 1 repeatedly until the satisfying precision be reachable. Multiple different estimation would be utilized.

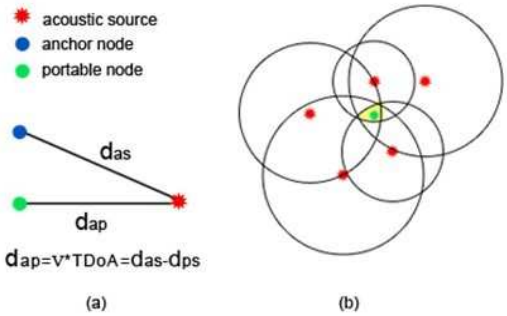


Figure 1. The relationship between the acoustic source and the portable node

Let i represent the index going over these n different sensor devices. The n circles which intersect at the region including the portable node can be shown as follows:

$$\begin{cases} (x_1-x)^2+(y_1-y)^2=d_1^{\,2} \\ (x_2-x)^2+(y_2-y)^2=d_2^{\,2} \\ \vdots \\ (x_n-x)^2+(y_n-y)^2=d_n^{\,2} \end{cases}$$

Each equation minus the last one:

$$\begin{cases} x_1^{\,2}-x_n^{\,2}+y_1^{\,2}-y_n^{\,2}-2(x_1-x_n)x-2(y_1-y_n)y=d_1^{\,2}-d_n^{\,2} \\ x_2^{\,2}-x_n^{\,2}+y_2^{\,2}-y_n^{\,2}-2(x_2-x_n)x-2(y_2-y_n)y=d_2^{\,2}-d_n^{\,2} \\ \vdots \\ x_{n-1}^{\,2}-x_n^{\,2}+y_{n-1}^{\,2}-y_n^{\,2}-2(x_{n-1}-x_n)x-2(y_{n-1}-y_n)y=d_{n-1}^{\,2}-d_n^{\,2} \end{cases}$$

Equations set above equals to the matrix representation of $AX=b_1$ which the definition of A, X, b_1 are as below:

$$A=\begin{bmatrix} 2(x_1-x_n) & 2(y_1-y_n) \\ \cdots & \cdots \\ 2(x_{n-1}-x_n) & 2(y_{n-1}-y_n) \end{bmatrix} \quad X=\begin{bmatrix} x \\ y \end{bmatrix} \quad b_1=\begin{bmatrix} x_1^{\,2}-x_n^{\,2}+y_1^{\,2}-y_n^{\,2}+d_n^{\,2}-d_1^{\,2} \\ \cdots \\ x_{n-1}^{\,2}-x_n^{\,2}+y_{n-1}^{\,2}-y_n^{\,2}+d_n^{\,2}-d_{n-1}^{\,2} \end{bmatrix}$$

Applying least square to the overdetermined equations set, we get the portable point:

$$\hat{X} = (A^T A)^{-1} A^T b_1$$

Each portable node coordinate can be finally got from this calibration. The more samples there are, the more accurate the estimated value is.

Vector-based estimation

Now both the anchor and portable nodes apply to the estimation. With the portable nodes participation, the estimated value would be much more perfect. Here we no longer use the hyperboloid method shown in step 1. Moreover, a vector-based estimation is brought forward.

The original coordinate system performs a rigid transition (rotation, reflection, translation) in order to situate an arbitrary node at coordinate origin. As shown in Figure 2, we define $\vec{r_s}$, $\vec{r_i}$, $\vec{r_{si}}$ as follows:

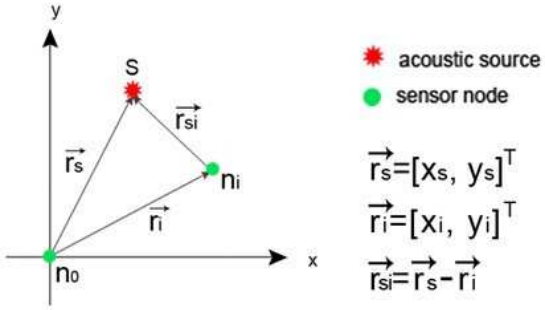


Figure 2. Vector transformation

d_i, the distance difference to S between n_i and n_0 is:

$$d_i = r_{si} - r_{s0} = r_{si} - r_s \Rightarrow r_{si} = d_i + r_s \tag{1}$$

In Equation (1), importing the vector notion we can get

$$(d_i + r_s)^2 = r_{si}^{\ 2} = \left|\vec{r_s} - \vec{r_i}\right|^2 \tag{2}$$

Moreover:

$$\begin{cases} (d_i + r_s)^2 = d_i^{\ 2} + 2d_i r_s + r_s^{\ 2} \\ \left|\vec{r_s} - \vec{r_i}\right|^2 = (\vec{r_s} - \vec{r_i})^T (\vec{r_s} - \vec{r_i}) = r_s^{\ 2} - 2\vec{r_i}^T \vec{r_s} + r_i^2 \end{cases} \tag{3}$$

According to (2)(3), we get：

$$r_i^2 - d_i^{\ 2} - 2d_i r_s - 2\vec{r_i}^T \vec{r_s} = 0 \tag{4}$$

Practically, experimental error ξ exists. Equation (4) depict as:

$$\xi = r_i^2 - d_i^{\ 2} - 2d_i r_s - 2\vec{r_i}^T \vec{r_s} \tag{5}$$

Applying Equation (5) to all nodes, there will be

$$\xi = \underbrace{\begin{bmatrix} r_1^2 - d_1^{\ 2} \\ r_2^2 - d_2^{\ 2} \\ \vdots \\ \vdots \\ r_n^2 - d_n^{\ 2} \end{bmatrix}_{n\times 1}}_{A_1} - 2\cdot \underbrace{\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_n \end{bmatrix}_{n\times 1}}_{A_2} \cdot r_s - 2\cdot \underbrace{\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix}_{n\times 2}}_{A_3} \cdot \vec{r_s} \tag{6}$$

Equation (6) equals to the representation below:

$$\xi = A_1 - 2A_2 r_s - 2A_3 \vec{r_s} = -2\left(\left(A_3 \vec{r_s} + A_2 r_s\right) - \frac{A_1}{2}\right) = -2\left(\begin{pmatrix} A_3 & A_2 \end{pmatrix}\begin{pmatrix} \vec{r_s} \\ r_s \end{pmatrix} - \frac{A_1}{2}\right) \tag{7}$$

We define:

$$A_5 = \begin{pmatrix} A_3 & A_2 \end{pmatrix}_{n\times 3}, \quad W = \begin{pmatrix} \vec{r_s} \\ r_s \end{pmatrix}_{3\times 1}, \quad b_2 = \frac{A_1}{2} \tag{8}$$

Combining (8) with (7), the final problem is to calculate the minimum value of $\|A_5 W - b_2\|_2^2$.

Due to parameter estimation, we get $\widehat{W}$:

$$\widehat{W} = (A_5{}^T A_5)^{-1} A_5{}^T b \tag{9}$$

As Equation (9) come out, $\vec{r_s}$ is got, which contains the final estimated acoustic source coordinate.

Experiment

Hardware platform

The hardware platform is built upon the MICAz mote devices manufactured by Crossbow Technology Incorporated running the TinyOS embedded operating system, a widely used component-based architecture targeting most wireless sensor network applications.

Sensor nodes (Figure 3) we used here consist of a processor/radio board, meaning a MICAz mote, and a sensor&data acquisition board, meaning the acoustic sensor board made by ourselves in our experiment. Another fundamental hardware is the base station/gateway (Figure 4) consists of a mote and a mote interface/gateway board (MIB 520 here) which is produced by Crossbow as well.

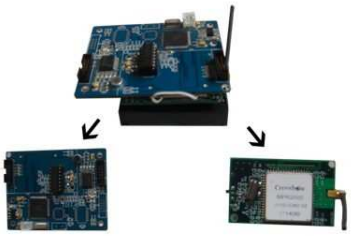


Figure 3. A sensor node.

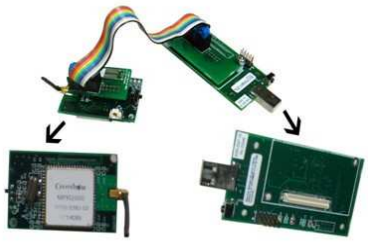


Figure 4. The gateway

Results analysis

To test the performance of the algorithm, we simulate the sound source in a room of dimensions of about 6*5*5 m. In out experimental setup, three anchor nodes are prepositioned on the border of the room to form a convex hull around the self-organization sensor network. Another three portable nodes are randomly deployed in the area.

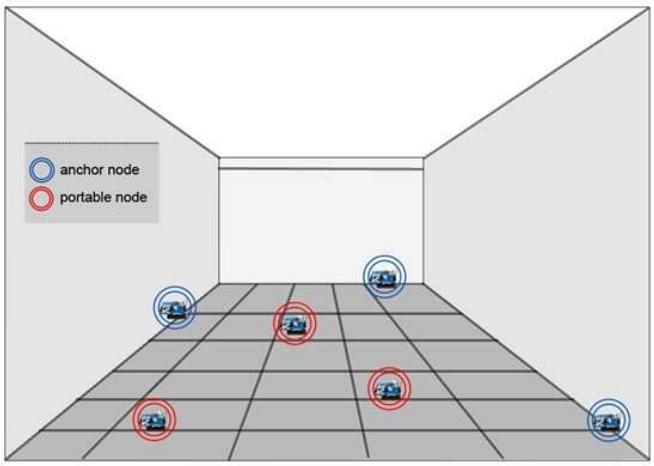


Figure 5. The experiment environment.

As the three steps in the algorithm, our analysis divides into three phases. In the first stage, the acoustic source position would be estimated only by the anchor nodes. The portable don't participate in calculate, notwithstanding they receive and transmit the data. As shown in Figure 6, the estimated values are reluctant to accept.

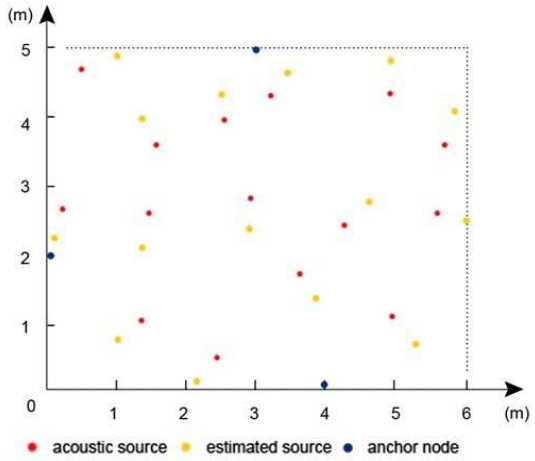


Figure 6. The 1st stage

The portable node coordinate are got in the second phase. As we mentioned, the more samples there are, the more accurate the estimated value is. The experimental results prove it in Figure 7.

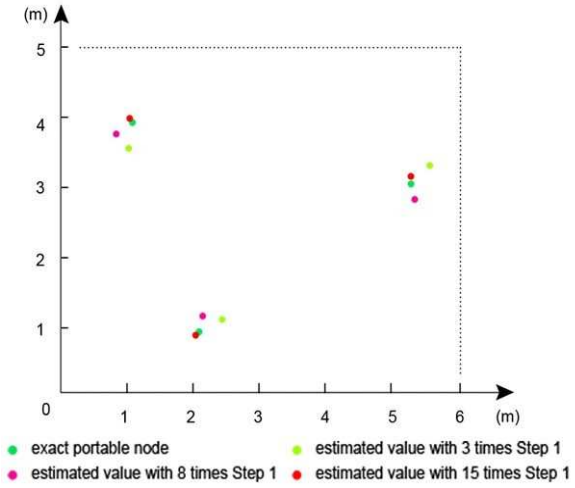


Figure 7. The 2nd stage

The third period is the time in which this system really put into use. All sensor nodes take part in cooperatively. The result (Figure 8) is much better than the result in the first stage which fully proves the effectiveness of the proposed algorithm.

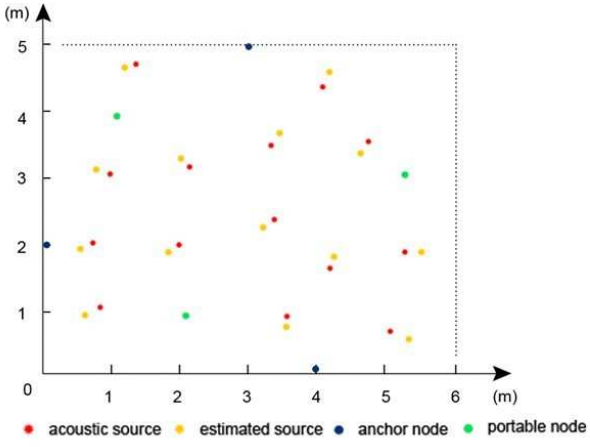


Figure 8. The 3rd stage

The still existent error is due to several reasons. The first issue is the time syschronization. The system is impossible to reach complete syschronization which a subtle discrepancy would result in centimeters to decimeters error. Secondly, our algorithm is based on the two-dimensional while our environment is three-dimensional which would lead to inaccurate TDoAs between each sensor pair. Besides, analysis of acoustic signals is essential, sampling and feature extraction are inevitable factors.

Conclusion

Localization is one of the fundamental problems in wireless sensor networks. For acoustic source localization, this paper presents an advanced algorithm which modifies the conventional TDoA estimation procedure to make it more accurate with more additional portable sensor nodes. Furthermore, the sensor node localization is integrated into the acoustic source localization which could dramatically reduce the algorithm complexity and decrease the network energy consumption. Our experimental results indicate the high precision of the proposed technique.

References

[1] J. Hill, R. Szewczyk, A. Woo, S. Hollar, D. Culler, and K. Pister. System architecture directions for networked sensors. In *Proceedings of ASPLOSIX*, 2000.

[2] A. Savvides, M. Srivastava, L. Girod, et al. Localization in sensor networks[M]//*Wireless Sensor Networks. Massachusetts: Kluwer Academic Publishers,* 2004: 327 – 349.

[3] J. Bachrach, C. Taylor. Localization in sensor networks[M]//*Handbook of Sensor Networks: Algorithms and Architectures. Hoboken: John Wiley & Sons,* 2005: 277 – 310.

[4] G. Mao, B. Fidan, B. D. Anderson. Wireless sensor network localization techniques[J]. *Computer Networks*, 2007, 51(10): 2529 – 2553.

[5] J. Wang, R. Ghosh, and S. Das. A survey on sensor localization, *Journal of IET CTA*, vol. 8, pp. 2{11, February 2010.

[6] M. Z. Rahman, G. C. Karmakar and L. S. Dooley, "Passive Source Localization using Power Spectral Analysis and Decision Fusion in Wireless Distributed Sensor Networks", *Proceedings of the International Conference on Information Technology: Coding and Computing* (ITCC'05), 2005.

[7] X. Sheng and Y. H. Hu, "Maximum Likelihood Multiple-Source Localization Using Acoustic Energy Measurements with Wireless Sensor Networks", *IEEE Transactions on Signal Processing,* vol. 53, no. 1, Jan. 2005.

[8] T. Ajdler, I. Kozintsev, R. Lienhart, and M. Vetterli, "Acoustic Source Localization in Distributed Sensor Networks", *Proceedings of the 38th Asilomar Conference on Signals, Systems and Computers*, vol. 2, Nov. 2004, pp. 1328–1332.

[9] F. Sivrikaya, and B. Yener, "Time Synchronization in Sensor Networks: A Survey", *IEEE Network*, vol. 18, no. 4, July-Aug. 2004, pp. 45-50.

[10] J. Elson, L. Girod, and D. Estrin, "Fine-Grained Time Synchronization using Reference Broadcasts," *Proceedings of the Fifth Symposium on Operating Systems Design and Implementation (OSDI 2002),* Boston, MA, December 2002.

doi:10.4028/www.scientific.net/AMM.50-51.49

A (*t, n*) Threshold Signature Scheme Based on Factorial Decompose Theorem of Polynomial

Yongquan Cai[a], Fulai Cheng[b]

College of Computer Science and Technology, Beijing University of Technology, Beijing, China

[a]cyq@bjut.edu.cn, [b]s200807005@emails.bjut.edu.cn

Keywords: Conspiracy attack; Threshold signature; Factorial decompose theorem; Traceability

Abstract. The traditional threshold signature scheme is a combination of the digital signature scheme and the secret sharing scheme. Any group of *t* (threshold value) or more players can reconstruct the secret together but no group of fewer than *t* players can in secret sharing scheme. Because of this property of secret sharing scheme, the current threshold signature scheme has a natural weakness in the anti-collusion attack. A new threshold signature scheme based on the factorial decompose theorem is designed successfully without the secret sharing scheme in order to get the ability of anti-collusion attacks. It not only has the property of group signature and threshold signature, but also has some good features including traceability，robustness and resisting forgery attacks.

Introduction

In a (*t, n*) threshold signature scheme which is firstly introduced by Desmedt and Frankel [1] in 1991, a signature key is shared among *n* players. And a signature can be generated only when the number of participators is equal to or more than the threshold value *t*, but any less than *t* players cannot generate a valid one. Then anyone who knew the public key can verify the authenticity of the message and if necessary, the signature also can be opened, so that the original signers can be traced without revealing the secret keys.

Many such schemes [1, 2, 3, 4, 5, 6, 7, 8] have been proposed. However, these traditional threshold signature schemes are based on digital signature scheme and secret sharing scheme. At Crypto'93, Li et al [2] pointed out that the threshold signature scheme suffers from the conspiracy attack, which means that any *t* or more group players can cooperate to reveal the group's secret key with high probability. Since then, design a new (*t, n*) threshold signature scheme withstanding the conspiracy attack has been a hot issue in cryptography.

In 2003, Bing Wang et al [5] proposed a new threshold signature scheme based on the Joint Secret Sharing (JSS) scheme in order to against conspiracy attack. In this scheme, every player chooses two keys $f_i(0)$ and k_i, and distributes the shares of $f_i(0)$ to others by Shamir's secret sharing scheme and keeps k_i in secret, such that the group's secret key is $\sum_{i=1}^{n}(f_i(0)+k_i)$.But a conspiracy attack and a forgery attack is proposed respectively in literature [6] and [7] to show that this scheme is insecure. Wenfang Zhang et al [8] designed a new scheme by improving Wang's. In the new one, the collaborators cannot get the other players' secret key, but they can still reveal the group secret key. Therefore, the schemes based on JSS still suffer from conspiracy attack by now.

Xinglan Zhang [9] proposed a new group signature scheme based on factorial decompose

theorem of polynomial with factors in finite field F_q in 2010. Even though this group scheme does not meet non-related requirements, it provides a new basis for the study of threshold signature scheme withstanding conspiracy attack. We proposed a secure (*t, n*) threshold signature scheme by improving the group scheme in literature [9] and we changed the group public key generation method, threshold signature generation and so on. Because of the useless of secret sharing scheme, this scheme can resist conspiracy attacks. And it not only has the property of group signature and threshold signature, but also has some good features including traceability，robustness and resisting forgery attacks.

Threshold Signature Scheme Based on Factorial Decompose Theorem of Polynomial

We assume that $U=\{U_1,U_2,...,U_n\}$ is the set of group of n players and the Group Manager (GM) is a trusted center. This scheme consists of five phases.

System Initialization Phase

GM chooses two large primes p, q such that, $2^{l-1}<p<2^l$, $512\leqslant l\leqslant 1024$, and $2^{159}<q<2^{160}$. Let g be a generator of a multiplicative subgroup of Z_p with order q and $H(\cdot)$ denotes a one-way collision resistant cryptographic hash function.

Every $U_i(i=1,2,...,n)$ in the group U chooses a random number $x_i\in Z_q^*$ and takes it as his secret key. Then, U_i computes his corresponding pubic key $y_i\equiv g^{xi} \pmod p$ and sends it to GM. On receiving all of the y_i, GM determines whether any two public keys y_i and y_j are equal. If $y_i=y_j$，$i\neq j$, GM asks U_j to choose another $x_j\in Z_q^*$. Otherwise, GM records y_i from player U_i.

Group Public Key Generation Phase

GM chooses two random parameters $r\in Z_q^*$, $K\in Z_q$ and an irreducible polynomial $g(x)\in F_q(x)$, computes $R\equiv g^r \pmod p$ and r' such that $rr'\equiv 1 \bmod q$. Next, GM takes any t elements from the collection of public keys $\{y_1,y_2,...,y_n\}$, computes

$$Y_i=\prod_{j=1}^{t} y_{ij}, y_{ij}\in\{y_1,y_2,\cdots,y_n\}, i=1,2,\cdots,C_n^t.$$

If $Y_i=Y_j$，$i\neq j$, GM breaks off the protocol and restarts it. Else, GM constructs a polynomial

$$f(x)=K+[g(x)\prod^{n}(x-Y_i^r)(\bmod p)](\bmod q) \quad (1)$$

After completion of these operations, $\{R, f(x), K\}$ will open as the group public key and $\{r, r'\}$ will not as the private key.

Threshold Signature Generation Phase

Assume that the subset $\{U_{j1}, U_{j2},..., U_{jn}\}$ of t group players wants to sign a message m on behalf of the group U, they perform the following steps to generate the signature. Anyone in the subset can be the designated clerk (DC) and we assume DC is U_{j1}.

(1)Each player U_{j1} chooses a random number $k_i\in Z_q^*$, computes $K_i=g^{k_i} \bmod p$, $K_i'=R^{k_i} \bmod p$ and broadcasts K_i to all the players in the subset. On receiving all of the K_i, each U_{ji} computes $K'=\sum_{i=1}^{t} K_i$. Then, U_{ji} computes the partial signature

$$S_{ji}=x_i\times H(m,K')-K'\times k_i \bmod q \quad (2)$$

and sends the partial signature (S_{ji}, K_i') to the DC U_{j1}.

(2)On receiving all of the partial signatures, the DC U_{j1} authenticates the validity of the partial signature (S_{ji}, $K_i^{'}$) by checking the following equation

$$g^{S_{ji}} K_i^{K^{'}} \equiv y_{ji}^{H(m,K^{'})} \bmod p \qquad (3)$$

If all the equation hold, the DC computes $K^{''} = \prod_{i=1}^{t} K_i^{'} \bmod q$, $S = R^{\sum_{i=1}^{t} S_{ji}} \bmod q$ and $\{m, S, K^{'}, K^{''}\}$ is the group signature for the message m. Else, (S_{ji}, $K_i^{'}$) is a forged partial signature.

Threshold Signature Verification Phase

On receiving the threshold signature $\{m, S, K^{'}, K^{''}\}$ from the DC, the verifier uses the public key to authenticate the validity of the group signature. The Verifier computes

$$Y_j^{'} = [SK^{''K^{'}}]^{H^{-1}(m,K^{'})} \bmod p \qquad (4)$$

then, checks the equation $K=f(Y_j^{'}) \bmod q$. The threshold signature is accepted if and only if $K=f(Y_j^{'}) \bmod q$.

Signer Identify Verification Phase

In case of disputes, the GM can use the secret key $r^{'}$ and $Y_j^{'}$ calculated according to Eq. 4 to compute $Y_j^{''} = (Y_j^{'})^{r^{'}} \bmod p$. So, the DC can trace back to find who the t players are from $Y_j^{''} = \prod_{i=1}^{t} y_{ji}, y_{ji} \in \{y_1, y_2, \cdots, y_n\}$.

Security Analysis of the Threshold Signature Scheme

The scholar G.L. Wang pointed out that a security threshold signature scheme should follow some basic properties: (1) any t players in the group can generate a valid group signature, but any t-1 or fewer players cannot, while they do not reveal any information about their secret keys and the group's secret key; (2) any verifier can verify the group signature with the group public key, but they cannot determine whether a player was involved in signing; (3) in case of disputes, the DC can check the signers. So, we will discuss our scheme according to those properties.

Theorem 1. Any t or more players in the group can generate a valid group signature.

Proof. The partial signature generated according to Eq. 2 satisfies Eq. 3.

$$g^{S_{ji}} K_i^{K^{'}} = g^{x_i \times H(m,K^{'}) - K^{'} \times k_i} g^{k_i K^{'}} \bmod p = g^{x_i \times H(m,K^{'})} \bmod p = y_{ji}^{H(m,K^{'})} \bmod p .$$

The group signature satisfies Eq. 4.

$$Y_j^{'} = [SK^{''K^{'}}]^{H^{-1}(m,K^{'})} \bmod p = [R^{\sum_{i=1}^{t} S_{ji}} (\prod_{i=1}^{t} K_i^{'})^{K^{'}}]^{H^{-1}(m,K^{'})} \bmod p$$

$$= [R^{\sum_{i=1}^{t} (x_i \times H(m,K^{'}) - K^{'} \times k_i)} (R^{\sum_{i=1}^{t} k_i})^{K^{'}}]^{H^{-1}(m,K^{'})} \bmod p$$

$$= [R^{\sum_{i=1}^{t} (x_i) \times H(m,K^{'})}]^{H^{-1}(m,K^{'})} \bmod p = g^{(\sum_{i=1}^{t} x_i) \times r} \bmod p$$

If the players were t or more players in the group, obviously, the result of Eq. 1 is K when $x=Y_j^{'}$. Else, the result is not equal to K according to the factorial decompose theorem of polynomial with factors in finite field F_q. So any t players in the group can generate a valid group signature, and any t-1 or fewer players cannot.

In case of disputes, the $Y_j^{'}$ can be opened. We know $rr^{'} \equiv 1 \bmod q$, so $(Y_j^{'})^{r^{'}} = g^{(\sum_{I=1}^{t} x_i)\times r\times r^{'}} \bmod p = g^{(\sum_{I=1}^{t} x_i)} \bmod p = \sum_{I=1}^{t} y_i$.

The threshold signature scheme has very strong security. We will discuss several possible attacks and none of these attacks can successfully break our new scheme.

(1)The adversary cannot forge a partial signature satisfying Eq. 3.

Suppose an attacker tries to imitate a player U_i of the group and generate a legitimate partial signature. According to Eq. 2, the attacker needs to get the player's secret key x_i. If the attacker uses the player's public key $y_i \equiv g^{x_i} (\bmod p)$ to compute x_i, he needs to solve the discrete logarithm problem (DLP), which is impossible by now. Even if the attacker got a partial signature $S_{ji}=x_i\times H(m,K^{'})-K^{'}\times k_i (\bmod q)$, an equation with two unknowns k_i and x_i should be used to calculated x_i. So this attack cannot work successfully. Furthermore, this scheme can resist substitution attack [8].

(2)The adversary cannot forge a group signature of message m satisfying Eq. 4.

Suppose an attacker tries to forge a group signature, then, he needs to make the equation $K = f(Y_j^{'}) \bmod q$ correct, which means that the signature must contain t legal secret keys because the equation $f(x)$ constructed and issued by the GM and cannot be tampered by attacker. We had known that to compute the players' secret key according to public information is equivalent to solving the discrete logarithm problem. So this attack cannot work successfully.

(3)Any t or more players cannot cooperate to obtain another player's secret key, let alone the group's secret key.

To avoid the conspiracy attack, each player chooses his secret key randomly in our scheme and this action has nothing to do with other players. Any t or more players can only cooperate to sign a message and they cannot compute any useful information about the system and another player's secret key. So the effectiveness of the cooperation of several players or adversaries has no difference with a single adversary's attack effect. Therefore our new (t, n) threshold signature scheme can resist collusion attacks.

Theorem 2. The scheme meets the anonymity and traceability.

Proof. $\{m, S, K^{'}, K^{''}\}$ is the valid group signature in our scheme. $K^{'}$ and $K^{''}$ are both calculated according to t random numbers which was selected by the players. And $Y_j^{'}$ can be calculated, but it cannot be opened because no one know $r^{'}$, the system secret except the DC. So the verifier couldn't determine the players who signed the message. In case of disputes, the GM uses $r^{'}$ to compute $(Y_j^{'})^{r^{'}}$ (mod p) which is equal to the product of t signers' public key. Then the GM can check the signers.

Conclusion

In this paper, we have constructed a new (t, n) threshold signature scheme based on factorial decompose theorem of polynomial and have proved its correctness. Because of the useless of secret sharing scheme, this scheme can resist conspiracy attacks. Besides, it not only has the property of group signature and threshold signature, but also has some good features including traceability，robustness and resisting forgery attacks.

Acknowledgment

This work was supported by the Beijing Municipal Natural Science Foundation (Grant No. 1102003) and the National Basic Research Program of China (Grant No. 2007CB311106).

References:

[1] Y. DESMEDT and Y. FRANKEL: Shared generation of authenticators and signatures. In: Proceedings of Crypto'91, Santa Barbara, California, USA, 1991. p. 457-469.

[2] C.M. LI, T. HWANG and N.Y. LEE: Remark on the threshold RSA signature scheme. In: Advances in Cryptology-Crypto'93 Proceedings, Berlin. 1994, LNCS773: p. 413-420.

[3] C.T. Wang, C.H. Lin and C.C. Chang: Threshold signature schemes with traceable signers in group communications. *Computer Communications*, 21(8), 1998, p.771.

[4] Q.L. XU: A modified threshold RSA digital signature scheme. *Chinese Journal of Computers*, 23(5), 2000, p.449.

[5] B. WANG, J.H. LI: (t, n) threshold signature scheme without a trusted party. *Chinese Journal of Computers*, 26(11), 2003, p.1581.

[6] Q.XIE and X.Y. Yu: A new (t, n) threshold signature scheme withstanding the conspiracy attack. *Wuhan University Journal of Natural Sciences*, 10(1), 2005, p.107.

[7] L.F. GUO, X.G. CHENG: Security analysis of a (t, n) threshold signature scheme without a trusted party. *Chinese Journal of Computers*, 29(11), 2006, p.2031.

[8] W.F. ZHANG, X.M.WANG, D.K. HE: Security analysis of the (t,n) threshold signature scheme without a trusted party and its improvement. *Journal of the China Railway Society,* 30(3), 2008, p.40.

[9] X.L. ZHANG: An efficient group signature scheme based on factorial decompose theorem of polynomial with factors in finite field. *Microelectronics & Computer,* 27(3), 2010, p.163.

[10] G.L.WANG, S.H.QING: Weaknesses of some threshold group signature schemes. *Journal of Software*, 11(10), 2000, p.1326.

doi:10.4028/www.scientific.net/AMM.50-51.54

A Smooth Flow of Traffic Circle Model

Nan Ji [1, a], Jie Zhang [2] and Yan Gao [3]

[13] College of Science, Hebei United University,Tangshan, Hebei Province,China

[2] Tangshan Tanggang Expressway Management office,Tangshan, Hebei Province,China

[a]pumpkinji@163.com

Keywords: traffic circle, control model, simulation

Abstract. This article sets up mathematic model for the traffic circle to determine how best to control traffic flow. In order to make the traffic at the circle intersection smooth in maximum, this model figures out the largest traffic capacity through Wardrop's formula. When the signal control is needed, a linkage time solving model is set up, which proves the number of seconds of every green light and work out the delay time. At last, the simulation software VISSIM is used for an example to work out the average delay time under the signal control. The result matches with this model.

Introduction

Traffic circles, were invented in the early 20th century and built in cities throughout Europe and the United States. Since 1903, Eugenl Hienard proposed that an island should be established at the joint of the roads and let all the vehicles run around the central island. Inside the interweave sections of roundabout intersections, the increase of the traffic volume, the converging and diverting of the vehicles lead to the disorder of traffic volume inside the roundabout. It is also possible that traffic accidents may happen due to the vehicles' forced driving in and out of circled road which result in the bottle-neck section of the roundabout. This article aims to find a best solution to relieve the roundabout's traffic pressure and guarantee the maximum smoothness of the roundabout's crossings in different situations.

Assumptions

To simples and generalize the model, an ideal traffic condition is given, in which only the main factors are considered.

◇the standard vehicle type as the basic measuring unit while measuring the passing volume

◇right driving habit is considered

◇only vehicles are considered, pedestrians and non-vehicles can go through the overpass or underground passage

Model Analysis and Development

Classification by Control: Types According to the diameter of the central of the circle, the types of traffic circle can be classified into three types[3]:

(1) Regular traffic circle: to the diameter of the center of the circle exceed 25m, weaving section is long and the entrance roads are not expanded into a bellbottom shaped one.

(2)Small traffic circle: the diameter of the center of the circle is 4~25m, the entrance roads are

expanded into a bellbottom shaped one, which is convenient for the vehicles driving into.
(3) Mini traffic circle: the diameter of the center of the circle is less than 4m, which is not necessarily to be the circular.

In usual situation, mini traffic circles are mostly single channels or other simple general ways, so in this paper, we don't consider this type, in regard to the regular and small traffic circle, according to the actual crossing capability and the maximal crossing capability, we judge that the traffic control system should be installed or not.

Calculating the passing ability The passing ability of interweave section means that the max number of vehicles incoming the cross from each entering road per hour, the unit is veh/h. To determine the passing ability of circle with traffic light, it is only needed to determine the passing ability of each entering road. We introduce the following Wardrop formula to calculate the passing ability. [4]

$$Q_m = \frac{354W\left(1+\frac{e}{W}\right)\left(1-\frac{P}{3}\right)}{1+\frac{W}{l}} \tag{1}$$

In which, Q_m be the max passing ability of interweave section (veh/h), e be the mean width of entering approach road of the circle (m), l and W be the length and width of interweave section (m) respectively, $e=(e_1+e_2)/2$ (m), where e_1 be the width of entering approach road (m), e_2 be the width of overhang parts of circle, P be the ratio of interweaving vehicles and all vehicles, denoted by percentage, symbols in the formula were shown in Fig 1.

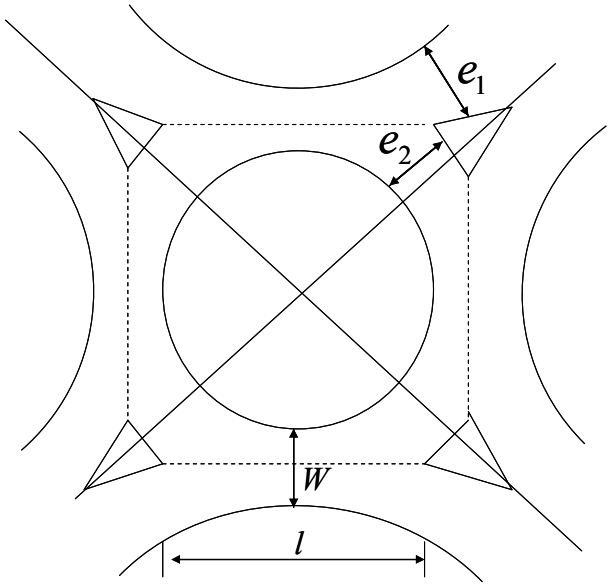


Fig 1 The geometric parameters in the circle

The control model without traffic lights. When the vehicle volumes at every entrance is lower than the volumes saturated itself, the traffic circle don't need outer control, we just need to let the vehicle obey the yield rules when rounding the circle, so that the whole traffic circle will keep in order.[1]

We take some weaving section in the circle for example, by taking the weaving section as the straight line. We take the weaving points as serving windows. Vehicles converge from different

directions and wait for passing through.

By analogy， we can get the mean delay time of the vehicles in any class

$$K_M = \frac{\sum_{i=1}^{M}\lambda_i}{\lambda_M}\overline{K} - \frac{\sum_{i=1}^{M-1}\lambda_i K_i}{\lambda_M} \quad (2)$$

In which, K_i stands for the mean delay time of the i-th priority right in system. $\overline{K}$ is the mean delay time of the all classes in system.

Because the method can avoid drivers' rushing out between different lanes, we set "yield" sign in entrance or exit of the circle, let the vehicles in higher class pass through first, which made the traffic circle driving in order, so the delay time calculated have the satisfactory effects.

The control model with traffic lights. Because traffic jam happens at the roundabout crossing, our purpose is to reduce the number of vehicles from different directions meeting at the weaving section by setting the traffic lights at every entrance to control the vehicles' entering time.

The period is the time when the red and green traffic lights finish their shining once and considered as an entire unit of the whole signal system. In the traffic light controlling system, the period can also be divided into the best period, the minimum actual green light time and the maximum actual green light time.

The best period is the time which causes the minimum delay when vehicles from different directions passing through the crossings. [2]

$$C_0 = \frac{1.5L + 5}{1 - Y}(s) \quad (3)$$

Where, $L = (nl_s + R)$ is the total loss time of every period; l_s is the average loss time of every phase. $Y = \sum_{i=1}^{M} Y_i$ is the sum of Y_i of every crossing's phase, Y_i is crossing's flow to Saturation. The minimum valid green light time of the main crossing as following,.

$$t_{A1} = \frac{Y_{11}}{Y}\left(C_0 - l_s\right) \quad (4)$$

$$t_1 = \frac{\left(T - t_2 \times 2\right)}{2} + t_3 \quad (5)$$

Where, Y_{11} is the crossing A's flow to saturation flow ratio, T is the period of traffic light, t_1 is the interval time of green light shining, t_3 is the yellow lights shining time.

According to the above, we can calculate the time of traffic lights. In figure 3, light A is the green light and light B is about to turn red. Because light A does not turn red immediately, vehicles still continue running a period of time Δt, Assume the radius of the roundabout crossing is r_0, the width of lane is ψ, the minimum distance the straight-through vehicles on the outside lane pass the

adjacent joint section $s = \frac{\pi\alpha_{ij}}{180}\left(r_0 + \frac{1}{2}\psi\right)$, So the interval time of green lights $\Delta t = \frac{s}{v}$

Under the condition in which the vehicles' arriving rate and the passing ability of the crossing remain the constant number, the delay of vehicles and the vehicles' arriving rate form the linear relationship. As in figure, vehicles' arriving rate fluctuates at random from one period to another period. In some certain period of time, saturation rate may be contemporarily over 1 due to the arbitrary fluctuation of the vehicles arriving rate, although the general average saturation rate is less than 1 during the whole period of time. Here according to Webster's average delay formula. [4]

$$t_w = \frac{T\cdot(1-\lambda)^2}{2\times[1-\lambda X]} + \frac{X^2}{2Q(1-X)} - 0.65\times(\frac{T}{Q^2})^{\frac{1}{3}}x^{(2+5\lambda)} \tag{6}$$

In which, t_w is the average delay of every vehicle (s); T is the time of period (s); λ is the ratio of valid green light time and the time of period ; X is the saturation; Q is the arriving rate (veh/s).

In the formula, the first part is the normal interval time of green lights caused when the vehicles' arriving rate remains unchanged. The second part and the third part are the additional interval time of green lights caused by the random fluctuation of the vehicles' arriving rate. When the saturation is low, the proportion taken by the second part and the third is small, but with the increase of the saturation, their influence on the formula's result increases.

Simulation Design

The traffic condition of roundabout crossings can be measured by the average delay time. The data of a roundabout crossing in Beijing is simulated. The following statistics is the passing volume of this roundabout crossing at rush-hour provided by Beijing Communication Bureau.

Table1. The list of directions

direction / items	east	west	south	North
vehicle	3620	3523	3940	3672

According to the wardrop equation, the maximum passing ability of interweave section is calculated using the MATLAB software.

Table 2. The list of directions

items / direction	R0	L	W	W/L	el	e2	e	E/W	Qm
East	30	45	17	0.37778	10.3	22	16.15	0.9706	4060
West	30	49	17	0.34694	16	14	15	0.8824	4009
South	30	40	16	0.4	10.4	22	16.2	0.9529	3881
north	30	42	16	0.38095	21.6	14	17.8	0.9412	4130

As can be seen, the maximum passing ability of the southern entrance Qm is less than its passing volume at rush hour. So it should be given passing priority.

The data of the interval time between red lights and green lights is calculated using MATLAB software.

Table 3.The list of directions

Items / direction	V	Ls	α	Period of traffic lights	Actual green light time	Interval time between green lights
East	34	8	90	76	36	1.4
West	34	8	90	76	36	1.4
South	34	6	90	76	42	1.4
north	34	6	90	76	42	1.4

The average delay time is calculated according to the above data d=5s

The simulated effect chart is as following:

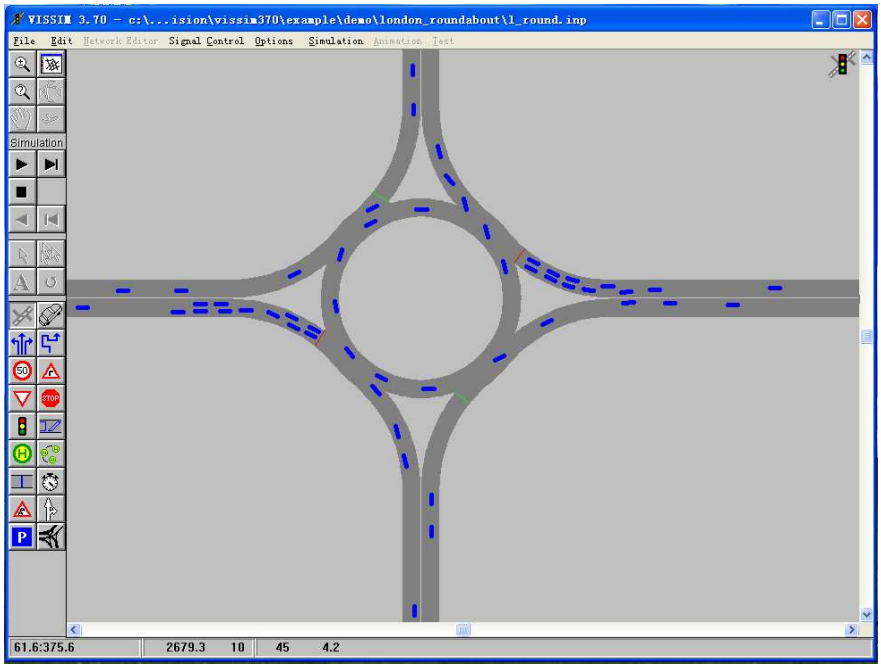


Fig3 The list of Simulation effect

The result of the simulation is consistent with the calculation result, which shows our model is feasible.

Acknowledgments. This paper is supported by Scientific Fund Project of Hebei Polytechnic University (z201004).

Reference

[1] GUO Ruijun, LIN Boliang *Traffic Operation Performances at Roundabout Weaving Sections,* Journal of Transportation Systems Engineering and Information Technology,pp.29-34(2010)

[2] Zheng rong Lu, *Design of Gang Control of Traffic Lights at Circular Intersection*, China Municipal Engineering, No. 2 ,pp.10-11(2002)

[3] Kuanmin Chen, Bao-jieYan, *Road Traffic Capacity Analysis*, People's Communications Publishing, Beijing (2003)

[4] Dianhai Wang, *The Theory of Traffic Flow*, People's Communications Publishing(2002)

doi:10.4028/www.scientific.net/AMM.50-51.59

Modal Analysis of Magnetic Suspended Control Moment Gyroscope

Wei Zhao[1, a], Libin Zhao[1,b], Jianyu Zhang[2,c] and Bangcheng Han[3,d]

[1]School of Astronautics, BeiHang University, Beijing 100191, China

[2]School of Aeronautic Science and Engineering, BeiHang University, Beijing 100191, China

[3]School of Instrument Science and Opto-electronics Engineering, BeiHang University, Beijing 100191, China

[a]zhaowei_138@sina.com, [b]lbzhao@buaa.edu.cn, [c]jyzhang@buaa.edu.cn, [d]hanbangcheng@buaa.edu.cn

Keywords: control moment gyroscope, modal analysis, finite element, ANSYS

Abstract. 3D finite element models of magnetic suspended control moment gyroscope structure are established by general finite element analysis software. Initial modal analysis scheme of the MS-CMG structure, in which the frame and gyro-chamber are analyzed separately, is described. A new modal analysis scheme that the MS-CMG is analyzed as a whole structure is presented. The connection between the gyro-chamber and frame is realized by defining a revolution pairs. The numerical predictions from two modal analysis schemes show obvious difference. The new scheme is more reliable since it simulates the real wok status of MS-CMG.

Introduction

Control moment gyroscopes (CMGs) are important actuators of attitude control systems in large spacecrafts due to their advantages of small volume, light mass, low power and long life. With the development of magnetic suspended technology, magnetic suspended CMGs (MS-CMGs), which provide high precision and eliminate contact by magnetic bearings between the gyro-chamber and gyroscope rotor, thus are increasingly used in spacecrafts. As a mechanism with high rotating speed, the structural dynamics behavior of MS-CMGs is very important for design [1,2]. Modal analysis could provide understanding on the structural dynamics characteristics, thus is performed in this paper using general software ANSYS®.

Basic theories of modal analysis

The purpose of modal analysis is to get the natural frequencies and modal shapes. The control equation of the system is:

$$\boldsymbol{M}\ddot{\boldsymbol{U}} + \boldsymbol{C}\dot{\boldsymbol{U}} + \boldsymbol{K}\boldsymbol{U} = \boldsymbol{F} \,. \tag{1}$$

where $\boldsymbol{U}$ is nodal displacement vector. The $\boldsymbol{M}$, $\boldsymbol{C}$ and $\boldsymbol{K}$ are mass matrix, damping matrix and stiffness matrix, respectively. $\boldsymbol{F}$ is force vector. If $\boldsymbol{F}$ is equal to $\boldsymbol{0}$, the free vibration equation of the structure is obtained. Generally, the damping can be ignored for the modal analysis. Thus, the control equation could be rewritten as:

$$\boldsymbol{M}\ddot{\boldsymbol{U}} + \boldsymbol{K}\boldsymbol{U} = 0 \,. \tag{2}$$

Then the basic equation of modal analysis could be obtained:

$$\boldsymbol{K}\boldsymbol{\phi} = \omega^2 \boldsymbol{M}\boldsymbol{\phi} \,. \tag{3}$$

where ω^2 is the eigenvalue, $\boldsymbol{\phi}$ is the modal vector.

It can be seen that eigenvalues and modal vectors are dependant on the structural stiffness and mass distribution. Therefore, the stiffness and mass distribution description in a structure are very important for accurately dynamic simulation.

The structural composing of MS-CMG

The MS-CMG is made up of gyroscope rotor, gyro-chamber, magnetic bearings, gimbal axes, frame, etc. Its connection relation and configuration are shown in figure 1. The gyroscope rotor is connected with gyro-chamber by magnetic bearings. Under working state, magnetic forces could be generated between the two components and keep the gyroscope rotor suspending and rotating at speed of 30000rpm. The gyro-chamber is fixed on gimbal axes, which are connected with frame structure by mechanical bearings. Thus, the gyro-chamber could rotate about the centerline of the mechanical bearings. Furthermore, the frame is fixed on the satellite by four bolts [3].

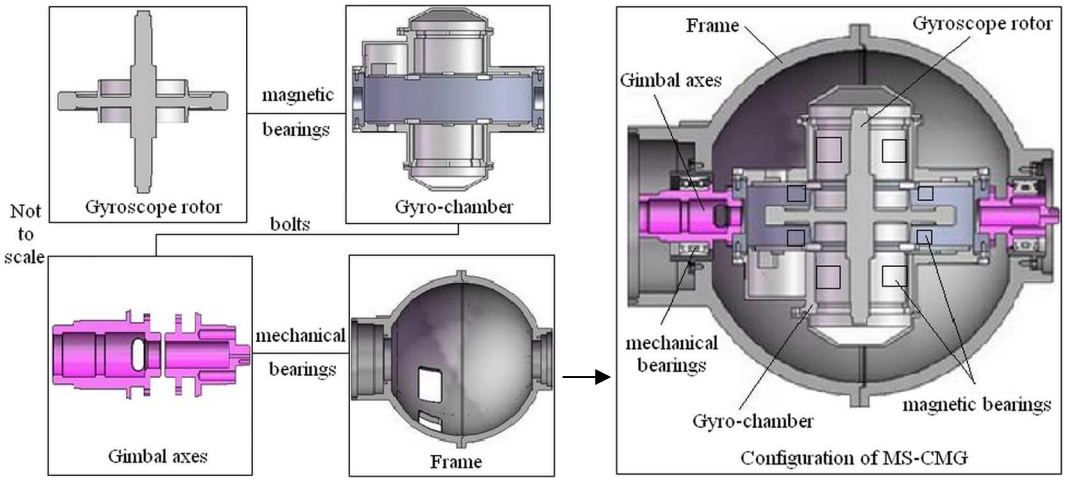
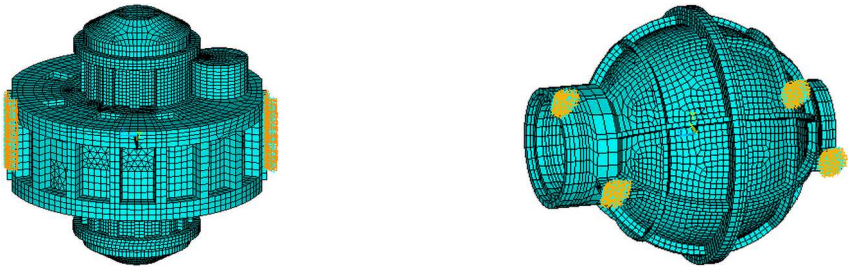


Figure 1. Schematic configuration of MS-CMG and corresponding connection relationship

Initial modal analysis scheme and results

Due to the rotation relation between the gyro-chamber and frame is difficult to simulate, a modal analysis scheme, which divided the MS-CMG into two components: the frame and the gyro-chamber with gyroscope rotor, is presented in reference [4,5]. Two separately modal analyses were performed to get knowledge of the dynamic behavior of MS-CMG.

Figure 2 shows the finite element models of gyro-chamber with gyroscope rotor and frame. The geometric model of gyroscope rotor, gyro-chamber, and frame were established respectively. Some little chamfers and bolt holes were neglected according to the basic principle of finite element modeling. In view of modal analysis, free meshing type was adopted to establish finite element model. The gyroscope rotor and gyro-chamber were connected by sixteen link8 elements, which represented the action of magnetic force. The element parameters of link8 were derived from the tested magnetic force value. The two sides of the gyro-chamber, which were connected with gimbal axes, were fixed in modal analysis, as shown in figure 2. Table 1 gives numerical results of modal analysis of different components.



(a) Gyro-chamber and gyroscope rotor (b) Frame
Fig.2. Finite element models of two components

Table 1. Natural frequencies of MS-CMG obtained from initial modal analysis scheme

Order	1	2	3	4	5	6	7	8	9
Frequency[Hz]	55.6	55.6	68.5	68.5	156	1137.9	1186.2	1250	1744
Structure Composing	Magnetic bearing					Gyro-chamber		Frame	

The initial five natural frequencies are corresponding to the modal shapes of magnetic bearing, which are very low compared with the work frequency 500Hz and could be ignored. The next two natural frequencies are referring to the gyro-chamber. After that, there are two natural frequencies associated with frame structure. These four frequencies are all very larger than the work frequency. Therefore, the MS-CMG structure fulfills the design requirements.

However, fixation boundary conditions in dynamic performance simulation of the gyro-chamber obviously increased the structural stiffness and inclined dangerous predictions were obtained. Thus more accurate simulation scheme is required for the structural design.

Modal analysis scheme of whole MS-CMG and results

A whole MS-CMG structural modal analysis scheme is presented aimed at the aforementioned problem. The whole finite element model of MS-CMG was established by connecting the gyroscope rotor, gyro-chamber, gimbal axes and frame. Several connection elements were adopted to simulate the assemble relationships in MS-CMG. The gyroscope rotor and gyro-chamber were connected by sixteen same link8 elements. In addition, a MPC technique was used to connect the gyro-chamber and the two gimbal axes. Furthermore, MPC184-Revolute element was selected to simulate the rotation relation between the gyro-chamber and frame structure. Two revolute pairs were defined between the mechanical bearings and the gimbal axes. MPC184-Revolute element has only one primary degree of freedom, i.e., the relative rotation about the revolute (or hinge) axis. Thus it could provide a relative rotation without axial displacement [6].

Fig.3 illustrates the finite element model of whole MS-CMG, which had 82,178 elements and 105,623 nodes. Zero displacement constraints were applied to nodes at which the frame was fixed on the satellite by four bolts. Block Lanczos method was used to perform modal analysis of MS-CMG, for its fast operation speed, high accuracy, and suitability for models including shell elements and solid elements. As a result, the initial ten modes were extracted for MS-CMG [7,8].

The initial ten natural frequencies are listed in table 2. The first natural frequency 0.001Hz, which is near zero, represents a rigid motion depicted by the relative rotation of gyro-chamber and frame structure. The next five natural frequencies are referring to the modal shapes of magnetic bearing. The seventh natural frequency is 755.6Hz. Its corresponding mode shape is the elastic vibration of the gyro chamber, which is shown in figure 4. The next two frequencies are also referring to gyro-chamber. The tenth frequency is with respect to the vibration of the frame. It can be seen that all the natural frequencies are far from the work frequency, thus no resonance occurs under work status of the MS-CMG.

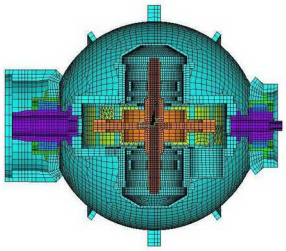


Fig.3. Finite element model of MS-CMG

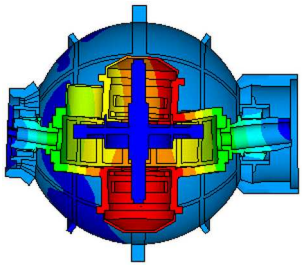


Fig.4. The seventh mode shape of MS-CMG

Table 2. Natural frequencies of CMG

Order	1	2	3	4	5	6	7	8	9	10
Frequency[Hz]	0.001	55.6	61	68.4	68.4	155	755.6	883	889	1156
Structure Composing	Gyro-chamber	Magnetic bearing					Gyro-chamber			Frame

Discussion and Conclusion

It is obvious that the natural frequencies corresponding to the magnetic bearing are relatively closed in table 1 and table 2. However, the natural frequencies associated with the gyro-chamber in table 2 are much less than that in table 1. The error arrives at 50%. Since a real rotation relation and displacement constraints are included in the new modal analysis scheme, the results in table 2 are more reliable. In addition, a small amount decreased for the natural frequency corresponding to the frame, which is due to the gyro-chamber and rotor components provide an additional mass in the modal analysis. Thus in the design of frame structure, frequency redundancy should be considered.

The FEM of the whole MS-CMG structure is established and a new modal analysis scheme is presented. Since a real work status of MS-CMG is described, the new scheme provides more reliable results. The predicted natural frequencies are far away from the work frequency. Thus the MS-CMG structure satisfies the design requirement. The modal analysis work in this paper provides understanding on the dynamic behavior of MS-CMG and further optimal design.

Reference

[1] B. Wie, K.W. Byun, V.W. Warren, D. Geller, D. Long and J. Sunkel: *Journal of Guidance, Control, and Dynamics* Vol. 12(5) (1989), p. 714

[2] Tang Liang and Chen Yiqing: *Acta Astronautica* Vol. 65 (2009), p. 1506

[3] Han Bangcheng, Hu Gang and Fang Jiancheng: *Optics and Precision Engineering* Vol. 14(4) (2006), p. 662

[4] Han Bangcheng, Zhang Shuna, Fang Jiancheng and Zhao Libin: *Journal of System Simulation* Vol. 20(3) (2008), p. 763

[5] Feng Hongwei and Fang Jiancheng: *Journal of Chinese Inertial Technology* Vol. 13(1) (2005), p. 72

[6] ANSYS Element Reference, Release 11.0, ANSYS Inc, 2007.

[7] Yang Guojun, Xu Yang, Shi Zhengang and Gu Huidong: *Nuclear Engineering and Design* Vol. 237(2007), p. 1363

[8] Y.A. Khulief and M.A. Mohiuddin: *Finite Elements in Analysis and Design* Vol. 26 (1997), p. 41

doi:10.4028/www.scientific.net/AMM.50-51.63

Set Pair Social Network Analysis Model

Hongmei Yang [1,a], Chunying Zhang [2,b], Ruitao Liang [2,c] and Fang Tian [3,d]

[1] College of Continue Education ,Hebei Polytechnic University,Tangshan Hebei 063009,China)

[2] College of Science,Hebei Polytechnic University,Tangshan Hebei 063009,China)

[3] College of architecture,hebei university of engineering.)

[a]yanghongmei@heut.edu.cn, [b]zchunying@heut.edu.cn,[c]lrtao335@163.com,[d]286909295@qq.com

Keywords: social networks; set pair analysis; set pair social networks;

Abstract: Through the study on social network information, this paper explore that there exists the certain and uncertain phenomena in the process of finding the relationship between individuals by using social networks, and the social networks are constantly changing. In light of there are some uncertainty and dynamic problems for the network, this paper put forward the set pair social network analysis model and set pair social network analysis model and its properties.

Introduction

Social networks, also known as social networks, are the relative steady relationship networks comprised by certain groups (people, businesses, organizations)[1]. As a new data mining technology, social network have applied in management, criminal organizations, the Internet and other areas. In the perspective of social networks[2-4], the references[2-4] found out the various social relationships existing as a single individual in the network, and the data mining for the social relationships. Zhou Weiguo[5] and other China scholars, find out the friendship between the network users through social network, and find out the shortest path between users through excavating the incidence relation between users at the exploited social network visual platform. Scholar Gao Peng[6] analyzed the timing sequence of chatting data to infer the relationship between users of social networking, and then judged and filtered the noise in the chatting data according to the reflected users' communicative relations in the social network. The references [5, 6] between individuals, exchange of information of some individuals to get the relationships between the individuals. And make use of social networking information excavate some useful information. However, in the process of assessment of the individual relations in the network through the social network, there still exist different and uncertain information elements. And the incidence relation between individuals cannot be reflected totally by using the certain and identical information assessment. The nodes in the real social networks and the connected properties usually changes over time[7]. The relation between network will change with the variations of factors, therefore, it's hard to dig the accurate information only in view of static state of the same elments between individuals at a time in network. Set pair analysis theory, a systematic analysis method to handle the uncertain problems, was put forward by Chinese scholar ZHAO Ke-qing in 1989. And then, SUN Jin-zhong built up a dynamic mould based on the Markow Chains set pair, and make huge progress in dealing with the dynamic uncertain problems of set pair. So , this article try to put the idea of the set pair analysis into the social network, and put forward the set pair social network analysis model and set pair social network dynamic changes, and related properties.

Basic concept

Social Network.

Social networks [8]can be represented as multiple points (social actors) and the set composed by lines between points (on behalf of the relationship between the actors), if order G as a social network, E the edge set, V point set, then the social network can be expressed as $G=(V,E)$. Social network

embodies a kind of structural relations, which can reflect the social relations between actors. Different types of relations forms different social networks, such as social communication network, social support network, discussion network, the right network.

Social networks can be divided into two categories: (1) the default social network (default social network): Social network information existing among the actors is not complete; actors and actors' relationsh value is default; (Figure 1a) (2) a complete social network (complete social network): actors and actors' relationship value is complete, that is, there exists relationship between any two actors. (Figure 1b).

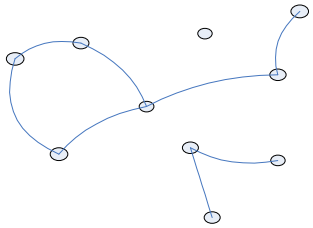


Figure 1a default social network

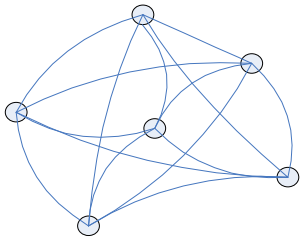


Figure 1b complete social network

Set pair analysis.

Definition 2.1 set pair [9] is linked by a certain set of two pairs, under the synergistic's uncertain systematic background, the two connected set A and B, their characteristics were analyzed, and N characteristics were obtained, which S owned by all set, and P characteristics were contrast (the opposite), the rest F characteristics relationship are uncertain, then the two sets of mathematical expression for the degree of connection u:

$$u = \frac{S}{N} + \frac{F}{N}i + \frac{P}{N}j \tag{1}$$

Type in: $\frac{S}{N}, \frac{F}{N}.\frac{P}{N}$, Denote them as a， b， c .The connection degree of the formula can be expressed as:

$$u = a + bi + cj \text{ and } a + b + c = 1$$

Where, i is the different markers in the interval $[-1,1]$ in which values depend on the circumstances, the different part tend to same in $[0,1]$, on the contrary in $[-1,0]$. j only acts as tag, set pair need to deal with the differences caused by uncertain between any two sets, and make use of relation to show the relavance degree.

Sets of network analysis model

In the social network relationships, the study of the properties of the two sets e_i and e_j, have N attributes (ie, the sum of the two objects' attributes), including S which are the same attribtes in two object properties set, and P are not the same property in two object properties, and F are the uncertain propertiesin two object properties. Then order:

$$\frac{S}{N} = a, \frac{F}{N} = b, \frac{P}{N} = c$$

As for the complex nodes in social networks, as well as the uncertainty between nodes, we give the following definition:

Definition 2.1 suppose all the study object in social network are domain set $U = \{e_1, e_2, \cdots e_n\}$, which e_k and e_s has the attribute set: $A(e_k) = \{x_{k_1}, x_{k_2}, \cdots, x_{k_n}\}$ and $A(e_s) = \{x_{s_1}, x_{s_2}, \cdots, x_{s_m}\}$, then the relation degree between e_k and e_s in the network at a time is:

$$\rho(e_k, e_s)(t) = a(t) + b(t)i + c(t)j \tag{2}$$

Where $e_k(t) \in U$ and $e_s(t) \in U$, a is the ration of the same properties and mutual properties in two objects; b is the ration of the uncertain properties and mutual properties in two objects; c is the ratio of the different properties and mutual properties in two objects; i is the different markers in the interval depending on the different circumstances on the values of $[-1,1]$, j only acts as tag and valued -1.

Definition2.1 We know the two study objects' relation degree is:

$$\rho(e_k, e_s)(t) = a(t) + b(t)i + c(t)j \tag{3}$$

If the given attribute certain weight $\omega_k (k = 1,2,\cdots,N, \sum_{k=1}^{N} \omega_k = 1)$, assuming the order of property by S, F, P and numbered them consecutively, then the connection degree is:

$$\rho(e_k, e_s)(t) = a(t) + b(t)i + c(t)j = \sum_{k=1}^{S} \omega_k + \sum_{K=S+1}^{S+F} \omega_k i + \sum_{k=S+F+1}^{N} \omega_k j \tag{4}$$

Where $i \in [-1,1]; j = -1; a + b + c = 1$

And in social networks, the network will change over time, the original property value changes, and some attribute value is invariant, and some turns into other properties. May wish to time in $t+1$, the original S_t properties were still in the same state S_{t1}, S_{t2} were transformed into an uncertain property, S_{t3} were converted to the opposite property. S_t in $[t, t+1]$ is the transfer vector (after normalization) as

$$\overline{a} = (M_{11}, M_{12}, M_{13}) = \left| \sum_{k=1}^{S_{t1}} \omega_k(t), \sum_{k=S_{t1}+1}^{S_{t1}+S_{t2}} \omega_k(t), \sum_{k=S_{t1}+S_{t2}+1}^{S_t} \omega_k(t) \right| \Big/ \alpha(t) \tag{5}$$

Type of $M_{11} + M_{12} + M_{13} = 1; \alpha(t) = \sum_{k=1}^{S_t} \omega_k(t)$

Similarly, the transfer vector $\overline{b}, \overline{c}$ can be obtained; all assumptions Ibid.

Therefore, at the time of $[t, t+1]$ the transfer matrix was M, at $t+1$ time, the degree of the two subjects was $\rho(e_k, e_s)(t+1)$:

$$\begin{aligned} \rho(e_k, e_s)(t+1) &= a(t+1) + b(t+1)i + c(t+1)j)^T \\ &= (a(t+1), b(t+1), c(t+1)) \cdot M \cdot (1, i, j)^T \end{aligned} \tag{6}$$

Definition 2.2: assuming matrices $R = (\rho(e_k, e_s))_{m \times n}$, R is set on the connection matrix, where $\rho(e_k, e_s)$ is the element on the connection matrix set, $R = (\rho(e_k, e_s))_{m \times n}$ matrix is the relational matrix which study the relation between object e_k and e_s, or the relation degree of every study object of a set of relational matrix confirmed in networks. This matrix is:

$$R = \begin{Bmatrix} \rho(e_1, e_1) & \rho(e_1, e_2) & \cdots & \rho(e_1, e_s) \\ \rho(e_2, e_1) & \rho(e_2, e_2) & \cdots & \rho(e_2, e_s \\ \vdots & \vdots & \vdots & \vdots \\ \rho(e_k, e_1) & \rho(e_k, e_2) & \cdots & \rho(e_k, e_s) \end{Bmatrix} \tag{9}$$

We can define the relation degree of set pair social network and network centrad as:

Definition 2.3 For a social network $G = (V, E)$, then the whole set of social networking on the link density is:

$$d = \frac{\sum_{k=1}^{n} \sum_{s=1}^{m} (\rho(e_k, e_1) + \rho(e_k, e_2) + \cdots + \rho(e_k, e_s))}{k \times s} \tag{7}$$

Definition 2.4: The relationship between social network study in any of the links with other subjects and take the maximum degree of any two subjects with the connection degree of difference from the previous maximum value is the central difference in the degree obtained:

$$C = \frac{\sum_{k=1}^{n}[c(\omega) - c(\omega_k)]}{\max \sum_{k=1}^{m}[c(\omega) - c(\omega_k)]} \quad k \neq s \tag{8}$$

Of which: $c(\omega) = \max(\sum_{s=1}^{m}\rho(e_1, e_s), \sum_{s=1}^{m}\rho(e_2, e_s), \cdots \sum_{s=1}^{m}\rho(e_k, e_s))$, $c(\omega_k) = \sum_{s=1}^{m}\rho(e_k, e_s)$.

3 Sets on the nature of social networks

Set on the Contact Matrix: In any social network links between the two study degree of $\rho(e_k, e_s)$, then the social network's relationship between objects forms a set of the contact matrix. table (5 * 5 networks)

Table 1 The Network Matrix of 5 * 5

	e_1	e_2	e_3	e_4	e_5
e_1	$\rho(e_1,e_1)$	$\rho(e_1,e_2)$	$\rho(e_1,e_3)$	$\rho(e_1,e_4)$	$\rho(e_1,e_5)$
e_2	$\rho(e_2,e_1)$	$\rho(e_2,e_2)$	$\rho(e_2,e_3)$	$\rho(e_2,e_4)$	$\rho(e_2,e_5)$
e_3	$\rho(e_3,e_1)$	$\rho(e_3,e_2)$	$\rho(e_3,e_3)$	$\rho(e_3,e_4)$	$\rho(e_3,e_5)$
e_4	$\rho(e_4,e_1)$	$\rho(e_4,e_2)$	$\rho(e_4,e_3)$	$\rho(e_4,e_4)$	$\rho(e_4,e_5)$
e_5	$\rho(e_5,e_1)$	$\rho(e_5,e_2)$	$\rho(e_5,e_3)$	$\rho(e_5,e_4)$	$\rho(e_5,e_5)$

From the table:

(1)set of the contact matrix $\rho(e_k, e_s)$ are arbitrary, $\rho(e_k, e_s) = \rho(e_s, e_k)$, the relation degree of the study objects e_s and e_k is the same as that of study objects e_k and e_s.

(2) the value of the diagonal constants, namely $\rho(e_k, e_s) = 1$, expressed the same degree of self-study 1.Diversity and opposition of 0.

(3) When $\rho(e_k, e_s) = -1$, then there is no connection between the two subjects.

In the social networks, the relation degree between any two study objects changes over time, the relationship between the two study may be strengthened, or the relationship between the two subjects may be reduced . Therefore, when the link between the two study objects become larger or smaller, then the relationship between the two study Strong Tendency or will appear weaker. Figure:

Figure 2 the Strongties and Weakties

Theorem 1: the link between the two study greater degree, then the relationship between the two objects will tend to strong ties.

Study the links between the two degrees are smaller, the relationship between the two objects will tend to weak ties.

Summary

This paper proposes related concepts of social networks and set pair analysis, and gives strong support to put forward the definition of set pair social network, meanwhile, it also proposes the relative properties according to the definition of set pair social network and dynamic change of network, and comes up with new method for the uncertainty problem in social network and the dynamic problem in network. Considering that the uncompleted set pair social network theory currently, there are a lot of work to do in this aspect.

Acknowledgment

The paper is financially supported by Natural Science Foundation of Hebei Province Department of Education(z2010142).

References

[1] Gou Hong-mei,Huang Bi-qing.A framework for virtual enterprise Operation management[J]. *Computers in Industry*,2003,50(3):333-352.

[2] Pan Hong, Zhai Dongsheng. Trust model based on social network research [J]. *Computer Engineering and Applications,* 2008,44 (12) ,215-220.

[3] Lu Zhiguo, Ma Guodong, Ren Shuhuai. Social network in the United States, the Library of the University [J]. *Library and Information Science*, 2009,155 (1), 69-72.

[4] Feng Qiu, Ye Xindong. Network Learning Community Social Network [J]. *Modern Education Technology,* 2010,20 (2), 37-41

[5] Zhou Wei Guo, Shi Rong Jian, Ye Ting. Web2.0 Web-based information mining in social networks [J]. *Micro Computer,* 2009,25 (2-3) :146-147.

[6] Gao Lin,Liu Jiming,Zhang Shiwu,et a1.Discovering the Dynamics in a Social Memory Network[C]//Proc.of 2008 International Conference on Web Intelligence and Intelligent Agent Technology.Sydney,Australia：[s,n],2008.

[7] Gao Peng, Xian-Bin Cao. Chat social network data based on the noise filter [J]. *Computer Engineering,* 2008,34 (5) :166-168.

[8] John Scott , Liu Jun translation, and so on. Social network analysis [M]. Chongqing: *Chongqing University Press*, 2007.

[9] Indefinite. Set Pair Analysis and its Preliminary Application [M]. Hangzhou: *Zhejiang Science and Technology* Press, 2000:8 - 90.

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.68

Optimal Design of the Vertical Machining Center Based on Stiffness Analysis

Yanmin Zhao, Zhijun Wu, Jianfu Zhang, Pingfa Feng and Dingwen Yu

Departent of Precision Instruments and Mechanology, Tsinghua University, Beijing, China

plont.zhao@163.com

Keywords: optimal design; vertical machining center; static stiffness; natural frequency; dynamic stiffness

Abstract. Stiffness analysis plays an important role in the optimization of the machine tool. By analyzing the static stiffness, dynamic stiffness and natural frequency of the key components, we discussed how to improve static and dynamic performance of the whole machine. Through the finite element analysis method, weak parts of the vertical machining center are firstly identified based on static stiffness analysis. Due to the purposes of increasing the stroke in Y-direction, improving the whole machine stiffness and without increasing the whole machine weight, optimal designs are carried out mainly on the structures of the spindle box and the column. And the performance of the parts and the whole machine before and after optimization is compared by testing whether the stiffness is improved or not. We also conducted experiments, and the results are consistent with the results of finite element analysis.

Introduction

Stiffness is one of the most important performance index for metal cutting machine [1]. Raising stiffness is beneficial to machine tools' efficiency, machining accuracy and suface processing quality [2]. Stiffness consists of static stiffness and dynamic stiffness, both of which need to be analyzed and optimized. Generally speaking, the larger static stiffness is, the harder one object will be deformed. And dynamic stiffness expresses the dynamic response when an object is under alternating load.

Stiffness could affect the machining accuracy significantly. So stiffness analysis has been widely used in design and optimization of machine tools, not only the traditional machine tools, but also machining center, parallel machines and hybrid machines.

In this paper, the research object is a vertical machining center (VMC) manufactured by TONTEC. The basic principles of stiffness analysis are based on finite element analysis theory. We used finite element analysis method to do static analysis, modal analysis and harmonic analysis separately. Static analysis is used to solve the static stiffness of the key parts, and harmonic analysis is used to solve the dynamic stiffness of the key parts. Modal analysis is used to solve the natural frequency of the key parts. To avoid some unnecessary computation, some models' structures, such as chamfering and small bosses, have been simplified.

Stiffness Analysis

Stiffness is the ability to resist deformation for mechanical parts and components. In the elastic range, stiffness is a coefficient which is proportional to the load applied to the parts, and the load is also proportional to the components' displacement [3]. In other words, stiffness is the force that could cause unit displacement. Stiffness consists of static stiffness and dynamic stiffness [4]. Static stiffness can be defined as the following form:

$$k_s = \frac{F}{d} \tag{1}$$

Where k_s is static stiffness, and F is the external force applied to the object, and d is the maximum deformation of the object in the external force's direction.

Eq.1 defines the calculation method of static stiffness for an individual component. Symbols in the equation are slightly different for the whole machine. Since we are most concerned about the deformation of spindle where the cutting tool is installed, d is the maximum deformation of the spindle in the external force's direction when calculating the static stiffness of the whole machine.

Dynamic stiffness is the capacity to resist deformation when dynamic load is applied to an object [5]. If the frequency of the interference force is close to the natural frequency of the object, resonance will appear [6]. The dynamic stiffness is the minimum at this moment, and the object is the most possible to be deformed. What's more, the deformation will be several times more than the static deformation. The dynamic stiffness is defined as the following form:

$$\overrightarrow{k_d} = \frac{\overrightarrow{F}}{\overrightarrow{d}} \tag{2}$$

Where $\overrightarrow{k_d}$ is dynamic stiffness, and $\overrightarrow{F}$ is the dynamic force applied to the object, which is usually a harmonic force, and $\overrightarrow{d}$ is the maximum deformation of the object in the dynamic force's direction. All of them are vectors, and dynamic stiffness is in complex form. Dynamic stiffness can be derived in the following expression:

$$\overrightarrow{k_d} = k - m\omega^2 + D(1-\lambda)\omega \tag{3}$$

The dynamic stiffness parameters include the rotor mass m, radial direct stiffness k, rotative speed ω, damping D, and fluid circumferential average velocity ratio λ. λ is the ratio of the average fluid angular velocity divided by the rotor angular velocity. We usually ignore the damping D, so the dynamic stiffness $\overrightarrow{k_d}$ only contains the real component. And $\overrightarrow{F}$ in Eq.2 is the peak of the harmonic force.

Optimization of the VMC

Optimization Goals

The vertical machining center is made up of a large number of components. In order to optimize the whole machine, the optimization goals include the following aspects:

- To improve the stiffness of the key components, aiming to achieve the goal of increasing the whole machine stiffness.
- To achieve the above goal without increasing the whole machine weight.
- To increase the stroke in Y-direction, to increase the machining range of the vertical machining center, while not reducing the whole machine stiffness in Y-direction.
- To improve the whole machine stiffness in X-direction, while not reducing the stiffness in the other two directions.

Static Stiffness Analysis of the Whole Machine

The relative positions between components will affect the static stiffness of the whole machine. Therefore the components are assembled in accordance with their relative positions under cutting conditions. The vertical machining center has three axes. So we need to do static stiffness analysis in three directions respectively. X-axis is the direction of the worktable's movement, Y-axis is the direction of the saddle's movement, and Z-axis is the direction of the spindle box's movement.

Fig.1 shows the machine model after being imposed of restrictions and loads. The bolt holes at bed's bottom are set to be fixed. The external force is applied to the spindle, and its reaction force is applied to the worktable. Loads are applied to the model in three directions respectively to solve the static stiffness of the whole machine. Then we can obtain the deformation of the whole machine in three directions. Fig.2 shows the deformation in Z-direction. According to Eq.1, we can calculate the static stiffness (N/μm) of the whole machine in X, Y and Z directions are 65.7, 186.9 and 97.5 respectively. It's easy to see that the static stiffness in X-direction is relatively poor, and Y-direction relatively good. We could reduce the static stiffness in Y-direction appropriately, so that the static stiffness in X-direction would be improved. The spindle box and column render the maximum

deformation, which can be seen from the deformation pattern displayed in the software. Therefore, we need to optimize the spindle box and the column.

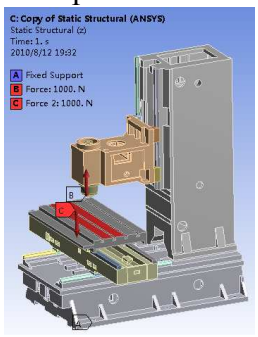


Figure 1. Restrictions and loads for simulation

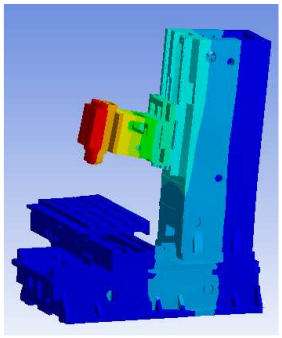


Figure 2. The deformation in Z-direction

Optimization of the Spindle Box

Fig.3 shows the structure of the spindle box. We will do static stiffness analysis with it at first. Its surfaces that contact with the sliding bottom on the sliding rails are set to be fixed. A remote force is applied to its torus. Thus its static stiffness can be calculated according to Eq.1.

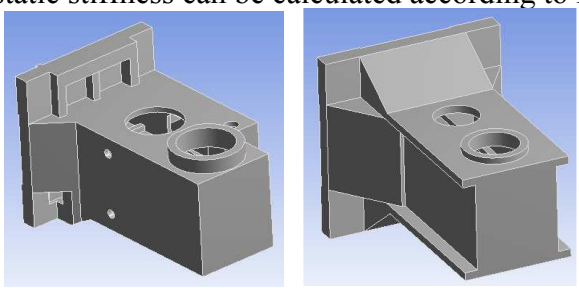


Figure 3. The spindle box model before and after optimization

We can see that the structure of the spindle box is asymmetric from Fig.3. In order to balance the stiffness in all directions, we changed its structure into symmetric, as shown in Fig.3. We could optimize the spindle box further, for example, to increase the size in X-direction, reduce the size in Y-direction, or increase the size of the wall thickness, etc. The results obtained are shown in Table II.

The meanings of the codes in Table 1 are as following.

- 1 represents the mass and static stiffness data of the spindle box before optimization.
- 2 represents the data after adding stiffeners.
- 3 represents the data after reducing the size in Y-direction.
- 4 represents the data after increasing the size in X-direction.
- 5 represents the data after increasing the size of the wall thickness.
- 6 represents the data after reducing the size in Y-direction while increasing the size in X-direction.
- 7 represents the data after reducing the size in Y-direction while increasing the size in X-direction and the size of the wall thickness.

Table 1. Static stiffness of the spindle box

Code	Mass (kg)	Static Stiffness (N/μm)		
		X-direction	Y-direction	Z-direction
1	280.3	134.9	717.2	325.3
2	319.7	148.1	755.1	386.4
3	269.8	153.6	755.3	386.3
4	290.3	179.3	743.7	341.9
5	309.9	167.4	877.1	409.5
6	281.6	201.1	775.3	400.6
7	312.8	261.9	1005.4	515.5

As can be seen from the data in Table 1, adding stiffeners has little effect on increasing stiffness, but it will increase the weight. In contrast, increasing the size of the wall thickness will increase the static stiffness, but it will also increase the weight. Based on comparison of the data in Table 2, we choose Option 6. Then, we did dynamic stiffness analysis and modal analysis with the spindle box before and after optimization. The results are listed in Table 2.

Table 2. Dynamic stiffness and natural frequency of the spindle box

	Natural Frequency(Hz)	Dynamic Stiffness (N/μm)		
		X-direction	Y-direction	Z-direction
Before Optimization	703.8	9.96	12.11	2.39
After Optimization	751.9	87.68	20.86	4.19

As can be seen from Table 2, the dynamic stiffness has been enhanced, and the natural frequency has been improved. The natural frequency in Table 2 is the first natural frequency, and the higher natural frequency has also been improved.

Optimization of the Column

Optimizing the column is similar to the steps of the optimization of the spindle box. During the optimization of the column, we increased its size in X-direction, reducing the size in Y-direction and Z-direction, aiming to improve the stiffness in X-direction. Fig.4 shows the structure of the column before and after optimization.

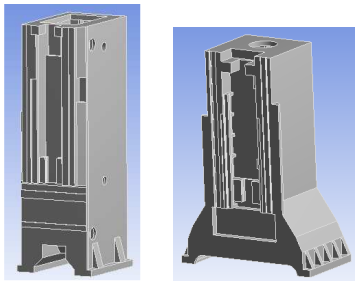


Figure 4. The column model before and after optimization

The surfaces of the column that contact with the bed are set to be fixed, and a remote force is applied to the surfaces that contact with the sliding rails. Then its static stiffness could be calculated according to Eq.1. The results of static analysis and modal analysis are shown in Table 3.

Table 3. Static stiffness and natural frequency of the column

	Natural Frequency(Hz)	Static Stiffness (N/μm)		
		X-direction	Y-direction	Z-direction
Before Optimization	257	380.1	549.4	748.4
After Optimization	376.5	724.1	632.9	724.3

As can be seen from Table 3, the static stiffness has been enhanced. The natural frequency in Table 4, which is the first natural frequency, has been improved, and the higher natural frequencies have also been improved. When the masses of two different objects are similar, the stiffness of the one with higher natural frequency should be better. So in this case, we can conclude that the column stiffness after optimization is better than its stiffness before optimization.

Performance Comparison Before and After Optimization

Now we will do stiffness analysis with the whole machine model after optimization. Only a harmonic force is imposed on the spindle in dynamic stiffness analysis. And the other settings are the same with static stiffness analysis. Based on the displacement obtained, we can calculate the dynamic stiffness of the whole machine according to Eq.2. And the results are listed in Table 4.

Table 4. Performance of the whole machine before and after optimization

	Static Stiffness (N/μm)			Dynamic Stiffness (N/μm)		
	X	Y	Z	X	Y	Z
Before Optimization	65.7	186.9	97.5	19.6	23.9	8.66
After Optimization	94.5	225.3	106.1	88.3	23.2	20.9

From the data in Table 4, we can see that the static stiffness of the whole machine after optimization has been all improved in three directions. Although the dynamic stiffness in Y-direction decreased slightly, but the dynamic stiffness in the other two directions has been improved.

Conclusions

We used finite element analysis method to analyze the static stiffness of the whole machine model, and we found the weak parts on static stiffness were the spindle box and the column. To improve the static stiffness of these parts, we carried out improvements with these key parts. Then we did static analysis, modal analysis and harmonic analysis with them separately to compare their performance, and found that their stiffness had been improved. Finally, we performed finite element analysis to calculate the static stiffness and dynamic stiffness of the optimized machine, as well as experiments to verify the results of the simulation.

We can see that all the optimization goals mentioned in Part III have basically been achieved. From the optimization results above, we can see that the effects of optimization based on stiffness analysis are quite obvious. And the analytical and the experimental results showed good consistency, which means that the optimal method is practical.

Furthermore we noticed that the errors between experimental results and simulation results are relatively large, although their trend is consistent. This is because the simulation we did does not take it into account that the impact of joint surfaces.

Acknowledgment

This project is supported by the Key National Science and Technology Projects of PR China (No. 2009ZX04014-035).

References

[1] Yue Liu and Jinsong Wang, “Static stiffness research and optimization on a hybrid machine tool considering the stiffness of bearings and guide ways,” *Chinese Journal of Mechanical Engineering,* Vol.43, No.9, Sep. 2007

[2] Yue Liu and Jinsong Wang, “Static stiffness research and optimization on a hybrid machine tool considering the stiffnesses of bearings and guideways,” *Chinses Journal of Mechanical Engineering,* Vol.43, No.9, Sep. 2007

[3] Yue Liu, Jinsong Wang and Liping Wang, “Stiffness optimization of a heavy hybrid machine tool named XNZH2430,” *Journal of Tsinghua University (Science and Technology)*, Vol.46, No.8, 2006

[4] “Understanding and using dynamic stiffness,” *A Tutorial, information on http://www.ge-energy.com/prod_serv/products/oc/en/orbit/downloads/2q00tutorial.pdf*

[5] Bin Zhu and A. Y. T. Leung, “Dynamic stiffness for thin-walled structures by power series,” *Journal of Zhejiang University Science A*, Vol.7(8):1351-1357, 2006

[6] G. V. Narayanan, “On direct computation of beam dynamic stiffness coefficients using MSC.Nastran,” *information on http://www.mscsoftware.com/support/library/conf/auto00/p01000.pdf*

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.73

Novel design for a biomimetic water-jetting propulsion vehicle actuated by SMA wires

Yangwei Wang, Zhenlong Wang, Jian Li and Fei Gao

State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin 150001, Heilongjiang, China

wywkly@126.com, wangzl@hit.edu.cn, lijian499@163.com

Keywords: Biomimetic; SMA wire; water-jetting propulsion; robot cuttlefish

Abstract. An experimental biomimetic water-jetting propulsion vehicle for robot cuttlefish is presented. Firstly the cuttlefish swimming mode and mantle muscle structure is analyzed as the theory basis. Secondly the vehicle is designed and manufactured. The vehicle is actuated by SMA wires. Through rational structural design, it can achieve flexible water-jetting propulsion like cuttlefish mantle. Thirdly the geometrical model of the biomimetic mantle cross section is built. At last, the performance of the propulsion machine is tested by experiment. The experimental results are mainly about the response curves of the mantle outside diameter, the strain of SMA wire and the swimming displacement with time. The maximum contractive distance of the biomimetic mantle along the vertical direction is 6 mm, the maximum strain of SMA wire is -1.8 %. The maximum swimming speed is 45 mm/s.

Introduction

Ocean resources are becoming more and more important as the earth resources are exhausted gradually in the 21st century. Unmanned underwater vehicles (UUVs) are the most effective devices to explore, to exploit and to protect the ocean resources. But the UUVs propelled by screws suffer from low propulsive efficiency, bad maneuverability, high noise, and aquatic life damage by their blades, etc. The locomotion of aquatic animals is a result of many millions of years of evolutionary optimization for rapid, efficient and adjustable movement. With the development of the computer technology, bionics, mechatronics and materials science, many kinds of biomimetic underwater robots imitating aquatic animal have been developed, for instance, robot fishes [1-3], the robot lobster [4], the undulatory tadpole robot [5], the jelly fish robot [6], the robot squid [7], etc.

Fish are the most common aquatic animals which propelled themselves by their undulatory/oscillatory fins of body. It is well known that fishes swim with high speed, high efficiency, high maneuverability and low noise. So many kinds of autonomic underwater vehicles tend to imitate fishes' locomotion modes to achieve high performance recently. Another class of aquatic animals swims using a kind of water-jetting propulsion mode. Thrust is generated by directly jetting water from the body cavity which is powered by contraction of muscles. Comparing with the fish swimming modes, water-jetting propulsion can achieve higher speed, for example escape-jet locomotion of squids. The jetting propulsion can also achieve the precise positioning during slow swimming with high efficiency for example pelagic locomotion of jellyfishes. Consequently, the underwater robot using biomimetic water-jetting propulsion has a very well research prospect. Researchers have developed a variety of water-jetting propulsion vehicles already. Mohseni et al from university of Colorado developed a vortex ring thruster (VRT) [8]. Thomas et al from California Institute of Technology developed a synthetic jet propulsion system used in a small scale autonomous underwater vehicle with low speed and high maneuverability [9].

In the present work, the mechanism of water-jetting propulsion locomotion and the structure in the mantle of the cuttlefish are firstly analyzed. Then a biomimetic water-jetting propulsion vehicle that imitates cuttlefish is developed. The vehicle is actuated by the Shape Memory Alloy (SMA) wires. At last the propulsion performance of the biomimetic water-jetting propulsion vehicle is experimented

and tested. The experimental results show that the vehicle has a simple structure. It also can propel itself by jetting water obviously following flexible contraction of the man-made mantle.

Mechanism of water-jetting propulsion mode of cuttlefish

Cuttlefish, which are soft-bodied marine cephalopods, propelled themselves by using a complex of jetting propulsion for high speed swimming and fin undulatory propulsion for low speed swimming and hovering. Figure 1 shows the jetting process of cuttlefish. The jetting propulsion process includes both the refill and jet periods. During the refill period, mantle becomes inflated firstly and consequentially water rushes into the mantle cavity by passing the sides of the head and the collar. In this period, the funnel orifice keep closed. Then the mantle contracts, so the water pressure in the mantle cavity increases and the collar is forced against the mantle wall. The route used for sucking water into the mantle cavity is then closed down, which causes the collar act as a one-way valve. The ventral wall of the funnel presses against the mantle as well and the funnel stays open during mantle contraction. Water, therefore, is forced through the funnel and out the funnel orifice. The thrust provided by the water leaving the funnel based on the principle of momentum conservation can be directed by muscularly directing the funnel. After the jet period, the cuttlefish refills its mantle cavity with water to prepare for the next jet period again. When the jet process begins, the jet orifice expands to its maximum diameter rapidly. During the remainder of the contraction cycle, cuttlefish actively reduces funnel diameter to maximize power output from a given volume change [10]. During the refill periods, the fins of cuttlefish will be undulatory with high frequency to reduce the loss of velocity. During the jet periods, the fins will be wrapped around the outside mantle to maintain an excellent streamlined body and reduce drag.

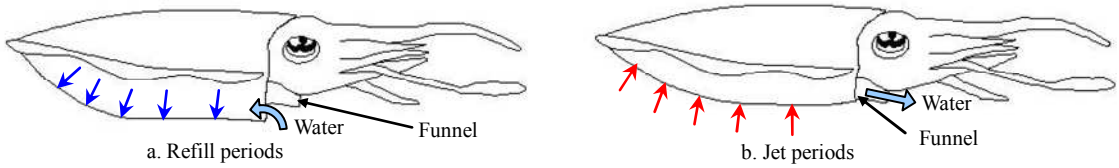


Fig.1 Scheme of jetting process for cuttlefish

Muscle structure of cuttlefish's mantle

Cuttlefish's body is not supported by common hardened skeletal elements, but by the muscular hydrostatic skeletal. Muscular hydrostatic skeletal differ from common musculoskeletal systems because muscles alternately stiffen and support a structure, or produce movement. The musculature therefore acts both as the effector of movement and as support for movement. The lofty swimming ability of cuttlefish benefits from its muscular hydrostatic skeletal system.

The mantle muscle structure of cuttlefish and squid is same. The mantle of loliginid squids includes two predominant muscle orientations: circumferential muscle fibers (CM) that constitute the bulk of the mantle wall and radial muscle fibers (RM) that extend from the inner to the outer surface of the mantle wall as partitions between the bundles of circular muscle fibers [11].

The design of biomimetic water-jetting propulsion vehicle

Actuator selection. Actuator selection is very important in designing the biomimetic water-jetting propulsion vehicle. In order to simulate muscle contraction well, the drive materials need to provide sufficient output force, power and deformation. Presently, widely used artificial muscle materials include SMA, ion conducting polymer metal composite and piezoelectric. SMA actuators have these advantages: small volume, simple structure, flexible movement, easy to control, less effected by the surrounding environment (except temperature). Ti-Ni based SMA is widely used because of its high resistance rate, long fatigue life, large recovery stress, large recovery strain and high energy density.

SMA wire is similar to the circular muscle fiber of cuttlefish. So it is easy to lay and embed into the structure as a suitable biomimetic water-jetting propulsion actuator.

Design guidelines. When cuttlefish swims in low speed，the action frequency of the mantle is also low. In this swimming strategy, the mantle contraction is active by circular muscle fiber contraction. But the mantel expansion is passive because it just relies on the energy stored in the mantle. The radial muscle does not contract at all. The present biomimetic water-jetting propulsion vehicle just simulates cuttlefish with slow frequency and low speed. The requirements of the water-jetting propulsion vehicle are given as follow:

(1) The cross section of the biomimetic mantle is semicircular. It can achieve uniformity contraction and release actions like cuttlefish's mantle. The contraction action is driven by SMA wires contraction. The refilling action relies on the elastic energy stored within the mantle.

(2) It is difficult to make the biomimetic mantle be a pure flexible hydrostatic skeletal structure due to existing technology, so in our present research a kind of quasi-flexible structure was used.

(3) As the thickness of the biomimetic mantle is limited, it is difficult to simulate the radial muscle and its contraction. Only circular muscle contraction and its movement are simulated.

Structure design. The biomimetic water-jetting propulsion vehicle has similar shape with real cuttlefish, consisting of the biomimetic mantle, shell, switch membrane and nozzle. Figure 2 shows the biomimetic water-jetting propulsion vehicle. The cross-section of the biomimetic mantle is semicircular. The rearward and the circle side of the biomimetic mantle are closed while the foreside is open. The refill and jet can be achieved by diametrical change of the biomimetic mantle. The shell is basis of the vehicle. The nozzle and switch membrane are installed at the foreside of the shell. The switch membrane is also a flexible structure which is open when refilling water and closed when jetting water.

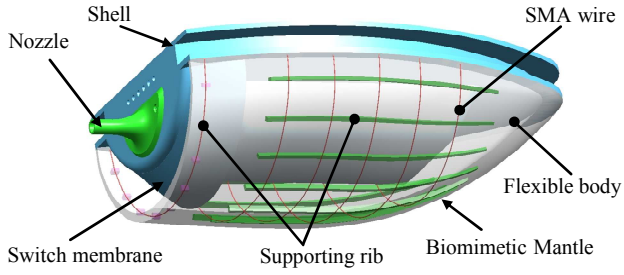


Fig.2 The structure of the vehicle

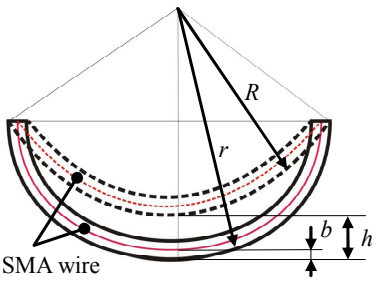


Fig.3 A cross section of biomimetic mantle

The biomimetic mantle consists of supporting ribs, SMA wires and flexible body. The flexible body is base of the biomimetic mantle, made of silicone gel. The silicone gel has well flexibility, resilience, stability, tear resistance. Its density is close to water. Supporting ribs are evenly distributed and embedded along the circumferential direction in the flexible body, playing a supportive role. Every two adjacent ribs are connected by a SMA wire. When the SMA wires contract, the distance between every two adjacent ribs will be reduced. Then the entire biomimetic mantle contracted. Figure 3 shows the contraction process of a cross section of biomimetic mantle.

The elastic energy resilience mechanism is employed in the mantle movement process. When the SMA wires are electrical heated, they contract because of their inner phase transformation. The forces generated by SMA wires in vertical direction of the ribs reduce the distance between adjacent ribs. This will induce an increase in the thickness of the biomimetic mantle wall. Because the biomimetic mantle wall is essentially constant in volume, it must result in a decrease in the biomimetic mantle diameter while elastic energy is stored in the flexible body. When the electrical heating ends, the SMA wires begin to transform into martensite phase. Then the energy releases from the flexible body and drives the SMA wires to return its initial length. The flexible body also returns to its initial thickness. This thinning results in an increase in the diameter of the biomimetic mantle and expansion of the mantle cavity. The process goes round and round. The pulse water-jetting propulsion is achieved by cycle of heating and cooling of the SMA wires.

Several hypotheses are made to describe the kinematics of the biomimetic mantle by a mathematic model: the longitudinal length of the biomimetic mantle always keeps at a constant; the thickness of the biomimetic mantle at every time t keeps uniform in the deformation process; the inside and outside walls of the biomimetic mantle are circular surfaces in the deformation process; the deformation of supporting rib and its impact on silicone gel deformation is ignored because the supporting rib is narrow; The contraction of SMA wires keep uniformity in the action process.

So the relationship of the SMA wires strain and the outside diameter of the biomimetic mantle can be expressed

$$\begin{cases} R(t) = \dfrac{2r^2 + h^2(t) - 2rh(t)}{2r - 2h(t)} \\ \varepsilon(t) = \dfrac{\pi(r-b) - 2\arcsin(\dfrac{r}{R(t)}) \cdot [R(t) - b]}{L} \end{cases}$$

where R is the outside diameter of the biomimetic mantle after the contraction of SMA wires, h is contractive distance of the outside diameter along the vertical direction, r is the is the outside diameter of the biomimetic mantle before the contraction of SMA wires, t is time, b is the distance from the installation location of the SMA wires to the outside wall of the biomimetic mantle, L is the length of a single SMA wire, ε is the strain of the SMA wires.

Experiments and discussion

In order to test the performance of the water-jetting propulsion vehicle, initial swimming experiment was done. The biomimetic mantle is 230 mm in length, 160 mm in width, 55 mm in height, 8 mm in thickness. 9 supporting ribs and 6 SMA wires are embedded in the biomimetic mantle. The diameter of the SMA wires is 0.2 mm. The nozzle is 30mm in length, 6 mm in inside diameter of the jetting orifice. The vehicle weight is 728 g and filled water weight is 700 g. Figure 4 shows the swimming state of the vehicle in water.

In the experiment DC power supply is used to drive SMA wires. The electric voltage is 20 V, 24 V and 28 V. The heating time is 1.5 s, and cooling time is 2 s. Digital video recordings of the motion of the biomimetic vehicle were used to determine very accurate jetting kinematics and biomimetic mantle deformations. Then the action images are processed through the video processing software. Figure 5 shows a test result of swimming along a straight line. The contractive distance of the outside diameter along the vertical direction and strain of SMA wires are shown in Figure 6.

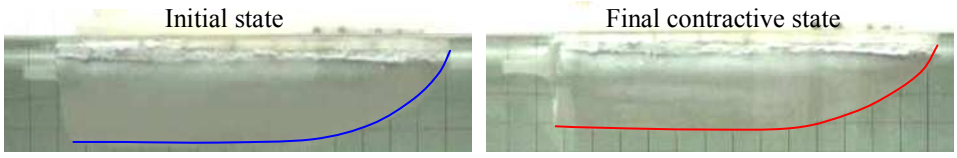


Fig.3 The biomimetic water-jetting propulsion vehicle

The experimental results show that: The radial contraction of the biomimetic water-jetting vehicle is uniform distribution, imitating flexible contraction of real cuttlefish mantle well. The maximum swimming speed is 45 mm/s. The maximum contractive distance of the biomimetic mantle along the vertical direction is 6 mm, the maximum strain of SMA wire is -1.8 %. With the increase of voltage, the maximum swimming speed, the contractive distance of the mantle and the strain of SMA wire also increase. The displacement curve in figure 5 shows: in 0 ~ 1.6 s interval, the swimming speed increase fast; after 1.6 s, the speed reduces gradually.

The contractive distance curve of the mantle and the strain curve of SMA wire in figure 6 show: in 0 ~ 1.6 s interval, the SMA wires contract driven by electric current heating, then the mantle contract

and achieve maximum strains; after 1.6 s, the heating of the SMA wire is over, the biomimetic mantle expand back to its initial state by elastic energy stored in the mantle.

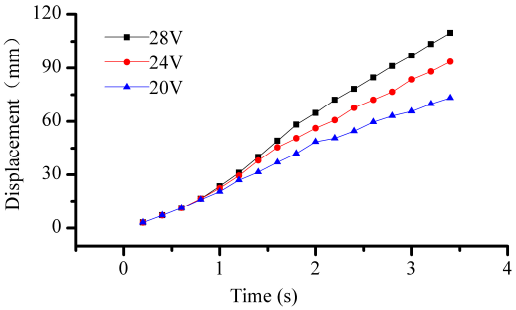


Fig.4 The curves of displacement

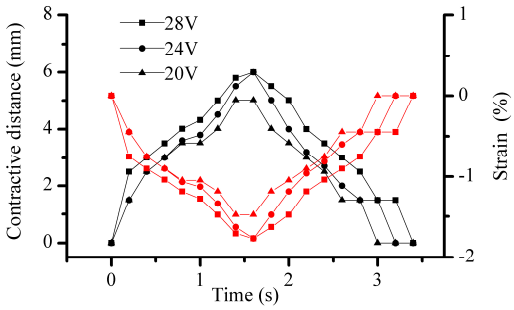


Fig.5 The curves of contractive distance and strain

Conclusions

The biomimetic water-jetting propulsion machine is developed by based on the analysis of cuttlefish's swimming mode and mantle muscle structure. The machine is actuated by SMA wires, through the rational structural design achieved a flexible water-jetting propulsion that simulated jetting action of cuttlefish mantle. The machine has a simple structure, and without any transmission. Experimental results show that the propulsion machine has a flexible action is good, noise-free swimming, etc., for the development of robot cuttlefish laid the foundation. With the gradual improvement of the design and perfect the biomimetic water-jetting propulsion machine for underwater robot development will provide a new and efficient underwater propulsion machine.

Acknowledgment

This work is supported by the National Natural Science Foundation of China under grant No. 50775049 and by Self-Planned Task (NO. SKLRS200805C) of State Key Laboratory of Robotics and System (HIT).

References

[1] A. H. Techet, F. S. Hover, M. S. Triantafyllou: Flow, *Turbulence and Combustion*, Vol. 71 (2003), p. 105~118

[2] J. D. Liu, I. Dukes, H. S. Hu: *International Conference on Intelligent Robots and Systems* (Edmonton, Canada, August 2-6, 2005)

[3] K. H. Low, A. Willy: *Journal of Vibration and Control,* Vol. 12(2006), p. 1337~1359

[4] K. K. Safak, G. G. Adams: *Robotics and Autonomous System*, Vol. 41(2002), p. 225~243

[5] Kim, D. H. Kim, J. Jung, et al.: *Smart Materials and Structures*, Vol. 14(2005), p. 1579~1585

[6] S. Guo, L. Shi, X. Ye, et al.: *IEEE International Conference on Mechatronics and Automation* (Harbin, China, August 5-8, 2007)

[7] Information on http://www.robotster.org

[8] M. Krieg, K. Mohseni*: Journal of Oceanic Engineering,* Vol. 33(2008), p. 123~132

[9] A. P. Thomas, M. Milano, M. G. G'sell, et al.: *IEEE International Conference on Robotics and Automation* (Barcelona, Spain, April 18-22, 2005)

[10] I. K. Bartol, M. R. Patterson, R. Mann: *Experimental Biology,* Vol. 204(2001), p. 3655~3682

[11] W. M. Kier, J. T. Thompson: Proceedings of the 1st International Symposium: *Coleoid Cephalopods through Time* (Berlin, Germany, 2003)

doi:10.4028/www.scientific.net/AMM.50-51.78

A Fuzzy Concept Similarity Measure Based on Lattice Structures

Lidong Wang[1,a], Dianxuan Gong[2,b] and Xin Wang[1,c]

[1]Department of mathematics, Dalian Maritime University, Dalian 116026, P. R. China

[2]College of Sciences, Hebei Polytechnic University, Tangshan 063009, P. R. China

[a] wld1979@yahoo.com.cn, [b] dxgong@heut.edu.cn, [c] xenawang7811@hotmail.com

Keywords: Concept similarity, fuzzy concept lattice, irreducible element, ontology

Abstract. With the rapid development of the semantic web, determining the degree of similarity between concepts from same or different ontologies plays an increasing crucial role. In this paper, a new similarity model based on lattice structural information is proposed to evaluate the similarity degree between fuzzy concepts in the framework of fuzzy formal concept analysis. The proposed method preserves more structural information, which can be viewed as another extension and development of de Souza and Davis's model in fuzzy context.

Introduction

Evaluating the similarity between two things is one of the most popular ways for people to compare the two things and acquire knowledge. A variety of similarity measures have been proposed for evaluating the similarity between RNA structures, fuzzy rule bases, domain concepts, documents, ontology mapping etc. [1]. In semantic web, measuring concept similarity is crucial for ontology mapping, ontology alignment. Ontology is a conceptualization of a domain into a human understandable, machine-readable format consisting of entities, attributes, relationships, and axioms. It is used as a standard knowledge representation for the semantic web. With the rapid development of the semantic web, it is likely that the number of ontologies will greatly increase during the next few years, which leads to the arising demand for rapid and accurate assessing concept similarity [2]. So, assessing the similarity between concepts is an important and interesting activity in ontology mapping and alignment

Many similarity measure approaches to deal with ontology mapping and alignment have been proposed. The main approaches are based on distance within an ontological structure, concept information content, concept feature matching and a combination of these approaches. Currently, FCA (formal concept analysis) techniques have been used to deal with evaluating concept similarity in ontology construction and merging [2-5]. Concept lattice-based similarity models can preserve more featural and structural information and have a higher correlation with human intuition. de Souza and Davis have introduced a new similarity model based on concept lattice, in which meet-irreducible elements (structural elements) are then used as features in similarity model proposed by Rodriguez and Egenhofer [6].

These measures can only deal with typical ontology merging and ontology mapping. However, the conceptual formalism supported by typical ontology may not be sufficient to represent uncertainty information commonly found in many application domains due to the lack of clear-cut boundaries between concepts of the domains. To tackle this type of problems, one possible solution is to incorporate fuzzy logic into ontology to handle uncertainty data [7]. Fuzzy ontology is generated and used in text retrieval and search engines, in which membership values are used to evaluate the similarities between the concepts in a concept hierarchy [8]. In particular, fuzzy ontology based on FFCA (fuzzy formal concept analysis) is a currently interesting topic [8-10]. Thus, it is necessary and interesting to assessing the similarity between concepts under fuzzy context.

Inspired by above works, a new similarity model based on lattice structural information is proposed, which preserves more structural information of fuzzy concept lattice. The paper is organized as

follows. In the next section, the notion of a fuzzy concept lattice is recalled. In Section 3, the related similarity models are briefly summarized. Successively, in Section 4, a new method for assessing fuzzy concepts similarity is proposed. A detailed analysis for the model is discussed. Finally, conclusions are drawn in Section 5.

Fundamentals of fuzzy formal concept analysis

The fuzzy formal concept analysis (FFCA) incorporates fuzzy logic into formal concept analysis to represent vague information. There exist few different definitions about FFCA, we adopt the definition introduced by [8]. In FFCA, a fuzzy concept is defined within the fuzzy formal context.

Definition 2.1. ([8, 11]) A fuzzy formal context is a triple $K = (X, M, I = \phi(X, M))$, where X is a set of objects, M is a set of attributes, and I is a fuzzy set on domain $X \times M$. Each relation $(x, m) \in I$ has a membership value $\mu(x, m)$ in [0,1].

Given a fuzzy formal context $K = (X, M, I)$ and a confidence threshold T, denote

$A^* = \{m \in M \mid \forall x \in A, \mu(x, m) \geq T\}$ for $A \subseteq X$ and $B^* = \{x \in X \mid \forall m \in B, \mu(x, m) \geq T\}$ for $B \subseteq M$.

Definition 2.2. ([8]) A fuzzy formal concept (or fuzzy concept) of a fuzzy formal context (X, M, I) with a confidence threshold T is a pair of $(\phi(A), B)$, where $A \subseteq X$, $B \subseteq M$, $A^* = B$, and $B^* = A$. Each object $x \in A$ has a membership μ_x defined as $\mu_x = \min_{m \in B} \mu(x, m)$, where $\mu(x, m)$ = membership value between object x and attribute m with I, $\phi(A) = \{\mu_x / x \mid x \in A\}$. If B is empty, then $\mu_x = 1$ for any $x \in A$. A and B are the extent and intent of the fuzzy concept $(\phi(A), B)$, respectively. $F(X, M, I)$ denotes the set of all fuzzy concepts of the fuzzy context (X, M, I) with the confidence threshold T.

Definition 2.3. For $(\phi(A_1), B_1)$, $(\phi(A_2), B_2) \in F(X, M, I)$, define $(\phi(A_1), B_1) \leq (\phi(A_2), B_2) \Leftrightarrow \phi(A_1) \subseteq \phi(A_2)$ (which is equivalent to $B_2 \subseteq B_1$). $(\phi(A_1), B_1)$ is called the sub-concept of $(\phi(A_2), B_2)$, $(\phi(A_2), B_2)$ is called the sup-concept of $(\phi(A_1), B_1)$.

Lemma 2.1. ([11]) Let (X, M, I) be a fuzzy formal context, $(F(X, M, I), \leq)$ is complete lattice, in which meet and join given by:

$$(\phi(A_1), B_1) \wedge (\phi(A_2), B_2) = (\phi(A_1 \cap A_2), (B_1 \cup B_2)^{**}), \quad (\phi(A_1), B_1) \vee (\phi(A_2), B_2) = (\phi((A_1 \cup A_2)^{**}), B_1 \cap B_2).$$

Definition 2.4. ([12]) Let L be a lattice, $a \in L$. If for any $a, b \in L$, $c = a \wedge b \Rightarrow c = a$ or $c = b$; c is called a meet-irreducible($\wedge$) element; if for any $a, b \in L$, $c = a \vee b \Rightarrow c = a$ or $c = b$; a is called a join-irreducible$(\vee)$ element.

In Hasse Diagram, meet-irreducible element is linked upwards by only one edge, join-irreducible elements are linked downwards by only one edge.

Related similarity models based on FCA and FFCA

In [6], Rodriguez and Egenhofer have proposed an assessment of semantic similarity among entity classes in different ontologies based on the normalization of Tversky's similarity model [13],

$$S_{RE}(a, b) = \frac{|B_1 \cap B_2|}{|B_1 \cap B_2| + \alpha(a, b)|B_1 - B_2| + (1 - \alpha(a, b))|B_2 - B_1|}, \quad (1)$$

where the function $|\cdot|$ represents the cardinality of a set, $\alpha(a, b) = depth(a) / (depth(a) + depth(b))$, if $depth(a) \leq depth(b)$; $\alpha(a, b) = 1 - depth(a) / (depth(a) + depth(b))$, if $depth(a) > depth(b)$.

Another modification of Tversky's model is presented by de Souza and Davis [3] to use the intents of structural elements of the lattice (meet-irreducible elements) as features. By using concept lattice operations (join $(\vee)$ and meet $(\wedge)$), the set of common meet-irreducible elements is given by the meet-irreducible elements which are the intent of $a \vee b$. This set will be represented as $(a \vee b)^{\wedge}$. The similarity model is as follows:

$$S_{SD}(a,b) = \frac{|(a \vee b)^{\wedge}|}{|(a \vee b)^{\wedge}| + \omega |(a-b)^{\wedge}| + (1-\omega)|(b-a)^{\wedge}|}, \quad (2)$$

Where $(a-b)^{\wedge}$ is the set of meet-irreducible elements which are in *a* but not in *b*, and $(b-a)^{\wedge}$ is the set of meet-irreducible elements which are in *b* but not in *a*. The value of ω is within the range [0, 0.5]. This kind of measure combines featural and structural information into decision, which is easy to understand and more reliable.

To further improve the performance of the concept lattice-based similarity measure, Zhao et al. [2] and Wang et al. [14] extended de Souza and Davis's measure by virtue of rough set theory, respectively. In their models, meet-irreducible elements are replaced by the intents (extents) corresponding to the different lower approximation of attributes (objects). These models are also suitable for concept approximation.

Above mentioned measures can only deal with typical ontology merging and ontology mapping. To tackle ontology merging and ontology mapping with uncertainty, few measures based fuzzy concept lattice have been given to evaluate the similarities between different fuzzy concepts. In [9], the similarity of a fuzzy formal concept $a = (\phi(A_1), B_1)$ and its sub-concept $b = (\phi(A_2), B_2)$ is defined as: $(a,b) = |\phi(A_1) \cap \phi(A_2)| / |\phi(A_1) \cup \phi(A_2)|$ .This similarity has been used in clustering quality evaluation and fuzzy ontology generation based on FFCA [8]. Although these models are easy to be implemented, the structure information of fuzzy lattice is not taken into consideration.

Similarity model based on FFCA

Inspired by above methods, a new model to evaluate the similarity of fuzzy concepts is introduced.

Similarity model based on lattice structures

For any $a, b \in F(X, M, I)$, define similarity measure of *a*, *b* as follows:

$$S_F(a,b) = \omega \frac{|a^{\vee} \cap b^{\vee}|_F}{|a^{\vee} \cap b^{\vee}|_F + \beta |(a-b)^{\vee}|_F + (1-\beta)|(b-a)^{\vee}|_F} + (1-\omega)\frac{|(a \vee b)^{\wedge}|}{|(a \vee b)^{\wedge}| + \gamma |(a-b)^{\wedge}| + (1-\gamma)|(b-a)^{\wedge}|}$$

where $a^{\vee}$ represents the fuzzy set of the extent of join-irreducible elements of *a*. $a^{\vee} \cap b^{\vee}$ represents the conjunction of the fuzzy sets $a^{\vee}$ and $b^{\vee}$. $(a-b)^{\vee}$ is the set of objects which are in $a^{\vee}$ but not in $b^{\vee}$, and $(b-a)^{\vee}$ is the set of objects which are in $b^{\vee}$ but not in $a^{\vee}$. $|a^{\vee}|_F$ represents the cardinality of the fuzzy set $a^{\vee}$. $|a^{\wedge}|, (a-b)^{\wedge}, (b-a)^{\wedge}$ defined as the model (Eq.2). The parameters β and γ are introduced in the model as Tversky's model to consider psychological elements that the assessment of similarity is not symmetrical. One example usually cited is that people consider North Korean more similar to Chinese than Chinese to North Korean. The weight $\omega\,(0 \le \omega \le 1)$ can be established by the user. By adjusting the value of ω, one can consider different contribution of objects and attributes to enrich the flexibility of the method.

Analyses for the model

In the proposed model, not only the influence from meet-irreducible concepts, but also that from join-irreducible elements is considered. Meet-irreducible elements play an important role in lattice. Every meet-irreducible element is corresponded to one new attribute being added [4]. It is isolated. Every concept of fuzzy concept lattice can be written as a meet of other meet-irreducible elements. However, join-irreducible elements can play the same important role with as meet-irreducible

elements in lattice. Every join-irreducible element is corresponded to one new object being added. Every concept of fuzzy concept lattice can also be written as a join of other join-irreducible elements. Meet and join-irreducible elements contain main structure information to the lattice. So, the proposed method combines featural and structural information into decision.

Table 1. Fuzzy context with threshold 0.8

x_i	a	b	c	d	e
x_1	0.8	1.0	—	0.8	1.0
x_2	0.9	0.8	1.0	—	—
x_3	—	—	—	0.9	—
x_4	1.0	1.0	0.8	—	—

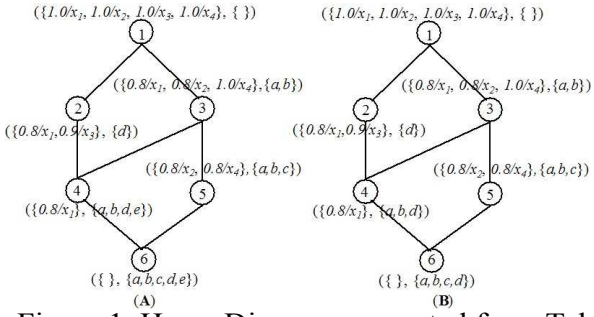


Figure 1. Hasse Diagram generated from Tab. 1.

For example, in (A) of the Fig. 1, the node 2 is meet-irreducible and introduces the attribute *d*, whereas the node 4 introduces the attribute *e* but is not meet-irreducible. That happens because *e* exists only in conjunction with the attributes *a*, *b* and *d*. They are not isolation. For this reason, the attribute *e* could be eliminated without any loss of structural information. The fuzzy concept lattice could be completely reconstructed without the presence of the attribute *e*. For example, in Fig.1, figure (B) generated from Tab.1 without the attribute *e* has the same structure with figure (A).

One point that is worth mentioning is that few attributes are usually added to one meet-irreducible element, for instance node 3 introduces attributes *a* and *b* in (A) of the Fig.1. From the point of structure of lattice, any of these attributes plays the same role with *e* in (A) of the Fig.1, so we take only one into account in the second part of the proposed model.

Example 4.1. As an example, we calculate the similarity measure between fuzzy concepts *a* (labeled by10) and *b* (labeled by 13) in Hasse Diagram (Fig. 2).

Table2. Fuzzy context with threshold 0.3 from [12]

	Price			Surface	
x_i	C_1	C_2	C_3	C_4	C_5
x_1	—	0.5	0.4	0.5	0.5
x_2	0.3	0.6	—	—	0.6
x_3	0.7	—	—	0.7	—
x_4	—	0.4	0.5	—	0.8
x_5	—	0.4	0.4	0.6	—
x_6	0.5	0.3	—	0.5	0.5

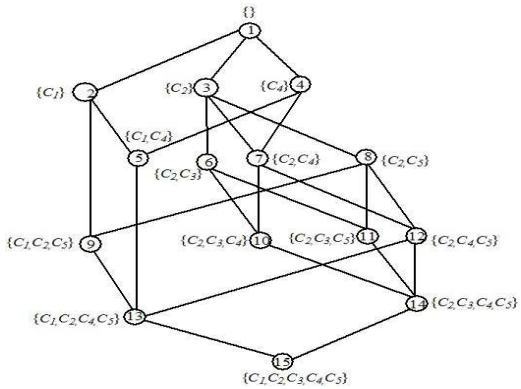


Figure2. Hasse Diagram of the Tab.2

$a=(\{0.3/x_2,0.3/x_6\},\{C_1,C_2,C_5\}), b=(\{0.5/x_1,0.3/x_6\},\{C_2,C_4,C_5\})$, $a^{\wedge}=\{C_1,C_2,C_5\}$,
$b^{\wedge}=\{C_2,C_4,C_5\}$, $a^{\vee}=\{0.3/x_2,0.3/x_6\}$, $b^{\vee}=\{0.5/x_1,0.3/x_6\}$, $(a\vee b)^{\wedge}=\{C_2,C_5\}$,
$(a^{\vee}\cap b^{\vee})=\{0.3/x_6\}$.

The value of β, γ and ω are fixed to be 0.5, which means that the assessment of similarity is symmetrical.

$$S_F^{\wedge}(a,b)=\frac{1}{2}\left(\frac{0.3}{0.3+0.3*0.5+0.5*0.5}+\frac{2}{2+0.5*1+0.5*1}\right)=0.5476.$$

In ontology mapping, let the similarity threshold be 0.54, then *a* and *b* can be mapped into each other. One can obtain mappings with different precision by adjusting threshold, and can get consider different importance of different particular attributes (objects) to the similarity model.

Conclusion

In this paper, a new similarity measure based on lattice structural information is established. The proposed method preserves more structural information of the source ontologies in ontology mapping, which is another extension and development of de Souza and Davis's similarity model in fuzzy context.

References

[1] X.J. Wan: Beyond topical similarity: a structural similarity measure for retrieving highly similar documents, Knowledge and Information Systems, 15(2008), p.55-73.

[2] Y. Zhao, W. Halang, X. Wang: Rough ontology mapping in E-Business integratio, Studies in Computational Intelligence, 37(2007), p.75-93.

[3] K. de Souza, J. Davis: Aligning ontologies and evaluating concept similarities, in: On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE, Springer Berlin (2004), p.1012-1029.

[4] G. Stumme, A. Maedche: FCA-MERGE: bottom-up merging of ontologies, in: Proc 7th Intl. Conf. on Artificial Intelligence, Seattle,WA,USA, (2001), p.1-6.

[5] T. Hsieh, K.Tsai, C. Chen et al.: Query-based ontology approach for semantic search, in: 6th International Conference on Machine Learning and Cybernetics, Hongkong, China (2007), p.2970-2975.

[6] M.A. Rodriguez, M.J. Egenhofer: Determining semantic similarity among entity classes from different ontologies, IEEE Transactions on Knowledge and Data Engineering, 15(2003), p.442-456.

[7] E. Sanchez, T. Yamanoi: Fuzzy ontologies for the semantic web, Lecture Notes in Computer Science, LNAI 4027, Springer Berlin (2006), p. 691-699.

[8] Q.T. Tho, S.C. Hui, A.C.M. Fong, T.H. Cao: Automatic fuzzy ontology generation for semantic web, IEEE Transactions on Knowledge and Data Engineering, 18 (2006), p. 842-856.

[9] R. Lau, Y.F. Li, Y. Xu: Mining fuzzy domain ontology from textual databases, IEEE/WIC/ACM International Conference on Web Intelligence, Silicon Valley, USA (2007), p.156-162

[10] W. Zhou, Z.T. Liu, Y. Zhao: Ontology learning by clustering based on fuzzy formal concept analysis, in: 31st Annual International Computer Software and Applications Conference, Beijing, China (2007), p.204-210.

[11] J.M. Ma, W.X. Zhang, S. Cai: Variable threshold concept lattice and dependence space, Lecture Notes in Computer Science, LNAI 4223, Springer Berlin (2006), p.109-118.

[12] X. Wang, W.X. Zhang: Relations of attribute reduction between object and property oriented concept lattice, Knowledge-Based Systems, 21(2008), p.398-403.

[13] A. Tversky: Features of similarity, Psychological Review, 84(1977), p.327-352.

[14] L.D. Wang, X.D. Liu: A new model of evaluating concept similarity, Knowledge-Based Systems, 21(2008), p. 842-846.

doi:10.4028/www.scientific.net/AMM.50-51.83

A Method of DEL Resources Matching DEL Needs

S.Q. Liu[1,2, a], L.Y. Zhang[1,b]

[1] School of Information Management, Wuhan University, Wuhan, China

[2]Strategic Planning College, Chongqing Technology & Business University, Chongqing, China

[a]lsqbrian@126.com, [b]lyzhang@whu.edu.cn

Keywords: DEL; urgency-attribute vector; AHP; fuzzy comprehensive evaluation method

Abstract. The dialectical relationship between disaster emergency logistics (DEL) resources and DEL needs shows that DEL tasks are rooted in DEL needs, and the implementations of DEL tasks are limited by DEL resources. What's more, the DEL resources shortage and the implementation of the DEL tasks directly affect the DEL needs. But how to reconcile the shortage of DEL resources and wide DEL needs will be very important, so a mathematical method for DEL contradiction is developed in this study. The aim is to coordinate logistics support for relief operations. The variety and extension of DEL needs can match the shortage of DEL resources. The urgency-attribute vector is made to represent the urgency degree of logistics task, and logistics task urgency order sequence can be got through the AHP based on logistics expert judging ability difference and fuzzy comprehensive method. So resources can be allotted to the task with the sequence of high priority, and with the sequence updating, the re-releasing logistics can be allocated again.

Introduction

Natural or man-made emergency events, such as floods, volcanic eruptions, earthquakes, or chemical spills, negatively affect society and environment, from the Taiwan Chi-Chi earthquake to the Indian Ocean tsunami, from the Katrina hurricane to Wenchuan earthquake, from Bhopal gas leak to Chernobyl accident. Human vulnerability, exacerbated by the lack of planning or appropriate emergency management, leads to financial, environmental or human losses [1]. The strengthening of emergency management is an inevitable choice for the disaster relief. Especially DEL management, the key to support emergency disaster relief, ensures the delivery of humanitarian assistance including the acquisition of commodities, transport arrangement and inventory management, etc. The paper first redefines and analyzes DEL from its needs, tasks and resources, and then puts forward urgency vector to show the urgency degree of disaster relief mission. AHP and fuzzy ranking method are used to rank various relief missions, so as to accomplish priority selection of DEL task in a shortage of logistics resources, and to alleviate their contradictions.

Fundamental analysis

After disaster happens, logistics has significant differences with the passage of time. The logistical responses in a disaster may be divided into emergent phase and stable phase. In emergent phase, the top priorities for logistics are to save people's lives and property, to displace people, to relocate victims and find shelter for refugees, and this emergent abnormal logistics should be government-led and be widely recognized and supported by the community; in stable phase, the primary tasks for logistics are to meet the social needs in disaster area, to recover and reconstruct disaster area, and this normal logistics should be enterprise-oriented, guided by government, and also supported by the community. The paper focuses on the first phase of DEL.

DEL is a series of DEL tasks the DEL providers implement to meet the DEL needs by using DEL resources. DEL providers include government departments, military units, enterprises, social organizations, groups, individuals, or even victims themselves. DEL providers, victims and the

Emergency Command Center is not only a general contingency logistics tasks, and users of logistics resources, or have emergency needs of the main logistics. Sustaining the lives of, victims surviving and emergency displacement property relief and people rescue personnel also produces logistics relief needs from their basic life need, their rescue act, etc. The movements of persons are a source of DEL needs. Emergency Command Center brings logistics needs based on geographic information, weather forecasting, disease outbreaks, transportation, medical rescue and environmental pollution, and the early warning and monitoring information.

DEL needs refers to the material requirements in space, time, and expense, related to transportation, inventory, packing, loading and unloading handling and the like within a certain time (such as gold 72 hours). And the needs are the interacting results of the socio-economic activities, suppliers and victims, are also the manifestation of inter-relationship of social economic development remedy. DEL need is derived, it is a derivative of the role of socio-economic activities in the disaster; it's universal, with a wide range of disaster relief to overcome the barriers of time and space in the regular activities; it's with diversity based on a various victims and rescuers; it's with the timeliness and region under a emergency scope, and in the same space, similar needs and similar disaster regional characteristics happen. All are related with the disaster regional features [2].

DEL tasks are various relief logistics services provided by community, determined by DEL needs. Such services need to adapt to diversity, universality, timeliness and region nature of DEL needs. These characteristics ensure that DEL tasks can be divided and combined. A logistics task randomly generated, if large, can be divided into a small one, and if small, can be combined into a large one. The granularity of DEL task decomposition or combination must have a relative independence of each other, and should not overlap, or overlap as little as possible to reflect the division of logistics task labor and its differential advantage; at the same time granularity should also match core functions and business units of actors in disaster. Excessively detailed decomposition results in the complexity of the problem, and is not helpful to the allocation of resources and actor choice in disaster, but also is to bring the complexity of management coordination; On the contrary, excessively large combination, can not reflect the personalization of logistics need and the competitive advantage of actors in specific emergency areas. According to DEL need and relief logistics resource, a task can be decomposed into different horizontal or vertical segments. Such a wide range of DEL tasks are created with relative independence and fit granularity.

To complete DEL tasks, actors must have relief logistics resources, including storage resources, transportation resources, equipment and all the logistics devices and facilities used for relief logistics activities. The narrow definition of logistics resources does not include human resources, which can be reflected in actors in disaster. The DEL resources have a shortage according to the DEL needs, and can not meet the need timely and efficiently. As is the case in the Wenchuan earthquake, the traffic disruption, road twisted, broken plumbing, air shut down, together with the interruption of communication, bad weather, caused the shortage of DEL resources. The shortage, sometime to some extent, the scarcity of DEL resources, roots in the economic scarcity of social resources, and exacerbated by the lack of appropriate DEL management and poor logistics infrastructure. The shortage also reflects the scarcity of social resources and the non-scarcity of disasters [2].

Thus, the relations between DEL needs, resources and actors are dialectical and interactive: the task derives from need; the task implementation is limited resources; the shortage of resource and the task implementing both influences the fulfillment of needs. So all the logistics tasks can not be completed, we must weigh the pros and cons and make a choice. We select those tasks that can save more lives and property. The urgent vector can be made to show the emergency degree of tasks. We can choose to complete the first row in the sequence of tasks, so that the logistics resources can be used effectively to alleviate the disaster. What's more, the urgent vector can also be used to show the disaster degree and the capacity of logistics resources. it will be useful and key to disaster strategic management.

Urgency degree vector

Sheu (2007) groups the affected areas using fuzzy-clustering techniques in which each affected area is characterized with multiple attributes to indicate the degree of relief demand [3] .Considering the existence of qualitative and quantitative attributes exhibited in characterizing the urgency of relief demands in the affected areas, four urgency attributes, used as the determinants of the grouping affected areas are specified. The degree of logistics relief urgency and disaster can show the urgency degree of logistics relief tasks, in which the logistics relief urgency can be shown by the time lead; and the disaster can be reflected by the logistics infrastructure, casualty number or casualty ratio, the number of the helpless and the destroying degree.

(1) Time lead U_t^1, represents the period of time between the initiation of any process of production and the completion of that process. Thus the lead time associated with the production of DEL task to the completion of the task may be anywhere from 1 hour to 1 week. The longer the lead time is, the more urgent the task is.

(2) Infrastructure U_t^2, represents the situation of logistics facilities, such as roads, airport and rail, inventory and so on. The poorer the infrastructure is, the more the more urgent the task is.

(3) Casualty number U_t^3, represents the ratio of, or the number of casualties observed relative to the total number of population trapped in a given affected area where the logistics task can cover. In general, a higher casualty ratio associated with the task may indicate higher urgency of the task.

(4) Helpless number U_t^4, represents the proportion of the helpless group including children and the elders relative to the total number of population trapped in a given affected area where the logistics task can cover. According to Shiono and Krimgold (1989), the survival probability of trapped victims in the disasters may decrease with time, depending on their physical conditions and severity of injuries. Therein, children and the elders can be regarded as the two most powerless groups who may need relief and rescue most urgently. Therefore, the corresponding proportion is taken into account in determining the emergency task priority. [4]

(5) Damage condition U_t^5, refers to the significance of building damage conditions, such as serious or complete destruction measured in a given time where the task covers In general, the damage conditions of the building may reflect the severity of the disaster effects on the survival probabilities of the trapped people. Accordingly, a relatively greater degree of building damage conditions may indicate a higher urgency of relief task.

Accordingly, we then have a 5×1 urgency-attribute vector associated with each given logistics tasks: $U_t = \left(U_t^1,\ U_t^2,\ U_t^3,\ U_t^4,\ U_t^5\right)^T$. Correspondingly, up to this stage, each given affected area can be represented by a multi-attribute datum characterized with a 4×1 urgency-attribute vector.

Urgency degree rating

Using the specified urgency-attribute vector, a fuzzy clustering-based algorithm is then employed to perform the function of affected-area grouping. The proposed algorithm is primarily composed of three sequential procedures including: (1) binary transformation, (2) generation of fuzzy correlation matrix, and (3) clustering. The fundamentals of these procedures are described in the following.

First, analytic Hierarchy Process（AHP）method can be used to weigh the influence of each factor in the urgent vector, and then fuzzy comprehensive method can be used to rank the priority of DEL tasks. We can evaluate the urgency degrees step by the following steps: To build the decision hierarchy; to consult experts; to calculate criteria weight; to rating the alternatives. We corporate related methods in each method. In the experts consultant step, On the basis of judging ability of experts, YanXi LI; Zhi TAO (2002) gave a value of judging ability, and then through a method of proper mathematical disposal, obtained the best judging matrix. If it doesn't fit the precision requirement, it will be returned and evaluated again by experts. [5]

Step1: to build the decision hierarchy

M (DEL experts) (E_k, k = 1, 2...m) participate in constructing judging matrix. The goal is the urgency degree of DEL tasks, and the criteria elements are the five urgency vector attributes.

Step2: to consult experts

We invite 25 experts from different fields, such as industries, universities, research institute, and government. There are many ways to get the comparison matrix results from experts. The Delphi method is usually applied. The Delphi method aims to obtain the consensus of experts by using a questionnaire survey. This survey method allows experts to express their opinion freely and privately.

Step3: to calculate criteria weight

(1) Get pair-wise comparison matrix A_k ($A_k = \left(a_{ij}^{(k)}\right)_{\mathrm{m}\times 5}$, k is the number of experts; 5 means the number of vector attributes) from each expert. The matrix is constructed by using a scale of relative importance. (2)Use eigenvector to calculate each comparison matrix weight. $A_k W = \lambda_{\max}^{(k)} W$, the eigenvector is denoted as w k is the number of experts. (3) Check the consistency of each comparison matrix. Judging ability is different, and the consistency index $\lambda_{\max}^{(k)}$ is calculated as follows: $CI^{(k)} = \dfrac{\lambda_{\max}^{(k)} - 5}{4}$ (k=1, 2… m), k is the number of experts. Obtain expert judging ability weight. Expert judging ability is inversely proportional to $CI^{(k)}$, because the value of $CI^{(k)}$ may be sequential, so on the condition that one appropriate decreasing continuous function f(x) may be chosen to regard as the relation function of between $CI^{(k)}$ and R_k. R_k can be determined by $CI^{(k)}$ .We define f(x) = $e^{-10(\mathrm{m}-1)x}$ R_k= $e^{-10(\mathrm{m}-1)CI^{(k)}}$ Normalize R_k, R_k= $\dfrac{e^{-10(\mathrm{m}-1)CI^{(k)}}}{\sum_{L=1}^{\mathrm{m}} e^{-10(\mathrm{m}-1)CI^{(k)}}}$ If one $CI^{(k)}$ is poor, the expert judging ability is poor. Then the R_k is smaller, and in the composite decision-making, he affects less; and vice versa. (4) Use R_k to combine group judgments. Combine each judging matrix A_k and get the group judgment matrix $\hat{A}_k$. (5) Obtain composite criteria weight.

The composite weights are obtained by multiplying the relative normalized weight of each factor with its corresponding normalized weight value for each alternative and making summation over all factors for each alternative.

Step 4: to rate the alternatives

The composite weights are obtained by multiplying the relative normalized weight of each factor with its corresponding normalized weight value for each alternative and making summation over all the factors for each alternative. At last but not least, we must comply with the allocation principles: (1) Integrate DEL resources. Logistics resources in a scattered, chaotic state must be first integrated, to be allocated and utilized effectively. (2)Higher in sequence the task is, allocation is in earlier priority; lower the task is, allocation is later in priority. Logistics resources can play the greatest social value. (3)With the same level in sequence, the tasks are assigned in priority, in accordance with the logic between the tasks was assigned the same priority. (4)Once the task is completed, the task should be deleted, and release logistics resources for reallocation. (5)Randomly new tasks are generated, reorder the sequence, and then follow the 2nd principle. (6)With the completion of the task, sequence should be gradually reduced until till the end.

DEL resource needs and tasks should be updated and fluctuate with the changes of disaster, so that the sequence of tasks varies in length.

Case analysis

25 experts participate in the decision-making of the four emergency tasks (T1, T2, T3, T4) rank.

(1)First of all six well-known experts selected in 25 (Ek, k = 1, 2... 6) participate in the construct the judging matrix and get each pair-wise comparison matrix A_k for the urgency vector elements,

(2) And then the eigenvector and max Eigen value can be got by using MATLAB. The normalized R_k is [0.57869644074753 0.41885037259417 0.00000002183029 0.00172070112311 0.00000905321945 0.00072341048545] (k = 1,2,...,6)

(3)Composite judging matrix $\hat{A}_k$ is got, and its max Eigen value and consistency are following: $\lambda_{max} = 5.39465182\ 287130$; $CI = 0.09866295571782$; $CR = 0.0409192474805 < 0.1$. The eigenvector is as following: [0.67397086547223 0.18752745433995 0.27534271812042 0.20310906047124 0.62731955453018]; normalized weights of urgency vector are [0.342592 0.095324 0.139962 0.103244 0.318878] (Hold 6 digits after the decimal point).

(4) 25 experts judge the first task with the Fuzzy comprehensive evaluation: membership matrix $R = (r_{ij})_{5\times4}$; urgency vector $W = (w_i)_{1\times5}$; Rank vector $D = (d_j)_{1\times4}$. Comprehensive membership vector S=W•R. To sum up, $\mu = S \bullet D^T$, and the score of T1 is 80.042521(table 1); also, T2's is 85.153062, T3's is 78.849009, and T4's is 65.736531. And the sequence of tasks is: T2>T1>T3>T4.

Table1: Fuzzy comprehensive evaluation of T1

Rank / tickets membership / Urgency vector		strongest 100	stronger 85	common 70	week 55
U_t^1	0.342592	9 0.36	14 o.56	2 0.08	0 0.00
U_t^2	0.095324	3 0.12	14 0.56	7 0.28	1 0.04
U_t^3	0.139962	5 0.20	15 0.60	5 0.20	0 0.00
U_t^4	0.103244	1 0.04	10 0.40	11 0.44	3 0.12
U_t^5	0.318878	2 0,08	11 0.44	12 0.48	0 0.00
	Comprehensive membership	0.12168	0.471050	0.362338	0.044924
Comprehensive Evaluation	Sum	80.042521			

Summary

From the prospective of DEL needs, tasks and resources, the paper redefines the DEL concept, and analyses the dialectical relationships between them. A new method is made to solve the contradiction in the shortage of DEL resources and the variety and extension of DEL needs. The urgency-attribute vector is made to represent the urgency degree of logistics task, and logistics task urgency order sequence can be got through the AHP based on logistics expert judging ability difference and fuzzy comprehensive method. So resources can be allotted to the task with the sequence of high priority, and with the sequence updating, the re-releasing logistics can be allocated again.

References

[1]L.H.Jiang and H.Y.Fang, in: *Military Economic Research*,Vol.2(2008), p. 38-40.

[2]A.H.He, in: *Disaster economics*, Northwest Universtiy Press(2000)， p. 3-6.

[3] J.B.Sheu, in: *E-Logistics and Transportation Review*(2007), p. 43, 687-709.

[4]Shiono and Krimgold, in: *Proceedings of the International Workshop on Earthquake Injury Epidemiology,* The Johns Hopkins University(1989).

[5]Y.X.Li and Z.Tao, in: *Natural Science Edition of Shenyang normal university journal*, Vol.20, No.2(2002), p. 86 - 90.

©(2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.89

Design, Implementation and Performance Testing of Vector Control for a Flexible Biomimetic Octopus Arm

Y.K .Wang [1,a], C.Guo [1], J.li [1] , Y.W. Wang [1] and Z.L. Wang[1,b]

[1]School of Mechanical Engineering, Harbin Institute of Technology, Harbin, China, 150001

[a]wangyukui@hit.edu.cn, [b]wangzl@hit.edu.cn

Keywords: Vector Control, Shape Memory Alloy (SMA), Biomimetic Octopus Arm

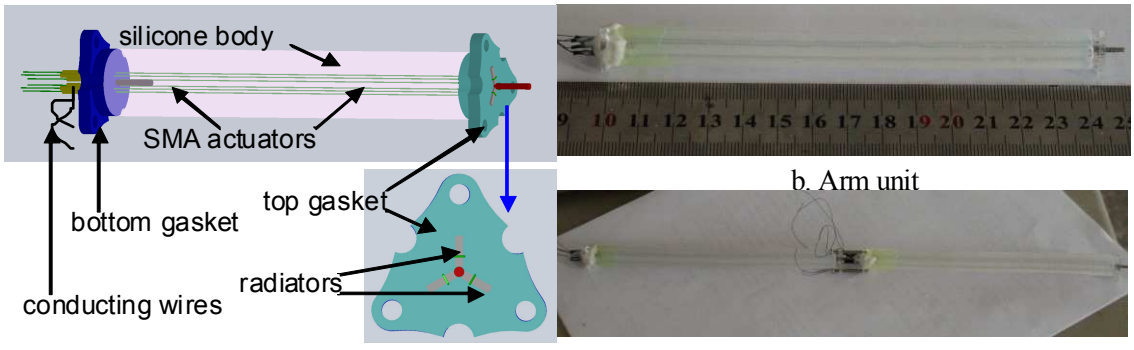
Abstract. According to the principle of muscular hydrostats, a flexible biomimetic octopus arm is put forward firstly. The method of vector control for the flexible biomimetic octopus arm is presented and analysed in detail. Based on the combination of ARM (advanced RISC microprocessor) and FPGA (Field Programmable Gate Arry), a control system for the flexibie biomiometic octopus arm is discussed, and the functions of every components is analysed. To verify the feasibility and reliability of the method of vector control, experimets are carried out. The experimemtal results demonstrate that this method applied in the control of the biomimetic octopus arm has a good performance.

Introduction

Flexible robots have been applied in unstructured and highly congested environments more and more, for they can exhibit complex movement with soft structures devoid of rigid components [1-3]. While, the control of these flexible robots is hard due to their soft structures. In this paper, a method of vector control for a flexible biomimetic octopus arm is put forward to solve the problem of orientation of the arm in space. Moreover, because the structure of the arm is similar to a kind of flexible robots[4-5], whose drivers is equally distributed in space, this method of vector control can be applied in these flexible robots. Hardware implementation for this method of control is also presented, the advantages of which is high-speed and low-cost. Finally, experiments of basic bending and rotation are carried out to verify the validity of this method.

Structure of the Flexible Biomimetic Octopus Arm

Inspired by the structure of octopus arm and bending function, the structure of the biomimetic arm unit is proposed (figure 1-a), which consist of a silicone body, two gaskets, three groups of SMA actuators and three radiators. The SMA actuators are embedded equally distributed in the silicone body along the longitudinal direction. In order to protect top gasket from overheating, radiator are fixed between SMA wires and top gasket, as is shown in figure 1 and in figure 1-b. If arm units connect in series, a flexible biomimetic octopus arm will be got (figure 1-c).



a. Unit structure c. Two-link structure

Fig.1 The biomimetic octopus arm prototype

SMA wire will contract when it is heated to by electric current or other methods, simultaneously, generate force. Its function is similar to the longitudinal muscle fibers of the octopus arm. This is the first necessary condition to flexible bending. Silicon is incompressible, whose volume is hardly changed by external force. So when one side of silicon is contracted, it will resist deformation, whose function is similar to the resistance muscle. The force generated by the side of the silicon contracted is the second necessary condition to flexible bending.

The Method of Vector Control

The vector control of the arm unit, in fact, is the control of the orientation angle β, which is generated by two groups of SMA actuators work in synchronism and the bending angle θ (figure 2). When these two parameters are determined, the spatial location of the arm unit is fixed. The orientation angle β is related to vector synthesis, so it can be changed by changing the proportion of t_1 and t_2, as is shown in figure 2.

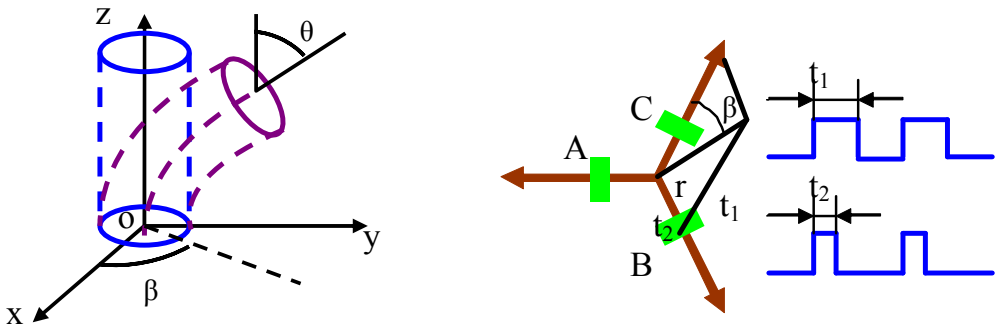


Fig.2 The meanings of parameters

When the required orientation is 0o, 120o or 240o, actuator A, B or C. In addition, the other required orientation angles need to be edtimated in which quadrant they are, then two groups of SMA actuators on boundaries of this quadrant work in synchronism, which can achieve the aim of vector control. In detail, firstly, the orientation angle determines the quadrant and actuators, secondly, referenced pulse width should be fixed according to bending angle, finally, pulse width of each actuator which is corresponding to each orientation angle is calculated. As SMA actuators are driven by PWM, pulse width is regarded as the controlling parameter when the voltage is fixed. The required orientation angle is β, γ, t_1 and t_2 are the pulse width corresponding to each actuator , r is referenced pulse width. Further more, referenced pulse width r is corresponding to bending angle θ, so r influences the size of t_1 and t_2. Orientation angle β influences the proportion of t_1 and t_2.

According to the mathematical relationship, as shows in figure 5, we can get the model of vector control as follows:

$$\cos\beta = \frac{t_1^{\,2} + r^2 - t_2^{\,2}}{2 \times t_1 \times r} \tag{1}$$

$$\cos\gamma = \frac{t_2^{\,2} + t_1^{\,2} - r^2}{2 \times t_2 \times t_1} \tag{2}$$

$$\frac{r}{\sin\gamma} = \frac{t_2}{\sin\beta} \tag{3}$$

Solve the abvoe equations regarding t_1 and t_2 as unknown numbers:

$$t_1 = \frac{\sqrt{3}}{3} \times r \times \sin\theta + r \times \cos\theta \tag{4}$$

$$t_2 = \frac{2\sqrt{3}}{3} \times r \times \sin\theta \tag{5}$$

Hardware Implementation

We implement algorithm for the method of vector control in an ARM (advanced RISC microprocessor) to achieve cost-effectiveness. An FPGA chip is designed for embedded-PWM controllers, by using a very high speed integrated circuit hardware description language (VHDL). Combining an FPGA chip and an ARM forms the low-cost intelligent controller. Using the high capacity of an FPGA chip, the additional hardware such as an encoder counter and a pulse-width-modulation (PWM) generator, can also be embedded into a single FPGA chip.

Datum of FPGA come from the bus of external memory of ARM, so the interface of FPGA and ARM should comply with time sequence. According to time sequence, when the signal of chip enable and the signal of write enable are low at the same time, the datum from address bus are translated into the signals of chip selection for PWM generator by the encoder counter embedded in FPGA. The selected PWM register gets the pulse width from the data bus (figure 3).Figure 3 shows a real figure of the embedded controller, consisting of an ARM, an FPGA, and drivers.

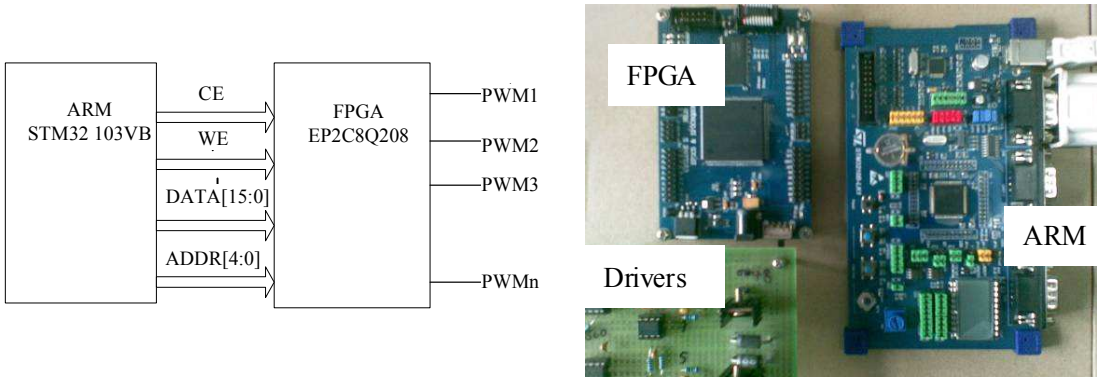


Fig.3 Interface between ARM and FPGA and Hardware implementation

Experimental Results

Basic Experiment of the Method of Vector Control. Vector control method show good accuracy ,as is shown in figure 4. Maximum error of control accuracy is 10%, The bending angles ,whose maximum error of control accuracy is 10% ,is far from 0°, 60° and 120°. The other three quadrant have the similar results.

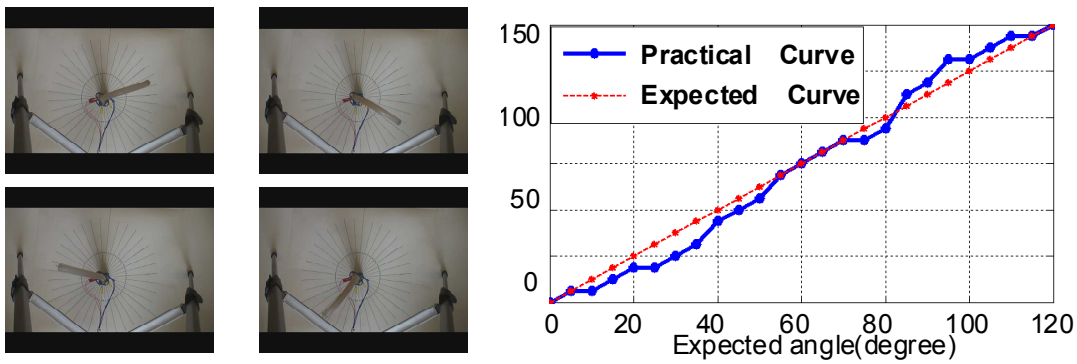


Fig.4 Experiment of vector control

Vector Control of Rotation. On the basis of the method of vector control, the arm unit can rotate(figure 5) by controlling the width of the pulse. In the experiment of rotation, a fixed 30% pulse is applied on one of the three groups of SMA actuators firstly, and a changable pulse is applied on another group of SMA actuators, whose width change from 0 to 90% in sequence. From the

experimental results, a small error exit between expected results and practical results, which is relative to the speed of the change of the width of the pulse which stimulate the driver and the environmental factors.

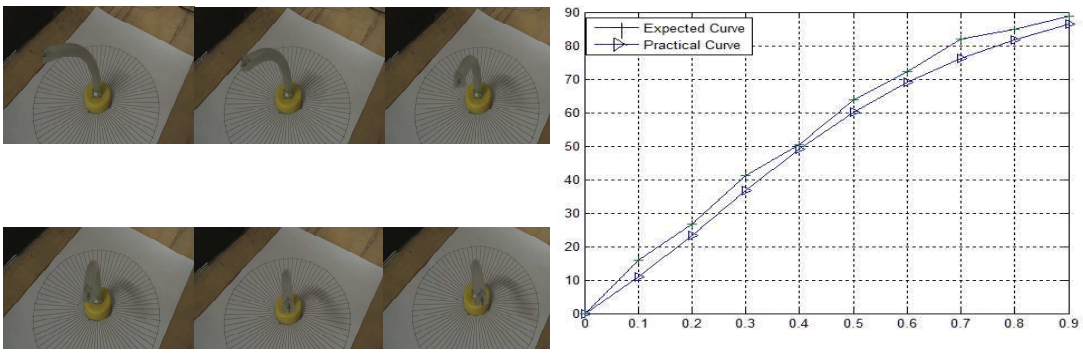


Fig.5 Experiment of rotation

Summary

A method of vector control for a flexible biomimetic octopus arm is presented, and the hardware implementation is also discussed. Based on the hardware, experimental results demonstrate the feasibility of the control method. Moreover, this metho of vector control can be used in the flexible robot, whose structure is similar to the arm unit by making some changes in future.

Acknowledgment

This work is supported by the National Natural Science Foundation of China under grant No. 50775049 and by Self-Planned Task (NO. SKLRS200805C) of State Key Laboratory of Robotics and System (HIT).

References

[1]Crespi, A. and Ijspeert, A. J. Amphibot ii: *In 9th International Conference on Climbing and Walking Robots (CLAWAR 2006)*, pages 19-27.

[2]Cieslak, R. and Morecki, A. Robotica (1999), 17:11-16.

[3]Trimmer, B., Takesian, A., and Sweet, B.(2006). *7th International Symposium on Technology and the Mine Problem, Monterey*, CA.

[4]William McMahan, Bryan Jones, and Ian D. Walker, . Pro. *IEEE/RSJ International Conference onIntelligent Robots and System (IROS), Edmonton, Canada,* 2005,pp. 3345-3352.

[5]Hideyuki Tsukagoshi, Ato Kitagawa, Mitsuru Segawa: *Proceedings of the 2001IEEE International Conference on Robotics & Automation,* Seoul, Korea, May21-26, 2001

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.93

The plastic work curvature criterion in evaluating gross plastic deformation for pressure vessels with conical heads

Mohammad Zehsaz [1, a], Farid Vakili Tahami [2,b] Yasser Ashraf Gandomi [3,c]

[1] Department of Mechanical Engineering, University of Tabriz, Tabriz, Iran

[2] Department of Mechanical Engineering, University of Tabriz, Tabriz, Iran

[3] Department of Mechanical Engineering, University of Tabriz, Tabriz, Iran

[a] zehsaz@tabrizu.ac.ir, [b] f_vakili@tabrizu.ac.ir, [c] ashrafgandomi86@ms.tabrizu.ac.ir

Keywords: Plastic work curvature criterion, Twice elastic slope criterion, Gross plastic deformation, Conical head, Plastic load

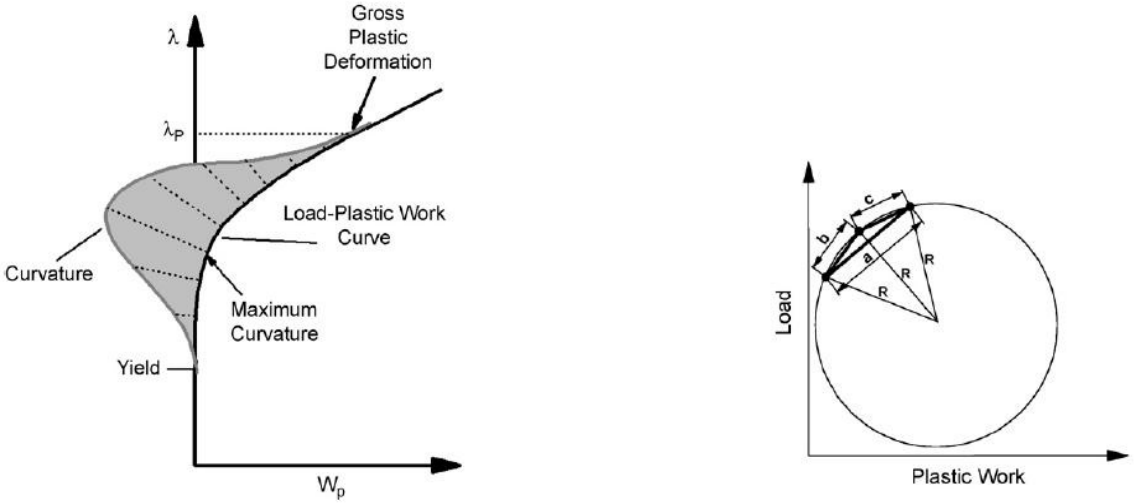
Abstract:
Conical shells are often joined to cylindrical shells and under internal pressure, the intersection between the large end of a cone and a cylinder is subjected to a large circumferential compressive stresses which can lead to its failure by whether axi-symmetric gross plastic deformation involving excessive inward deformations or non-symmetric buckling which accompanies by circumferential waves around the intersection. In Pressure Vessel Design by Analysis, the designer is required to address both these behavior modes when specifying the allowable static load. In this paper, plastic collapse or gross plastic deformation load is evaluated for a typical pressure vessel with conical head using plastic work curvature criterion and the results are compared with other criteria suggested by international standards such as ASME. The plastic work curvature criterion is based on the plastic work dissipated in the structure as loading progresses and may be used for structures subject to a single load or a combination of multiple loads. The results of analyzing using this new criterion show that the plastic collapse load given by the plastic work curvature criterion is robust and consistent and is in the close agreement with the results from international codes. The most significant aspect of the proposed method in which plastic work curvature criterion has been used is that the plastic collapse loads are determined purely by the inelastic response of the structure and therefore they are not influenced by the initial elastic response: a problem with some other established plastic criteria. The results also show that the design load is mostly limited by the formation of an axi-symmetric gross plastic deformation in the intersection of the vessel prior to the formation of non-symmetric buckling modes.

1-Introduction

In the process of designing pressure vessels, it is important to consider all possible failure modes of a special structure. Metallic pressure vessel components such as dished ends show different modes of failure depending upon the geometrical and loading conditions. These are mainly gross plastic deformation (GPD) under static load, instability (Buckling), fatigue crack initiation at highly stressed locations under cyclic loading, time dependant material behavior at high temperature (Creep), progressive plastic deformation (Ratcheting) and so forth. In the majority of the practical cases, GPD under static loading is the fundamental failure mode that has been widely focused by international standards such as PD5500 unfired fusion welded pressure vessels [1], ASME boiler and pressure vessel code, section III and IV [2] and EN13445-3: unfired pressure vessels [3]. GPD occurs when a vessel (of ductile material) experiences excessive static load. Unless the load is limited, this excessive load ultimately leads to plastic collapse or rupture of the vessel.
According to the international design codes, two distinct stress analysis may be used in order to prevent GPD failure mechanism: elastic analysis and inelastic analysis. The elastic analysis has the advantage that it is a straightforward method but the complicated part of the procedure is relating the elastic stresses to the inelastic GPD failure mechanism.

The inelastic analysis contains post yield stress redistribution and the formation of the GPD mechanism. In the case of performing an inelastic analysis, the GPD load is defined directly from the simulated structural response. The inelastic analysis has the advantage that the results are obtained directly from the structural response and it does not need any stress categorization.
The inelastic analysis may be performed in two distinct styles to characterize the structural response: Limit analysis and Plastic analysis. In the Limit analysis, the material model is defined as elastic-perfectly plastic or rigid-perfectly plastic and the small deformation theory is applied. In this procedure, the load carrying capacity of the structure is limited by equilibrium considerations. Although this style is easily applicable, but it may be very conservative due to the fact that the real structures may show geometrical nonlinearities or may exhibit post-yield strain hardening material model. In order to simulate the real behavior of a structure, the plastic analysis is applied which covers the geometrical and material nonlinearity. Material nonlinearity is modeled as bilinear strain-hardening model. Geometrical nonlinearity or large deformation effect may enhance or diminish the load carrying capacity of the vessel. For example, geometrical nonlinearity can strengthen the cone-cylinder intersection of steel vessels, as the shell is dominated by tensile membrane stresses except at the intersection [4]. In this case the load-deflection curve keeps rising by increasing deformation, and therefore, the definition of failure has to be artificial and has to be based on limiting displacements or stiffness losses. Some of these definitions have been adopted in Gerdeen [5].
In this paper, the plastic behavior of a typical conical-head vessel has been analyzed using four different methods and also a new proposed one which is based on the plastic work curvature criterion. The results have been compared to verify the accuracy of this new method.

Plastic work curvature criterion: The plastic work curvature (PWC) criterion is based on the relationship between the plastic work's dissipation in the vessel and elastic strain energy using a characteristic curve shown in Fig. 1.



a: Plastic work curvature b: circumradius evaluation of curvaure

Fig. 1 Plastic work curvature criterion and circumradius evaluation of curvature [6]

The PWC criterion applies a geometric construction to define the GPD load. According to Li and Mackenzie [6], the curvature of the characteristic load-plastic work curve is used to define the GPD load, as illustrated in Fig. 1a. In the plot, the PWC is normalized with respect to the maximum value of PWC calculated in the analysis. In the elastic region, the curvature is zero indicating zero plastic deformation. Post yield, plastic stress redistribution occurs and the plastic work curvature (PWC) increases to a maximum as the plastic deformation mechanism develops. The maximum stress redistribution occurs at the load corresponding to the maximum PWC, where after it begins to decrease as the plastic deformation mechanism is established. When the PWC reaches a minimum constant or zero value, relatively little or no further plastic stress redistribution occurs in the vessel

unless a second plastic deformation mechanism is initiated in a formerly elastic region. At this stage, the structure exhibits constant or gross plastic deformation with increased loading and the corresponding load is therefore designated as the plastic load.
The load-PWC plot may be created from the numerical results of a FE analysis and plotted against applied pressure using a simple technique based on the circumradius of three points [7]. The plastic work corresponding to the applied load is calculated by the FE program for each load step. Then, the results are written to a data file as a series of load-plastic work points. The curvature of a sector of curve defined by three consecutive points is the inverse of the circumradius of the three points. Eq. 1 gives the circumradius R of a triangle with side lengths of a, b and c, shown in Fig. 1b.

$$R = \frac{abc}{4\sqrt{s(a+b-s)(a+c-s)(b+c-s)}} \tag{1}$$

where s is the semiperimeter given by Eq. 2.

$$s = \frac{a+b+c}{2} \tag{2}$$

2-Model description

The application of PW and PWC criteria in evaluating the GPD load of a thin torispherical heads has been discussed by Mackenzie et all. [6, 8]. Furthermore, the GPD and buckling loads of thin wall conical heads has been thoroughly investigated by Teng et all [9,10]. In this paper we intend to investigate the plastic load of a conical vessel under internal pressure using the PWC plastic criterion. The geometry of the model has been defined in Fig. 2 and more details are given in Table 1. For the intersection of the large end of a cone and a cylinder, internal pressurization cause large circumferential compressive stresses in the intersection (Fig. 3). Due to this compressive hoop stresses, the intersection may undergo a failure in the form of axi-symmetric collapse involving excessive inward axi-symmetric deformations or non-symmetric buckling causing periodical waves around the circumference. Therefore, according to the possibility of buckling failure mechanism, it is necessary to model the structure in 3D.

Table 1: Geometrical Properties of the model

Symbol	Definition	Quantity
α	cone apex half angle [degree]	63.43
t1	thickness of the conical head [mm]	5
t2	thickness of the cylinder [mm]	5
R	radius of the cylinder [mm]	500
L	length of the cylinder [mm]	488

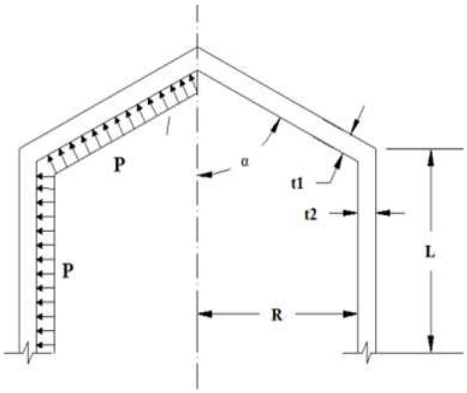


Fig. 2 Pressure vessel with conical head

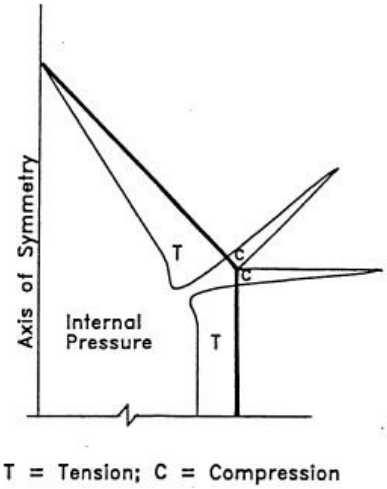


Fig. 3: Distribution of circumferential membrane stress in the vessel under internal pressure [11]

A FEA has been performed using the ANSYS program [12]. The 3D model has been developed using ANSYS 8-noded structural solid elements. The mesh of the model consisted of 23400 elements having 2 elements through thickness and is refined at the region next to the intersection, because, this region is likely to undergo the buckling failure mechanism. In order to model the evolution of the buckling deformation, the model was analyzed with 3D ANSYS 8-noded shell93 elements. The mesh of the 3D model consisted of 12312 elements. For the GPD analysis, displacements of the nodes in the far part of the cylinder are restricted in all the directions except radial direction. The end boundary conditions of the cylindrical shell in the buckling analysis, were defined to represent continuity in the shell (free for radial displacement and restrained against meridional rotation and axial and circumferential displacement). The length of the cylinder (L) is assumed to be 4λ (where λ is the linear elastic meridional half-wavelength and it is equal to $2.44\sqrt{Rt}$) to ensure that the boundary effects at the far end of the cylinder would not affect the behavior near the cone-to-cylinder joint [13]. In addition, a bilinear hardening material model was used in the analysis with material properties elastic modulus $E = 198.5kN/mm^2$, yield stress $\sigma_y = 265N/mm^2$, Poisson's ratio $\nu = 0.294$ and plastic modulus $E_p = 0.014E$.

3-Results

Plastic analysis: In the plastic analysis, as it was mentioned, the large deformation theory was applied and the bilinear hardening material model was considered as well. In addition, the 3D models were used to simulate the structural behavior.

Plastic analysis: TES and TI criteria: When applying the TES and TI criteria, it is necessary to specify a deformation parameter at a point on the vessel. Conical heads experience membrane deformation in the crown and cylindrical region and plastic hinge deformation in the intersection and the choice of location of deformation parameter should be made according to which region first experiences GPD. Three normal displacement parameters at the crown, intersection and cylinder were considered. In this case, the maximum applied pressure has been 1 MPa. The corresponding structural responses using TES and TI criteria have been illustrated in Fig. 4 and Fig. 5 respectively and the calculated plastic loads for the model have been given in Table 2.

According to Table 2, TES criterion suggests that the intersection region first experiences GPD and the corresponding plastic load is 0.8 MPa. TI criterion, on the other hand, points out that the crown region is as critical as the intersection region and predicts the plastic load to be 0.71 MPa. Both criteria, similarly, suggest that the cylindrical part does not experience GPD at all (because the lines do not intersect each other).

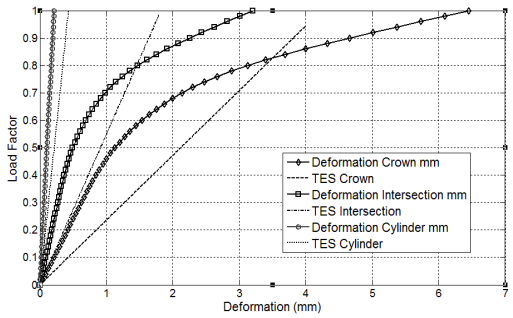


Fig. 4: TES criterion: Load Factor-Deformation

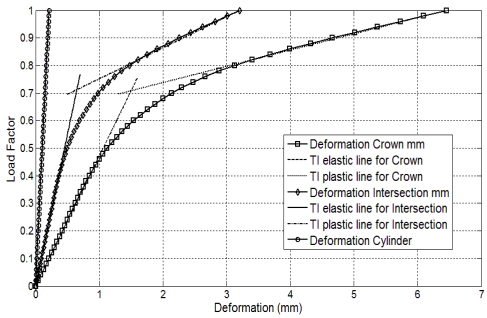


Fig. 5: TI criterion: Load Factor-Deformation

Table 2: Plastic loads based on TES and TI criteria [MPa]

	TES criterion	TI criterion
Plastic load (crown)	0.82	0.71
Plastic load (intersection)	0.8	0.72

Plastic analysis: PW criterion: The PW criterion, however, does not need any load or deformation parameter. In this analysis, the maximum applied pressure for the vessel has been 0.86 MPa. In this method, the variation of the Load Factor with Plastic Work has been obtained. When the Load Factor approches 0.5, the material yielding initiates and with further increase in the Load Factor, the plastic regions develop. When the structural response becomes steady (Fig. 6), the unlimited plastic flow establishes in the vessel. Therefore, the plastic load according to the PW crirerion, is 0.71 MPa.

Plastic analysis: PWC criterion: Similar to the PW criterion, the PWC criterion does not require any load or deformation parameter. In order to apply the PWC criterion, the accumulative plastic work for every substep obtained from ANSYS software has been saved and have been used as input data for a computer code developed in MATLAB software [14]. Using this code, the PWC of the vessel has been calculated. and the results are shown in Fig. 7. In this analysis the maximum applied pressure has been 0.86 MPa. According to Fig. 7, when the Load Factor is less than 0.37, the deformation of the vessel is completely elastic; by further increse in the pressre, the plastic regions initiate. In Fig. 7, every peak in the curve corresponds to the initiation of a new plastic region. The unlimited plastic flow develops in the vessel when the plastic-work-curvature reaches its almost steady or constant value of 0.2 at $Load\ Factor = 0.91$ and as a result, the structure does not bear any further pressure.

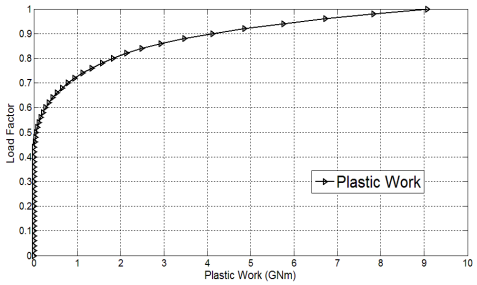


Fig. 6: Plastic analysis-PW criterion: Load Factor-Plastic Work

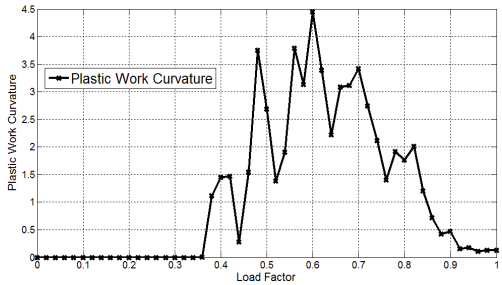


Fig. 7: Plastic analysis-PWC criterion: Plastic Work Curvature-Load Factor

Limit analysis: In contrast to the plastic analysis, in the limit analysis, the material of the vessel is assumed to be elastic-perfectly plastic and the small deformation theory is applied. In this analysis, normal displacement at the intersection region is considered as the deformation parameter. For this model, the maximum applied pressure has been 0.8 MPa and according to the results, in the limit Load Factor, the structural response approches a steady value and this, illustrates that the

vessel has reached its limit load. Therefore, the calculated limit load for this vessel based on limit Analysis method is $P_{\text{limit load}} = 0.54MPa$.

Buckling analysis: In conical vessels, owing to the internal pressure, the compressive hoop stresses arise in the intersection region and theses stresses may lead to the buckling of the vessel. In order to evaluate the dominant failure mechanism of the vessel, the buckling analysis has been carried out with the ANSYS software. According to Teng [13] when the structural response of the similar adjacent nodes in the intersection part of the vessel starts to diverge from each other, the buckling mode establishes in the vessel and the corresponding pressure is regarded as the buckling pressure. For this model, the calculated buckling pressure is 1.1 MPa. The plastic strain distribution prior to instability is shown in Fig. 8.

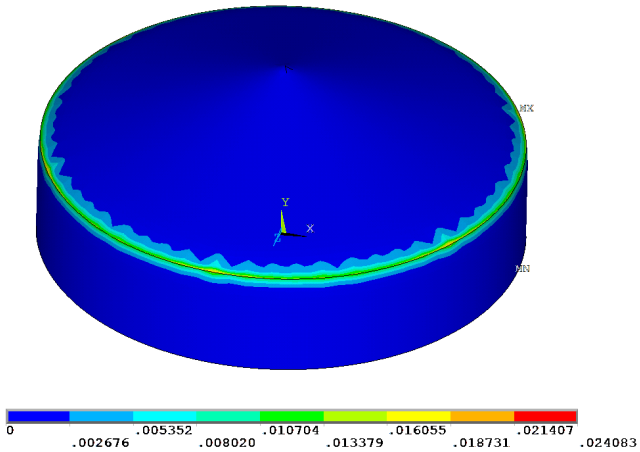


Fig. 8: Equivalent plastic strain distribution at numerical instability pressure

Table 3: Calculated failure loads for the conical vessel [MPa]

Plastic criterion	Plastic Load
Limit Analysis	0.55
TI criterion	0.71
TES criterion	0.8
PW criterion	0.71
PWC criterion	0.78
Instability Load	1.1

It can be seen in Fig. 8 that although the major portion of the vessel has remained elastic, the comperessive sresses in the intersection region has made the vessel to buckle.
In Table 3, the calculated plastic loads from different plasatic criteria and the corresponding numerical instability load have been summorised.
According to Table 3, the Limit analysis suggests the minimum plastic load and in contrast, the TES criterion predicts the maximum plastic load among the other plastic criteria. The GPD load predicted by the PWC criterion is slightly lower than the Plastic Load given by TES. The comparison of the limit and plastic analysises, shows that including non-linear aspects of geometry significantly affects the simulated response and the results of the analysis. This table also shows that buckling pressure is 1.1MPa which is greatly in excess of the limit load, indicating that the geometric strengthening occures when large deformation effects are considered. When the load is increased beyond the GPD pressure, the plastic deformation in the intersection becomes non-axisymmetric as local buckling occurs around the head. In addition, according to the Table 3, the GPD mechanism forms in the intersection region prior to circumferential buckling deformation.

4-Discussion and conclusion:

In this paper, in addition to the PWC criterion, the TES, TI, and PW criteria have been used to not only investigate the gross plastic deformation mechanism, but also to define the plastic limit loads for a typical conical-head vessel. The plastic load calculated using the TES or TI criterion depend on the location of the deformation parameter. Three deformation parameters were considered: normal displacement in the crown, intersection and cylinder regions of the vessel. The intersection parameter was found to give the most conservative value of plastic pressure. All methods, except TI, show that the intersection deformation parameter is more critical. Furthermore, the PW and PWC criteria indicated that GPD failure mechanism occurred in the intersection region of the vessel

prior to the formation of non-symmetric buckling and gross plastic membrane deformation of the crown or shell. This indicates that the use of the intersection deformation parameter is appropriate in the TES criterion.

In addition, it was shown that unlike the TES and TI criteria, in the PW and PWC criteria, the plastic loads were determined purely by the inelastic response of the structure, and were not influenced by the initial elastic response. This means that in the finite element analysis, the designer is allowed to omit the low stress regions from the model and represent them by suitable boundary conditions. The results, also, demonstrated that the PWC Pressure-Plastic Work plot would help the designer to identify the evolution of distinct plasticity mechanisms in different regions of the vessel under loading conditions.

Comparing the results of the small and large deformation analysis of the conical vessels show that the thin pressure vessels usually experience geometric strengthening in the process of loading. As a result, in analyzing the thin pressure vessels, the application of the large deformation theory is necessary.

In conclusion, the results show that the plastic pressure level obtained based on PWC criterion for the conical vessel, has a good agreement with the value obtained using ASME approved TES criterion.

References:

[1] BSI, PD5500:2000 specification for unfired fusion welded pressure vessels. London: British *Standards Institution* (2000).

[2] ASME Boiler and Pressure Vessel Code, Sec VIII, Div 1 & 2. *American Society of Mechanical Engineering*, New York (2005).

[3] EN 13445. Unfired pressure vessels-Part 3: design. *European Committee for Standardization,* (2002).

[4] Jin-Guang Teng. Cone-Cylinder Intersection Under Pressure: axi-symmetric failure. *Journal of Engineering Mechanics;* 120:1896-1912 (1994).

[5] Gerdeen JC. A critical evaluation of plastic behavior data and a united definition of plastic loads for pressure vessel components. *WRC Bull* ; 254 (1979).

[6] Donald Mackenzie, Duncan Camilleri, Robert Hamilton. Design by analysis of ductile failure and buckling in torispherical pressure vessel heads. *Thin-Walled Structures*; 46: 963–974 (2008).

[7] Weisstein EW. Circumradius. From MathWorld—a Wolfram web resource, Information on http://mathworld.wolfram.com/Circumradius.html

[8] Hongjun Li, Donald Mackenzie. Characterizing gross plastic deformation in design by analysis. *International Journal of Pressure Vessels and Piping* ; 82:777–786 (2005)

[9] Y. Zhaoa, J.G. Teng. A stability design proposal for cone–cylinder intersections under internal pressure. *International Journal of Pressure Vessels and Piping* ; 80:297–309 (2003)

[10] J.G. Teng, Y. Zhao. On the buckling failure of a pressure vessel with a conical end. *Engineering Failure Analysis* ;7:261-280 (2000)

[11] Jin-Guang Teng. Cone-Cylinder Intersection Under Pressure: non-symmetric buckling. *Journal of Engineering Mechanics ;*121:1298-1305 (1995).

[12] ANSYS version 10.0 (2005).

[13] Y. Zhao, J. G. Teng. Buckling Experiments on Cone-Cylinder Intersections Under Internal Pressure. *Journal of Engineering Mechanics* ;127:1231-1239 (2001).

[14] MATLAB version 7.5 (2007).

doi:10.4028/www.scientific.net/AMM.50-51.100

Study on the Effects of Train Live Loads on Isolated and Non-isolated Simply Supported Railway Bridges

Changyong Zhang[1, a], Tieyi Zhong[1, b], Kejian Chen[2, c] and Yunkang Gong[1, d]

[1] School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China

[2] China Railway Eryuan Engineering Group Co. Ltd, Chengdu, Sichuan 610031, China

[a]coolzcy1983@163.com, [b]tyzhong2008@163.com, [c]ckjian@163.com, [d]09121704@bjtu.edu.cn

Keywords: simply supported railway bridge, isolated bridge, non-isolated bridge, train live load, lead rubber bearing.

Abstract. In this paper, based on the finite element program ANSYS, the model of a simply supported railway bridge with and without isolation using lead rubber bearing is established. Seismic response time-history analyses of the bridge subjected to high-level earthquakes are carried out considering and not considering train live loads. Through the comparison and analyses of the results, the effects of train live loads on seismic calculation of non-isolated railway bridges and isolated railway bridges are obtained. The results of the research will support the further study on seismic design and isolation design of simply supported railway bridges.

Introduction

As known, several terrible earthquakes happened in the past few years which had brought huge casualty and loss to the local people. Especially, more than 69,000 people lost their lives in the earthquake which shook Sichuan province of China on May 12th, 2008 and more than 11,000 people died in the earthquake which struck Concepcion of Chile on January 12th, 2010 with many injured If bridge, as an important part of the lifeline projects, is damaged in the earthquake, that will cause enormous social and economic loss. There were many spectacular failures of bridges happened in every major seismic event, which severely affected the efficiency of the rescue operation after earthquakes. In the earthquake happened in Sichuan province, several important railway bridges were destroyed which made it quite difficult to transport the disaster relief goods and materials to disaster regions. Thus, seismic calculations and analyses of railway bridge structures have attracted more and more attention, especially for the railway bridges located in earthquake areas.

With the rapid development of modern railway engineering technology, train speed is increasing rapidly, while more and more railway bridges are built especially for high speed trains. At the same time, the possibilities of trains running on bridges during earthquakes are highly increased, for which the train live loads are becoming important parts in the seismic calculations of railway bridges..

According to the Chinese standard 'Code for Design on Subsoil and Foundation or Railway Bridge and Culvert' (2006) [1], train live load is not considered in seismic calculations for high-level earthquakes. With the increment of the possibility of trains running on bridges when earthquakes occur it is necessary to study the influence of train live loads on seismic calculations of railway bridges for high-level earthquakes. In this paper, the transverse seismic responses of a 32m simply supported railway bridge with and without isolation bearings subjected to high-level earthquakes are calculated based on FEA program ANSYS considering and not considering the transverse inertia forces induced by train live loads. Some meaningful results are acquired, which will support the further study on effects of train live loads on seismic design of railway bridges.

Calculation Models

A 32m simply supported railway bridge located in an earthquake area is taken as the example and the transverse vibrations of the bridge are mainly studied in this paper.

Structure model of the railway bridge. The chosen railway bridge is a 32m T-shape section box girder bridge and the pier is a round-end variable section concrete pier. Single pier model is adopted in this study. In the single pier model, the superstructure of the bridge is simplified as a 520 tons lump mass and the pier is fixed on the ground. Two different lead rubber bearings are used as isolators.

The FEA model of the bridge is established based on finite element program ANSYS. The superstructure is simulated by MASS21 element. The pier is modeled by BEAM3 element. In the model of isolated bridge, the lead rubber bearings (LRB) are simulated by nonlinear spring element COMBIN39. The sketch of the models without and with isolation is shown in Fig.1.

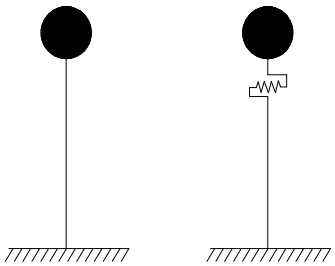


Figure 1 Sketch of calculation models without and with isolation

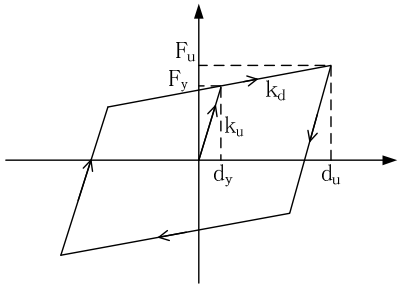


Figure 2 Sketch of bilinear model of LRB

Bilinear model of lead rubber bearing. As shown in Fig. 2 [2, 3, 4], F_y represents the yield stress of the bearing and d_y is the yield displacement of the bearing. k_u is the initial stiffness as well as k_d is the post-yield stiffness of the bearing. $\eta = k_d / k_u$ is the harden ratio of the bearing.

Two lead rubber bearings are considered in the research and the basic parameters of the bearings are given in Table 1.

Table 1 Mechanical parameters of LRB

Item	Initial stiffness k_u (10^7N/m)	Post-yield stiffness k_b (10^7N/m)	Stiffness ratio η	Yield force F_y (kN)
LRB1	6.1	0.183	0.03	610
LRB2	8.6	0.258	0.03	960

Train live loads. Two types of locomotive marshalling are considered in the research:

Train live load 1: Domestic 25G train or similar type. The composition of SS8 locomotive + 12 double deck passenger cars (SYZ25G) is adopted in this locomotive marshalling type.

Train live load 2: Domestic high speed train. CRH2 train is taken in this locomotive marshalling type.

In calculation, 50% of the train weights are applied on the beam in the transverse direction of the bridge according to Code for Design on Subsoil and Foundation or Railway Bridge and Culvert (2006) [1]. The basic parameters of the trains mentioned above are given in Table 2 [5].

Table 2 Basic parameters of locomotives

Train	Type	Total length (m)	Constant distance (m)	Wheel track (m)	Axle load (kN)
SS8+25G	Locomotive	17.516	9.000	2.900	228.00
	Passenger car	26.570	18.500	2.400	152.68
CRH2	Locomotive	25.000	17.500	2.500	135.00
	Passenger car	25.000	17.500	2.500	120.00

Seismic excitations. The railway bridge mentioned above locates in mountain areas where the site conditions are 'Site Ⅱ'. Ten ground motion records obtained on ground of Site Ⅱ are chosen as the earthquake excitations and they are regulated corresponding to high-level earthquakes with the intensities of 8 degree. Ten different working conditions are defined based on the aforesaid ground motion records. Table 3 gives out the basic characteristics of the earthquake records.

Table 3 Basic characteristics of earthquake records

Earthquake name	Work condition	Component	Magnitude	PGA(cm/s2)	Date
Santa Barbara	1	N45E	5.9	233.8	1941.6.30
N.California	2	N98W	5.2	193.8	1949.3.9
Kern County	3	S48E	7.7	128.6	1952.7.21
San Fernando	4	N45E	5.3	45.9	1957.3.22
Parkfield Califor.	5	N50E	5.6	232.6	1966.6.27
San Fernando	6	N29E	6.4	96.7	1971.2.9
San Fernando	7	S89W	6.4	154.1	1971.2.9
San Fernando	8	N37E	6.4	195.6	1971.2.9
Southern Califor.	9	N36W	6.0	52.9	1952.11.21
Parkfield Califor.	10	N36W	5.6	14.2	1966.6.27

Results and analyses

In this paper, pier vibration responses of the railway bridge with and without isolation considering and not considering train live loads in ten working conditions are calculated and compared. Maximum values of transverse shear forces at pier bottom of non-isolated bridge in ten working conditions are shown in Fig. 3.

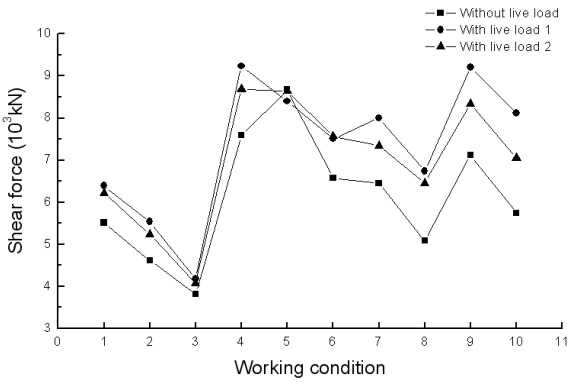


Figure 3 Maximum shear forces at pier bottom of non-isolated bridge with and without train live loads

Fig.3 shows that the transverse shear forces of non-isolated bridge considering live loads are larger than those not considering live load. The max margin considering live load 1 is 41.46% in working condition 10 and when considering live load 2 the max margin is 27.17% in working condition 8.

Maximum values of transverse shear forces at pier bottom of isolated bridges with bearing 1 and bearing 2 considering and not considering train live loads in ten working conditions are shown in Fig. 4 and Fig. 5.

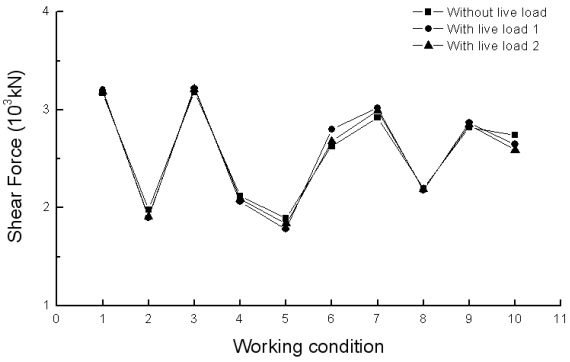


Figure 4 Maximum shear forces at pier bottom of the bridge isolated by bearing 1 with and without train live loads

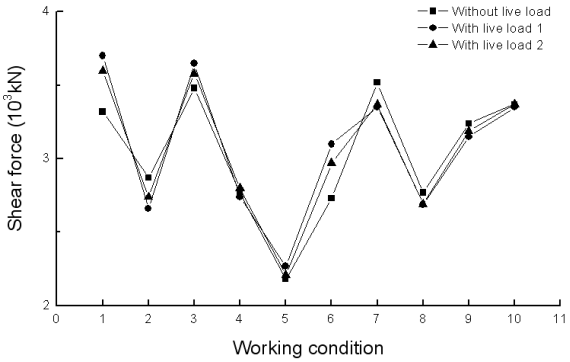


Figure 5 Maximum shear forces at pier bottom of the bridge isolated by bearing 2 with and without train live loads

It can be seen from Fig. 4 and Fig. 5 that shear forces at pier bottom of isolated bridge are very close considering or not considering train live loads. Especially when the bridge is isolated by bearing 1 whose stiffness is small, the shear forces are almost the same considering or not considering live loads. When isolated by bearing 2, the max margin between considering live load 1 and not considering live load is 13.55% in working condition 6 and for live load 2, the max margin is 8.79% in working condition 6.

Isolation ratios of bearing 1 and bearing 2 considering and not considering train live load in ten working conditions are shown in Fig. 6 and Fig. 7.

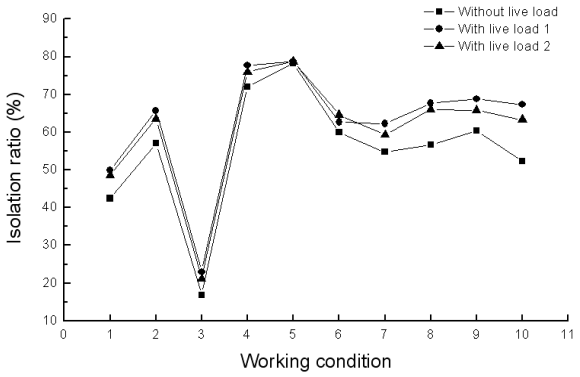


Figure 6 Isolation ratio of bearing 1 in ten working conditions with and without train live loads

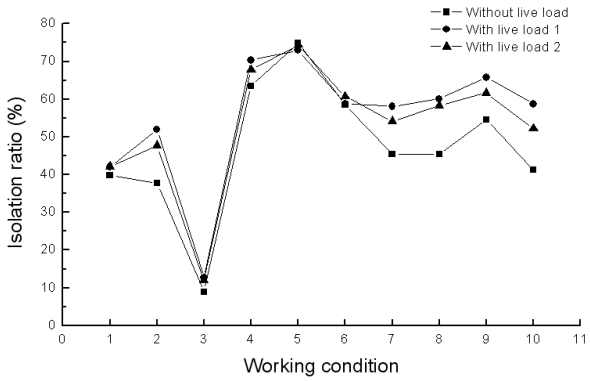


Figure 7 Isolation ratio of bearing 2 in ten working conditions with and without train live loads

It is shown in Fig. 6 and Fig. 7 that when considering live loads, isolation ratios of bearings are higher than those not considering live loads. For bearing 1, the max margin between isolation ratios is 37.08% and for bearing 2 the max margin is 42.46%. It is because after considering train live loads, the shear forces of non-isolated bridges are much enhanced but those of isolated bridges are less changed.

Apparently from Fig. 3 ~ Fig. 7, it can be seen that the inertia forces induced by train live loads have obvious effects on seismic responses calculation of non-isolated railway bridges, and the larger train live loads are, the more significant the effects become. For isolated railway bridges the effects of train live loads become much less, and the smaller the stiffness of isolated bridge is, the less the effects become.

Conclusions

In this paper, considering two types of locomotive marshalling and using two lead rubber bearings, the effects of train live loads on seismic responses calculation of a 32m simply supported railway bridge are studied. By comparing the calculation results of vibration responses in different working conditions it is found that the transverse shear forces of non-isolated bridge considering live loads are larger than those not considering live loads with the max margin up to 41.46% in ten different working conditions. For isolated bridge, shear forces at pier bottom are very close considering or not considering live loads.

Obviously, train live loads have significant effects on seismic calculation of non-isolated railway bridges and small effects on seismic calculation of isolated railway bridges. For non-isolated bridge, the larger train live loads are, the greater the effects become and for isolated bridge, the smaller the stiffness of isolated bridge is, the smaller the effects become.

Thus, for non-isolated railway bridges and the isolated railway bridges whose stiffnesses are big, train live loads should be considered in the seismic calculation when subjected to high-level earthquakes as well as low-level earthquakes and design earthquakes.

Acknowledgment

This study is sponsored by the Science & Technology Research Development Project of China Railway Ministry (Grant No. 2008G010-A) and the Science & Technology Research Development Project of China Railway Ministry (Grant No. 2008G028-C).

References

[1] Code for seismic design of railway engineering (GB 50111-2006).

[2] T.Y. Zhong, F.L. Yang and B. Wu, Bidirectional Seismic Response Analysis of Railway Simple Supported Beam Bridge, China Railway Science, 28(3): 38-43, 2007.

[3] F.L. Yang, T.Y. Zhong and H. XIA, Study on Optimization of the Design Parameter of Seismic Absorption and Isolation bearing for Railway Simple Supported Beam Bridges, Journal of the China Railway Society, 28(3): 128-132, 2006.

[4] T.Y. Zhong, F.L. Yang and H. XIA, The Design Method of the Seismically Isolated Bridge by Lead Rubber Bearing Based on Energy Method, China Railway Science, 30(2): 43-48, 2009.

[5] C.Y. Liu, C.Y. Zhang, M.B. Zhang and T.Y. Zhong, The Passive Control Study on Traffic-induced Vibration Responses of Bridge Pile, 4th International Symposium on Environmental Vibrations: Prediction, Monitoring, Mitigation and Evaluation, October 28-30, 2009, Beijing Jiaotong University, China, p. 1566, (2009).

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.105

Study on the Impact of Lead-Rubber Bearing Parameters on Seismic Responses of Seismically Isolated Pier

Yuanqing Xu[1, a], Tieyi Zhong[1, b], WengangJi[2, c], Xu Li[1, d]

[1]School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China,

[2]CCCC ROAD AND BRIDGE CONSULTANTS CO.LTD, 100029, China

[a]xuyuanqing470@163.com, [b]tyzhong2008@163.com, [c]wengangji@sina.com,

[d]lilixuxuwudi@126.com.

Keywords: lead-rubber bearing (LRB), isolation ratio, hardening ratio, initial shear stiffness, yield strength, hysteresis curve.

Abstract. Based on the finite element program SAP2000, 120 models of eight different piers with different lead-rubber bearings (LRB) are established. The seismic response analyses of piers isolated by LRB in axial direction of the bridge are carried out under the excitations of seven earthquake waves. The calculations of seismic responses of the piers with and without isolation are implemented by using dynamic time-history analysis method, and the comparison of the results is given out. Considering the isolation ratio of the moment at the bottom of pier as observation samples, the variances of the three LRB mechanical parameters are analyzed. Furthermore, the ability of the parameters of LRB to adjust the seismic response of different piers is discussed. The analysis results show that all the mechanical parameters of LRB have significant effect on the isolation ratio of the moment at the bottom of pier, and the parameters of LRB can adjust the seismic forces of the isolated piers whose height difference is less than 30 meters to be almost equal. This paper provides some valuable references for the further study of isolation design for railway bridges.

Introduction

Isolated bridge, designed with LRB, plays an effective role in the protection of itself during an earthquake [1, 2]. The seismic mitigation and isolation techniques have been applied in some railway bridge project in China. J.S. Zhuang and B. Wu have studied on mechanical characteristics of LRB and its application for seismic mitigation and isolation in bridge engineering [3]. D.S. Zhu has studied the influence of ductility and hardening ratio on isolation effectiveness. [4]. Professors G.P. Yan and T.Y. Zhong have studied on the performance of seismic mitigation and isolation for railway simple supported beam bridge by LRB [5]. Z.G. Zhang has focused on the performance of the vibration absorption and isolation of continuous beam railway bridge systematically by using the nonlinear model which considers the pier ductility of isolated bridge. However, the foundation stiffness was not discussed. Dr. F.L. Yang considered the soil-structure interaction and achieved some important conclusions [7]. However, there is no quantitative research about the significant influence of LRB on seismic responses and the adjustment ability of LRB on seismic force.

In this paper, based on an actual bridge, a series of single pier models are established and then dynamic time-history analyses are carried out by using the finite element software SAP2000. The calculations of seismic responses of the piers with and without isolation are implemented, and the comparison of results is present. Depending on the experimental datas about the parameters of LRB, which were studied by China Academy of Railway Sciences, the value range of the three parameters is set: hardening ratio $\eta \in (0.08, 0.5)$, initial shear stiffness K∈ (2.0E+7N/m, 1.0E+08N/m), and yield strength F∈ (1.0E+05N, 5.0E+05N). Also, the design of the bearing is based on the parameters which are selected with five different values from in the range above.

The calculation model and the selection of seismic waves

In this paper, several piers of a railway simply supported beam bridge are taken as models for calculations. Those piers are poured by C30 concrete and they are hollow, cylindrical with variable variable cross-section. The heights of those piers are 30 meters, 35 meters, 38 meters, 42 meters, 45 meters, 50 meters, 55 meters and 58 meters, respectively. The superstructure is 32 meters equal span beam, and weight of each span is 279.3ton as well as secondary dead load is 7.5×32.6＝244.5t. The bilinear model is adopted in LRB. In the single pier model, the superstructure of the bridge is simplified as a lump mass and the pier is firmly adjoined which can be seen from Fig. 1. Fig. 2 shows the mechanical model of lead rubber bearing. Seven typical seismic waves at Site II are taken as the seismic excitation. The basic characteristics of the seismic waves are showed in table 1.

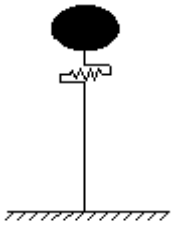


Figure 1. Model of single pier with rigid foundation.

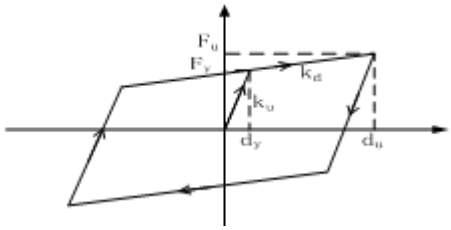


Figure 2. Mechanical model of LRB.

Table 1. Basic characteristics of the earthquake waves under the second ground

Name	Component	Magnitude	Acceleration [cm/s2]	Time
Kern County	S48E	7.7	128.6	1952.7.21
San Fernando	N45E	5.3	45.9	1957.3.22
Park field Cali for.	N50E	5.6	232.6	1966.6.27
San Fernando	N29E	6.4	96.7	1971.2.9
San Fernando	S89W	6.4	154.1	1971.2.9
San Fernando	N37E	6.4	195.6	1971.2.9
Southern Cali for.	N36W	6	52.9	1952.11.21

Significance analysis of parameters of LRB

The dynamic time-history analyses of bridges without isolation are carried out first. Then the seismic force under the 7 seismic waves can be obtained. Last the mean value can be considered as the seismic force of the pier. Table 2 shows the mean seismic force of the bridges without isolation.

Table 2. Average seismic force at the bottom of pier without isolation.

Height [m]	30	35	38	42	45	50	55	58
Moment [N·m]	7.22E+07	7.37E+07	8.83E+07	1.12E+08	1.19E+08	1.38E+08	1.77E+08	1.86E+08

Hardening ratio comparisons and significance analysis. In this calculation, the initial shear stiffness and the yield strength are fixed, while the hardening ratio is divided into five levels which are 0.08, 0.2, 0.3, 0.4 and 0.5, respectively. The average moment at the bottom of pier is calculated by using the time-history analysis method. Fig. 3 shows the relationship between hardening ratio and the moment at the bottom of pier. Fig. 4 shows the relationship between hardening ratio and isolation ratio of the moment at the bottom of pier.

According to Fig. 3, the seismic response of an isolated bridge pier increases correspondingly with its hardening ratio. What's more, the increasing trend is very stable. When the hardening ratio $\eta>0.2$, the increasing rate become larger. Therefore, we can increase the hardening ratio of low piers as well as reduce the hardening ratio of high piers in bridge design in order to the seismic responses of bridge piers with different height almost the same. Fig. 4 shows that the isolation ratio of the moment at the bottom of pier is decreasing while the hardening ratio is increasing, and this tendency decreases linearly after $\eta=0.3$. According to the model analysis, hardening ratio has no influence on structure period. The shock absorption is realized through the change of the seismic responses which is caused by the hysteresis of bearing. The smaller the hardening ratio, the bigger and fuller the area of the

hysteresis is, and the more energy consumed, as a result, the better the isolation effect is. The 42 meter high bridge pier is taken as an example, which is showed in Fig. 5.

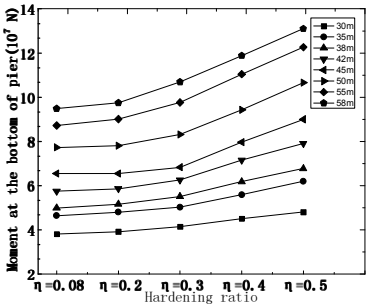


Figure 3. Relationship between hardening ratio and seismic responses.

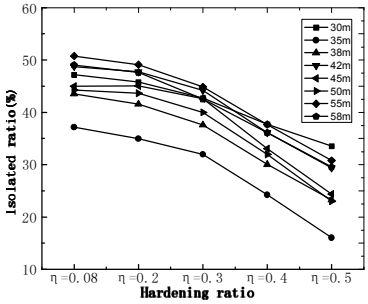


Figure 4. Relationship between hardening ratio and the isolation ratio of the moment at the bottom of pier.

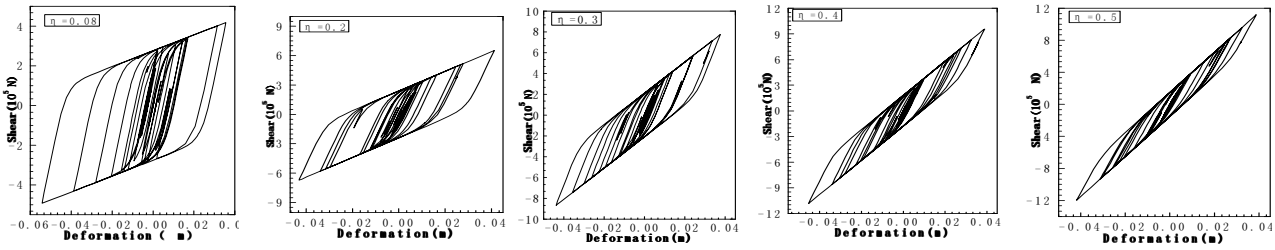


Figure 5. Hysteretic curves of LRB isolating single pier in five different levels of hardening ratio η.

Before discussing the significant effect of hardening ratio on the seismic responses of the isolated bridge pier by LRB, we should check the samples which come from the time-history analysis assistance whether they obey the normal distribution or not. In this paper, Kolmogorov-Smirnov test (Dn) is applied, and the calculation results are presented in table 3.

Table 3. Kolmogorov-Smirnov test of hardening ratio.

Hardening ratio	Sample size	Normal parameter		Extreme difference			Kolmogorov-Smirnov Z	Asymptotic significant
		Nverage	Standard deviation	Absolute value	Positive	Negative		
η=0.08	8	45.71	4.28	0.18	0.12	-0.18	0.51	0.96
η=0.2	8	44.43	4.53	0.18	0.15	-0.18	0.51	0.96
η=0.3	8	40.81	4.25	0.28	0.17	-0.28	0.80	0.55
η=0.4	8	33.37	4.62	0.22	0.17	-0.22	0.63	0.82
η=0.5	8	26.25	5.63	0.21	0.13	-0.21	0.60	0.87

According to table 3, the hardening ratio samples are in obedience to the normal distribution. By taking hardening ratio as a single factor for variance analysis, and using F-test method, we can get $F=24.38>3.91=F_{(1-0.01)}(4, 35)$. It proves that the isolation rate can be significantly affected by hardening ratio.

Comparisons and the significance analysis of the initial shear stiffness. In this calculation, hardening ratio and the yield strength are fixed. Then, let the initial shear stiffness K has five levels value: K=2.0E+07N/m, K=4.0E+07N/m, K=6.0E+07N/m, K=8.0E+07N/m and K=1.0E+08N/m, respectively. The average moment at the bottom of the pier is calculated by using the time-history analysis method. Fig. 6 shows the relationship between initial shear stiffness and moment at the bottom of fixed pier. Fig. 7 shows the relationship between initial shear stiffness and isolation ratio of the bottom moment of pier.

Fig. 6 shows the seismic response of the isolated bridge pier increases basically corresponding to the initial shear stiffness K. Fig. 7 shows that the isolation ratio of moment bottom of pier is decreasing while the initial shear stiffness K increasing. When K<4.0E+07N/m, some piers, such as 35m, 38m, 42m, present the reverse trend; some piers, such as 35m, 38m, 55m and 58m, present abnormal results. Taking initial shear stiffness as a single factor analysis of variance, and using F-test

method to get F=19.63>3.91=F $_{(1-0.01)}$ (4, 35). We can get a conclusion that the seismic responses can be significantly affected by initial shear stiffness.

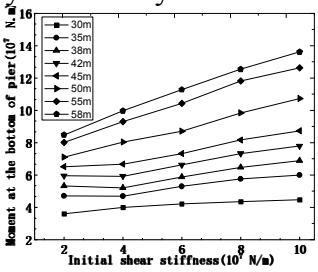


Figure 6. Relationship between initial shear stiffness and seismic responses.

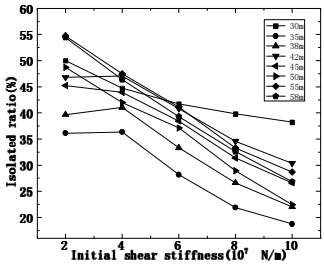


Figure 7. Relationship between initial shear stiffness and the isolation ratio of the moment at the bottom of pier.

The comparison and the significance analysis of the yield strength. In the calculation, the hardening ratio η and the initial shear stiffness K are fixed. Then the yield strength F is divided into five levels which are F=1.0E+05, F=2.0E+05, F=3.0E+05, F=4.0E+05 and F=5.0E+05, respectively. The average bottom moment of the pier is calculated by using the time-history analysis method. Fig. 8 shows the relationship between the initial shear stiffness and the moment at the bottom of fixed pier. Fig. 9 shows the relationship between the initial shear stiffness and the isolation ratio of the moment at the bottom of fixed pier.

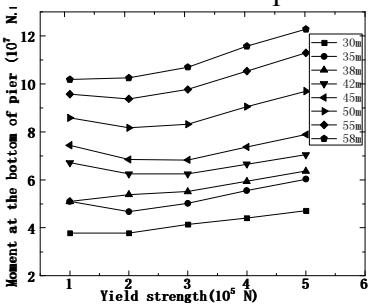


Figure 8. Relationship between yield strength and seismic

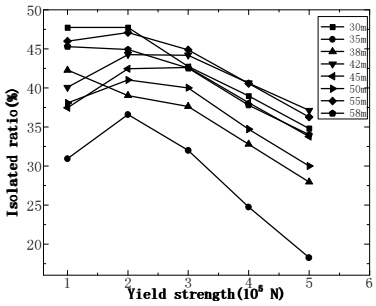


Figure 9. Relationship between yield strength and the isolation ratio of the moment at the bottom of pier.

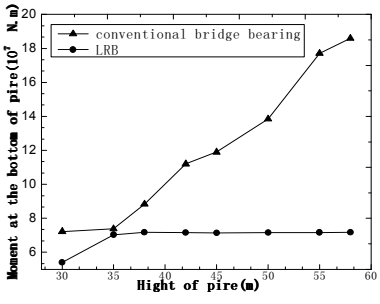


Figure 10. Comparison of the seismic responses of an isolated single pier with and without isolation.

Fig .8 shows that seismic responses increase with yield strength when F>2.0E+05N; the seismic responses of all bridge piers, except those with height 35m, 58m, decrease with the yield strength when F<2.0E+05N. Fig. 9 shows that the isolation ratio of the moment at the bottom of pier is decreasing when F>2.0E+05N. It is known that the yield strength has no influence on structure period. However, the shock absorption is realized through the change of the seismic responses which is caused by the hysteresis of bearing. The smaller the hardening ratio is, the bigger and fuller the area of the hysteresis is, , and the more the energy is consumed. By taking the initial shear stiffness as a single factor for variance analysis, and using F-test method, we can get F=7.39>3.91=F $_{(1-0.01)}$ (4, 35). It is concluded that seismic responses can be significantly affected by initial shear stiffness.

The adjustment of the parametersof LRB for seismic force.

Based on the theory and conclusions above, optimum parameters for LRB are designed to test the correctness of those conclusions and the adjustment ability of LRB parameters for seismic force. The LRB parameters of eight isolated bridge piers are presented in table 4. The average seismic force of different bridge piers isolated by LRB is presented in table 5. From Fig.10, it can be concluded: (1) some seismic responses of piers isolated by LRB with different heights are almost the same through the design of the LRB optimum parameters. As a result, it can concluded that seismic force of an isolated bridge pier can be regulated by LRB parameters; (2) the average seismic force of 30m bridge pier is much less than other bridge piers which have equal seismic force, in other word, the regulation

ability of LRB parameters is limited, when the height difference is over 30 meters, parameters of LRB cannot regulate the seismic force of different piers to the same.

Table 4. LRB parameters design for piers with different heights

Height[m] / Parameters	30	35	38	42	45	50	55	58
η	0.5	0.5	0.4	0.3	0.3	0.3	0.2	0.08
K[N/m]	1.00E+08	1.00E+08	7.00E+07	7.00E+07	6.00E+07	3.00E+07	3.00E+07	2.00E+07
F[N]	5.00E+05	5.00E+05	4.00E+05	4.00E+05	3.00E+05	2.00E+05	2.00E+05	1.00E+05

Table 5. The average seismic force of piers isolated by LRB with different height pier.

Height [m]	30	35	38	42	45	50	55	58	
Moment [N·m]	5.40E+07	7.03E+07	7.18E+07	7.15E+07	7.14E+07	7.16E+07	7.16E+07	7.17E+07	

Conclusion.

- The parameters of LRB have significant influences on the isolation ratio of the moment at the bottom of a pier. In addition, the seismic force of an isolated pier can be regulated, even equaled when the height differences of piers are below 30 meters.
- Seismic response of a pier is increasing with hardening ratio's increasing. It is noted that the seismic response increases more dramatically after the hardening ratio η exceed 0.2; the isolation ratio of the moment at the bottom of pier is decreasing with the hardening ratio increasing.
- Seismic response of an isolated bridge pier increases with initial shear stiffness K, while the isolation ratio of the moment at the bottom of pier is decreases with initial shear stiffness K. When K<4.0E+07N/m, some piers , such as 35m, 38m, 42m, have the opposite tendency.
- Seismic responses are increasing with yield strength when F>2.0E+05; the seismic responses of all bridge piers, except those with height 35m and 58m, are decreasing while F<2.0E+05; the isolation ratio of the moment at the bottom of pier is decreasing when F>2.0E+05N.

Acknowledgement

This study is sponsored by the Science & Technology Research Development Project of China Railway Ministry (Grant No. 2008G028-C).

References

[1] R.S. Jangid, Optimum lead–rubber isolation bearings for near-fault motions, Engineering Structures 29, pp: 2503–2513, 2003.

[2] J. S. Hwang, L. H. Sheng, Equivalent elastic seismic analysis of base-isolated bridges with lead-rubber bearings, Engineering Structures, Volume 16, Issue 3, April 1994, pp: 201-209, 1994.

[3] B. Wu, J.S. Zhuang, Study on Mechanical Characteristics of Lead-rubber Bearing and Its Application for Shock Reduction/Isolation in Bridge Engineering, China Railway Science, 25(4)pp: 138-140, 2004.

[4] D.S. Zhu, Y.C. Lao, D.Y. Shen, Q. Li, Research on Some Design Parameters of Seismically Isolated Bridges, China Civil Engineering Journal, Vol.33 (5)pp:52-56, 2002.

[5] T.Y. Zhong, F.L. Y, B. Wu, Analysis of the Bidirectional Seismic Responses for Seismically Isolated Railway Simple Supported Beam Bridge by Lead Rubber Bearing, China Railway Science, 28(3) pp: 38-43, 2007.

[6] Z.G Zhang, T.Y. Zhong, L. Wang, Study on Vibration Absorption and Isolation of Continuous Beam Railway Bridge by Lead Cored Rubber Buffer, China Safety Science Journal, 13(5)pp:41-45, , 2003.

[7] F.L.Yang, Study on Optimization of the Design Parameter of Seismic Absorption and isolation bearing for Railway Simple Supported Beam Bridges, Journal of the China Railway Society,Vol.28(3), pp:128-132, 2007.

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.110

Robust Control for a Class of Nonlinear Systems Based on Backstepping

Naibao He, Qian Gao

[1.] Huaihai Institute of Techology, Lianyungang 222005, China

E-mail: henaibao@126.com

Key Words: nonlinear system; adaptive control; backstepping

Abstract: Based on coordinate transform, the paper deduced the principle with which Chua's chaotic system can be translated into the so-called general strict-feedback form. Combining the backstepping method with robust control technology, an adaptive parameter control law is developed and thus the output tracking is successfully accomplished for the system with unknown parameters and dynamic uncertainties. It is proved that all states of the closed-loop system are globally uniformly ultimately bounded, and lead the system tracking error to a small neighborhood. Finally simulation results are provided to show the effectiveness of the proposed approach.

Introduction

Great progress in adaptive control of nonlinear systems, which plays an important role due to its ability to compensate for parametric uncertainties, has been witnessed in recent years [1-4]. In the past decade, backstepping has become one of the most popular design methods for adaptive nonlinear control and synchronization [5-6], because it can guarantee global stabilities, tracking and transient performance for the broad class of strict-feedback systems [7-8]. In this paper, we firstly show that the Chua's system in the research of chaos can be transformed into a kind of nonlinear system in the so-called general strict feedback form. Secondly, we focus on realizing chaos synchronization of a class of uncertain system by using only one controller and estimate parameters update law. The method adopted here is adaptive backstepping design procedure. Finally, numerical simulations are provided to show the effectiveness of the proposed approach.

Problem statement

The nonlinear system is described by

$$\begin{cases} \dot{z}_1 = p_1(z_2 - z_1 - f(z_1)) \\ \dot{z}_2 = z_1 - z_2 + z_3 \\ \dot{z}_3 = -p_2 z_2 \end{cases} \tag{1}$$

Let $g_1 = -p_2, g_2 = 1, g_3 = 1,\ \theta = 0, \Delta_1 = 0, \Delta_2 = x_1 - x_2$

$$\Delta_3 = p_1 x_2 - p_1(1 + p_4 x_3) - \frac{1}{2} p_1(p_3 - p_4)(|x_3 + 1| - |x_3 - 1|)$$

In this way, the controller Chua's circuit (1) can be easily written into a n-order feedback form as follows.

$$\begin{cases} \dot{x}_i = g_i(\bar{x}_i, t)x_{i+1} + \theta^T \phi_i(\bar{x}_i, t) + \Delta_i(\bar{x}_i, t) \\ \qquad 1 \le i \le n-1 \\ \dot{x}_n = g_n(\bar{x}_n, t)u + \theta^T \phi_n(\bar{x}_n, t) + \Delta_n(\bar{x}_n, t) \\ y = x_1 \end{cases} \tag{2}$$

where, $x = [x_1, \cdots, x_n]^T \in R^n$, $u \in R$, and $y \in R$ are the states, input and output, respectively; $\bar{x}_i$ denotes $\bar{x}_i = [x_1, \cdots, x_i]^T \in R^i \quad i = 1, \cdots, n$; $\theta = [\theta_1, \theta_2, \cdots, \theta_l]^T \in R^l$ is the vector of unknown constant parameters of interest; $\Delta_i(\bar{x}_i, t)$ is unknown Lipschitz continuous dynamic uncertain functions; $g_i(\cdot) \neq 0$, $\phi_i(\cdot)$, $i = 1, \cdots, n$ are known smooth nonlinear functions.

In this paper, we present a robustification method for the usual adaptive backstepping design for system (2) in the presence of parametric and dynamic uncertainties.

Assumption1: For each $1 \le i \le n$, there exists an unknown positive constant $\overline{p}_i$ such that, for all $\Delta_i(\overline{x}_i, t)$, satisfies $|\Delta_i(\overline{x}_i, t)| \le \overline{p}_i \psi_i(x_1, \cdots, x_i)$, where $\psi_i(x_1, \cdots, x_i)$ is a known nonnegative smooth function.

Assumption2: Let y_r be a bounded reference signal whose n-th derivatives is also bounded, and for random known positive constant d, satisfies $|y_r^{(n)}(t)| \le d$

Lemma 1[8]: For any x and y in R^n, and for any positive real number δ, we have

$$x^T y \le \frac{1}{4\delta}|x|^2 + \delta |y|^2$$

.

Controller design

In this section we extend the adaptive backstepping method to the non-automous strict-feedback system in form (2). The backstepping design procedure contains n steps. At step i, an intermediate control function α_i shall be developed by using an appropriate Lyapunov function V_i. Let us first consider the equation when $i = 1$.

Step1. Define the first error variable

$$z_1 = x_1 - y_r \tag{3}$$

Its derivative is $\dot{z}_1 = \dot{x}_1 - \dot{y}_r$

We choose the following Lyapunov function candidate

$$V_1 = \frac{1}{2} z_1^2 + \frac{1}{2} \tilde{\theta}_1^T \Gamma^{-1} \tilde{\theta}_1 + \frac{1}{2} r^{-1} \tilde{p}_1^2 \tag{4}$$

Where $\Gamma = \Gamma^T > 0$ is the adaptive gain matrix , $\tilde{\theta}_i = \theta_i - \hat{\theta}_i$, $\tilde{p}_i = p_i - \hat{p}_i$, $\hat{\theta}_i$ and $\hat{p}_i$ represent the estimate vector of unknown parameter vector θ_i and p_i, respectively ,r is a design parameter.

The time derivative of the Lyapnuov function V_1 satisfies

$$\dot{V}_1 == z_1(g_1 x_2 + \theta_1^T \phi_1(x_1) + \Delta_1 - \dot{y}_r) - \tilde{\theta}_1^T \Gamma^{-1} \dot{\hat{\theta}}_1 - r^{-1} \tilde{p}_1 \dot{\hat{p}}_1$$

According to assumption 1, assumption 2 and lemma 1, we have

$$|\Delta_1| \le \overline{p}_1 \psi_1(x_1) \quad , \quad |z_1 \dot{y}_r| \le \delta + \frac{z_1^2}{4\delta} d^2$$

$$\dot{V}_1 \le z_1(g_1 x_2 + \hat{\theta}_1^T \phi_1(x_1)) + 2\delta + \frac{z_1}{4\delta}(d^2 + \hat{p}_1 \psi_1^2(x_1)) - \Gamma^{-1} \tilde{\theta}_1^T (\dot{\hat{\theta}}_1 - \Gamma z_1 \phi_1(x_1)) - r^{-1} \tilde{p}_1 (\dot{\hat{p}}_1 - r \frac{z_1^2}{4\delta} \psi_1^2(x_1)) \tag{5}$$

Step i .($2 \le i \le n-1$) After some simple algebraic manipulations, the derivative of $z_i = x_{i+1} - \alpha_i$ can be expressed as $\dot{z}_i = \dot{x}_{i+1} - \dot{\alpha}_i$,where α_i is a virtual control. We consider the following Lyapunov function

$$V_i = V_{i-1} + \frac{1}{2} z_i^2 + \frac{1}{2} \tilde{\theta}_i^T \Gamma^{-1} \tilde{\theta}_i + \frac{1}{2} r^{-1} \tilde{p}_i^2$$

then we can obtain the derivative of V_i

$$\dot{V}_i = \dot{V}_{i-1} + z_i \dot{z}_i - \tilde{\theta}_i^T \Gamma^{-1} \dot{\hat{\theta}}_i - r^{-1} \tilde{p}_i \dot{\hat{p}}_i \tag{6}$$

$$\le -\sum_{k=1}^{i-1} c_k z_k^2 + \sum_{k=1}^{i-1} g_k z_k z_{k+1} + z_i(g_i x_{i+1} + \hat{\theta}_i^T \phi_i(x_1, \cdots; x_i) + \frac{z_i}{4\delta} \hat{p}_i \psi_i^2(x_1, \cdots; x_i) - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} g_j x_{j+1} - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} \theta_j^T \phi_j(x_1, \cdots; x_j)$$

$$+ \frac{z_i}{4\delta} d^2 (\frac{\partial \alpha_{i-1}}{\partial y_r}) - \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}_i} \dot{\hat{\theta}}_i - \frac{\partial \alpha_{i-1}}{\partial \hat{p}_i} \dot{\hat{p}}_i) - \tilde{\theta}_i^T \Gamma^{-1} \dot{\hat{\theta}}_i - r^{-1} \tilde{p}_i \dot{\hat{p}}_i + z_i \tilde{\theta}_i^T \phi_i(x_1, \cdots; x_i) + \frac{z_i^2}{4\delta_i} \tilde{p}_i \psi_i^2(x_1, \cdots; x_i) + 2i\delta \tag{7}$$

Choose virtual control law α_i as

$$\alpha_i=\frac{1}{g_i}(-c_iz_i+\sum_{j=1}^{i-1}\frac{\partial\alpha_{i-1}}{\partial x_j}g_jx_{j+1}+\sum_{j=1}^{i-1}\frac{\partial\alpha_{i-1}}{\partial x_j}\theta_j^T\phi_j(x_j))-\hat{\theta}_i^T\phi_i(x_i)-\frac{z_i}{4\delta}\hat{p}_i\psi_i^2(x_1,\cdots,x_i)-\frac{z_i}{4\delta}d^2(\frac{\partial\alpha_{i-1}}{\partial y_r})^2+\frac{\partial\alpha_{i-1}}{\partial\hat{\theta}_i}\dot{\hat{\theta}}_i+\frac{\partial\alpha_{i-1}}{\partial\hat{p}_i}\dot{\hat{p}}_i) \quad (8)$$

Substituting (12) into(11) results in

$$\dot{V}_i\le-\sum_{k=1}^{i-1}c_kz_k^2+\sum_{k=1}^{i-1}g_kz_kz_{k+1}+\tilde{\theta}_i^T(z_i\phi_i(x_1,\cdots,x_i)-\Gamma^{-1}\dot{\hat{\theta}}_i)+r^{-1}\tilde{p}_i(\frac{z_i^2}{4\delta}r\psi_i^2(x_1,\cdots,x_i)-\dot{\hat{p}}_i)+2i\delta$$

The $(\theta_i-\hat{\theta}_i)$and $(p_i-\hat{p}_i)$ term can be eliminated with the update law

$$\dot{\hat{\theta}}_i=\Gamma z_i\phi_i(x_1,\cdots,x_i)\qquad\dot{\hat{p}}_i=\frac{z_i^2}{4\delta}r\psi_i^2(x_1,\cdots,x_i) \quad (9)$$

which yields $\dot{V}_i\le-\sum_{k=1}^{i-1}c_kz_k^2+\sum_{k=1}^{i-1}g_kz_kz_{k+1}+2i\delta$ (10)

The coupling term $\sum_{k=1}^{i-1}g_kz_kz_{k+1}$ can be eliminated in the final step.

Step n. Define Lyapunov function candidate as

$$V_n=V_{n-1}+\frac{1}{2}z_n^2+\frac{1}{2}\tilde{\theta}_n^T\Gamma^{-1}\tilde{\theta}_n+\frac{1}{2}r^{-1}\tilde{p}_n^2 \quad (11)$$

Its derivative is

$$\dot{V}_n=\dot{V}_{n-1}+z_n\dot{z}_n-\tilde{\theta}_n^T\Gamma^{-1}\dot{\hat{\theta}}_n-r^{-1}\tilde{p}_n\dot{\hat{p}}_n$$

The choice of control u is given by

$$u=\alpha_n=\frac{1}{g_n}(-c_nz_n+\sum_{j=1}^{n-1}\frac{\partial\alpha_{i-1}}{\partial x_j}g_jx_{j+1}+\sum_{j=1}^{n-1}\frac{\partial\alpha_{i-1}}{\partial x_j}\theta_j^T\phi_j(x_1,\cdots;x_j))-\hat{\theta}_n^T\phi_n(x_1,\cdots;x_n)-\frac{z_n}{4\delta}\hat{p}_n\psi_n^2(x_1,\cdots;x_n)-\frac{z_n}{4\delta}d^2(\frac{\partial\alpha_{n-1}}{\partial y_r})+\frac{\partial\alpha_{n-1}}{\partial\hat{\theta}_n}\dot{\hat{\theta}}_n+\frac{\partial\alpha_{n-1}}{\partial\hat{p}_n}\dot{\hat{p}}_n) \quad (12)$$

Selecting the update laws for $\dot{\hat{\theta}}_n$ and $\dot{\hat{p}}_n$,

$$\dot{\hat{\theta}}_n=\Gamma z_n\phi_n(x_1,\cdots,x_n),\ \dot{\hat{p}}_n=\frac{z_n^2}{4\delta}r\psi_n^2(x_1,\cdots,x_n) \quad (13)$$

then $\dot{V}_n\le-\sum_{k=1}^{n}c_kz_k^2+2n\delta$ (14)

The c_n selected is the smallest positive candidate constant, μ_n is a positive candidate bounding number, then we have

$$\dot{V}_n\le-c_nV_n+\mu_n \quad (15)$$

let $\lambda=\mu_n/c_n$, then $0\le V_n(t)\le\lambda+[V_n(0)-\lambda]e^{-c_nt}$

By means of the Barbalat lemma, we have

$$\lim_{t\to\infty}V_n=0 \quad (16)$$

So far the entire design procedure is completed.

Theorem1: Under assumption 1, assumption 2 and lemma 1,there exists a state feedback controller u and adaptive control law $\dot{\hat{\theta}}_i=\Gamma z_i\phi_i(x_1,\cdots,x_i)\quad\dot{\hat{p}}_i=\frac{z_i^2}{4\delta}r\psi_i^2(x_1,\cdots,x_i)$ acting on system. The closed-loop system is globally stable for all admissible uncertainties, and its output tracking error satisfies

$$|e|=|y-y_r|\le\sqrt{2V_n(0)}e^{-\frac{c_nt}{2}}+\sqrt{2\mu_n/c_n}$$

Proof: Let $V_n = V_{n-1} + \frac{1}{2}z_n^2 + \frac{1}{2}\tilde{\theta}_n^T \Gamma^{-1}\tilde{\theta}_n + \frac{1}{2}r^{-1}\tilde{p}_n^2$ Equation(16) yields $\dot{V}_n e^{c_n t} + c_n V_n e^{c_n t} \le \mu_n e^{c_n t}$

both sides quadrature and multiply $e^{-c_n t}$

$$V_n(t) \le [V_n(0) - \mu_n / c_n]e^{-c_n t} + \mu_n / c_n \le V_n(0)e^{-c_n t} + \mu_n / c_n$$

we have $|e|^2 \le 2V_n(0)e^{-c_n t} + 2\mu_n / c_n \le (\sqrt{2V_n(0)}e^{-\frac{c_n t}{2}} + \sqrt{2\mu_n / c_n})^2$

$$|e| = |y - y_d| \le \sqrt{2V(0)}e^{-\frac{c_n t}{2}} + \sqrt{2\mu_n / c_n}$$

Simulation example

To examine the effectiveness of the proposed design procedure, extensive computer simulations were carried out for uncertainties nonlinear system.

$$\begin{cases} \dot{x}_1 = x_2 + \theta_1^T x_1^2 + \Delta_1(t) \\ \dot{x}_2 = x_1^2 + 0.1(1 + x_2^2)u + \sin(0.5u) \\ y = x_1 \end{cases}$$

The design controller is asymptotically to track the reference output signal

$$y_d = 0.3\sin\left(t + \frac{\pi}{2}\right) + 0.5\sin(0.6t)$$

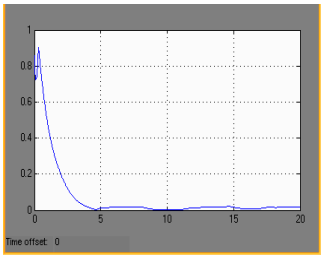


Fig 1. The control law u

For simulation purposes, we let $\theta(0) = 0.1$ and $\Delta_1(t) = 0.6\sin(2t)$, the initial conditions are chosen as $r = 1$, $c_1 = c_2 = 1$, $\delta = 0.1$ $x(0) = [0.2, 0]^T$, $\Gamma = 1$.

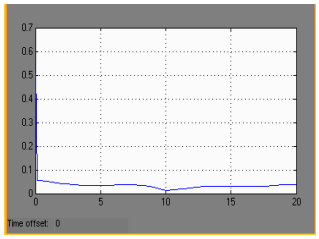


Fig2. Output tracking error $y - y_r$

In view of simulation on Fig 1and 2,it implies that the closed-loop system satisfies global uniformly ultimate boundness for all admissible uncertainties, estimate parameter convergent to certain constant. Simulation result demonstrate, the robust adaptive controller designed is feasible and effective.

Conclution

We have presented an adaptive control design procedure for the regulation of nonlinear systems with both parametric uncertainty and unknown nonlinearities. In the design, backstepping method is used to get the parameter adaptive control law. The proposed systematic backstepping design method has been proved to be able to guarantee global uniform ultimate boundedness of closed-loop signals. In addition, the output of the system has been proven to converge to an arbitrarily small neighborhood of the origin. Simulation results have been provided to show the effectiveness of the proposed approach.

References

[1] Yu YG, Zhang SC. Adaptive backstepping control of the uncertain Lū system,.Chin Phys ,Vol.11, No.12, (2002). P.1249-1253.

[2] Y.-S. Zhong, Globally stable adaptive system design for minimum phase SISO plants with input saturation, Automatica, Vol.41, No.9, (2005). p.1539-1547.

[3] C.L. Yang, R.Y. Ruan and M.K. Gong, Robust adaptive state feedback controller design and analysis of a class of uncertain nonlinear systems,Acta Mathematica Scientia, Vol.24, No.1, (2004), p.26-37.

[4] Jiang Z P,Parly L, Design of robust adaptive controller for nonlinear systems with uncertainties, Automatica, Vol.34, No.7, (1998). P.825-840.

[5] Krishnamurthy P, Khorrami F, Jiang Z P,Global output feedback tracking for nonlinear systems in generalized out-feedback canonical, IEEE Trans. On Automatic Control, Vol.47, No.5, (2002).p.814-819.

[6] Hai-Bo Ji and Hong-Sheng Xi, Adaptive output-feedback tracking of stochastic nonlinear systems, IEEE Trans. On Automatic Control, Vol.51, No.2,(2006),p.355-360.

[7] J. Zhou, C. Wen, Y. Zhang, Adaptive output control of a class of timevarying uncertain nonlinear systems, Nonlinear Dynamics System Theory, No.5, (2005),p.285-298.

[8] Krishnamurthy P, Khorrami F, Jiang Z P,Global output feedback tracking for nonlinear systems in generalized out-feedback canonical ,IEEE Trans. On Automatic Control, Vol.47, No.5, (2002). P.814-819.

Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.115

Numerical Simulation of Mechanical Behavior on Landslides Triggered by Heavy Rainfall

Zhenqiang Ni[1,2,3,a] , Jiming Kong [1,3,b], Fayou A [1,2,3,c]

[1]Institute of Mountain Hazards and Environment, Chinese Academy of Science & Ministry of Water Conservancy, Chengdu 610041, China;

[2] Key Laboratory of Mountain Hazards and Surface Process, Chinese Academy of Science, Chengdu 610041, China;

[3]Graduate University of Chinese Academy of Sciences, Beijing 100101

[a]nizhen1230@sina.com, [b]jimingk@imde.ac.cn, [c]afayou@163.com

Keywords: rainfall, landslide, deformation instability, mechanical behavior, numerical simulation

Abstract. Deformation instability of slope due to rain was one of the most common geological disasters. Speaking from the mechanism, the destructive action and triggering action of rainfall affected the stability of slope itself, which made the slope deform and even lead to instability. June 28, 2010, Guanling in Guizhou province a large landslide occurred which triggered by continuous heavy rainfall, it caused that in two village-groups 107 villagers were buried, all the tragedy of houses were completely destroyed. Using finite element software ANSYS to simulate the mechanical behavior of the landslide, to analyze the deformation mechanism, we got that the large landslide deformation mechanism was a complex mechanical geological process. From the analysis of interaction between soil and rock, the landslide showed complex mechanical geological features. As rainfall increased the strength of unsaturated soil and weak intercalation reduced, this strengthened the role of the slope deformation and failure. The lock-fixed section first appeared brittle fracture, then that brought shear failure to the creeping section. The interaction between lock-fixed section and creeping section was the key of slope failure.

Introduction

Deformation instability of slope due to rain was one of the most common geological disasters. Speaking from the mechanism, the rainfall destructive action and triggering action affected the stability of slope itself, which made the slope deform and even lead to instability. The impact of rainfall acted on the slope mainly through its external kinetic energy and internally generated interaction after infiltration, both of two led to the rainfall-induced-landslides [1]. The loss caused by rainfall-induced-landslides was enormous in China, according to statistics in the years 1310 to 1987, the number of deaths grew up to 256,000 people, and the annual economic loss reached 10 to 30 billion [2].

Influencing factors on slope stability

Engineering geological conditions. The lithology of Guanling landslide area mainly included clastic rocks, carbonates and quaternary broken sediments, from old to new as follows:

P_{2l+c}, the lithology comprised mudstone, siltstone and coal seam. Mainly distributing in the landslide accumulation area, surface was weathered seriously.

T_{ly}, the lithology comprised siltstone, mudstone. It was the layer where slide plane being, surface was weathered seriously, with joint fissures.

Fund: National Natural Science Foundation Project (50639070)
Author introduction: Zhenqiang Ni, born in Shandong in 1983, a doctoral candidate majoring geotechnical engineering.
Corresponding author: Jiming Kong is a researcher mainly conducting the studies on mountain disaster formation mechanism theory and mitigation project, born in Chongqing in 1956. E-mail: jimingk@imde.ac.cn

T_{lyn}, the lithology comprised limestone, dolomite. It mainly distributed in the south landslide area.

Q_{4el+dl}, lithology comprised silty clay with gravel. Widely distributing in all the landslide area, they mainly came from weathered clastic rocks, thickness ranged from 1 to 20m.

The landslide located on the southwestern side of Jianzhuang village fracture, 1.1km from it. The fracture strike was northwestwards, it extended length about 2.5km. The whole regional stratigraphic inclined to south, it was a monoclinal structure. The landslide aspect was 305-329°, stratigraphic occurrence was 170°∠45°, the slope was an escarpment slope. Unloading fracture developed, occurrence was 310°∠70°.

Rainfall. From June 27, 2009 to March 18, 2010, local rainfall had never been, however by 2010 on June 26 to 28, local area suffered continuous heavy rainfall. Especially in the night on 27, maximum rainfall reached 230mm, slope weight and physical mechanics parameter of rock and soil were liable to change rapidly by the situation that long drought and long period of rainfall. Rain would be very easy to quickly penetrate into the mudstone and sandstone in lower part mountain, and it made the soil saturated, bulk density increased, strength reduced. So that sliding force of the slope increased, the anti-sliding force reduced, safety factor was lower. Besides, increasing fissure water pressure and uplift force accelerated the further development of the landslide [3].

Foreign scholars [4] suggested using two independent stress state variables $(\sigma_n - u_a)$ and $(u_a - u_w)$ to express the state of unsaturated soil, and the corresponding shear strength formula was:

$$\tau_f = c' + (\sigma_n - u_a)\tan\varphi' + (u_a - u_w)\tan\varphi^b \quad (1)$$

σ_n—The total normal stress that affected on the bottom slope; c'—The effective cohesion; φ'—The internal friction angle that related to the normal force; φ^b—The internal friction angle that changed with suction; $(u_a - u_w)$—The matrix suction. With equation (1) we can analyze soil shearing strength with changes in rainfall infiltration process. The total cohesion $c' + (\sigma_n - u_a)\tan\varphi'$ was proportional to matrix suction $(u_a - u_w)$. Usually, slope in unsaturated soil, the matrix suction was large. When rain happened, because the rain water infiltration to the soil was near saturation or saturation, pore water pressure u_w was close to pore air pressure u_a, the suction $(u_a - u_w)$ tended to 0, that matrix suction tended to 0, soil shear strength declined. This theory explained the decrease of its shear strength due to rainfall in the rainy season, and how it induced slope deformation failure.

Mechanism of the landslide

Guanling landslide was a scare high-speed long distance landside debris flow disaster, and its formation was very complex. The occurrence of landslide throughout the process was mainly infected and controlled by slope structure, lithology, rainfall and other factors.

In the long-term continuing action and driving under gravity, the internal friction angle of the weak plane was reduced by weathering and invasion of heavy rainfall, which made part of the rock over the geological weak plane fall along this plane. At this moment the middle of the intact rock as its friction and shear resistance ability formed the locked-fixed section; furthermore, since the structure of the escarpment slope, therefore the rock of front creep section played a role in piles, the *piles* prevented landslide happened to some extent; With the further expansion of crack, lock-fixed section stress reached limit state, and the rock appeared plastic deformation, then damage occurred. This made slope happen peak stress and shear the *piles*, then it released the potential energy suddenly, the high-speed landslide formed at last [5, 6].

Numerical simulation of slope instability mechanism

Calculation principle. As the rain continued immersing, the physical and mechanical parameters of soil layers would be dropped (bulk density increased), using strength reduction method, to decide the slope instability conditions: when the slope loss stability, the body slide, slip body went into motion

from the stable stationary state. In the same time it came with great displacement and plastic strain; and its displacement and plastic strain was no longer a constant value, but in the state of infinite plastic flow. This was the characteristic when slope was broken, that reflected in the calculation was calculation didn't converge [7, 8].

Finite element model. According to the geological profile of Guanling and information provided by geological departments, selecting reasonable calculation parameters (Table 1, Table 2), to establish the finite element model. We adopted plane strain finite element model to deal with this problem, using the contact element in ANSYS to simulate unloading fracture and solid element to simulate landslide body. There were 4,779 nodes and 4,575 units after finite elements were meshed, choosing a lock-fixed section point A as the control point (Figure 2, Figure 3).

The boundary conditions were used in calculation model: setting the upper part of model as free boundary; left and right as horizontal restraint; bottom boundary as immovably confined. As the regional tectonic stress had little effect on the slope, this idealized model only considered initial stress trigged by gravity in the calculation.

Table 1 Physical and mechanical properties of initial soils and rock

soil \ index	E[GPa]	ν	γ[KN·m^{-3}]	Shear strength	
				φ [°]	C[KPa]
Gravel soil	2.0	0.40	16.1	37.4	47
limestone	12.0	0.24	34	56	320
mudstone	8.0	0.25	26	42	240
Mudstone intercalated with coal	7.0	0.30	24	34	210
Fracture	0.4	0.35	18.2	31.5	29

Table 2 Physical and mechanical properties of different saturated soils

soil	index	E[GPa]	ν	γ[KN·m^{-3}]	Shear strength	
					φ [°]	C[KPa]
11.2%	Gravel soil	2.0	0.40	16.1	37.4	47
	Fracture	0.4	0.35	18.2	31.5	29
35%	Gravel soil	1.7	0.40	16.4	36.0	46.5
	Fracture	0.36	0.35	18.5	29.0	27.5
70%	Gravel soil	1.3	0.40	16.8	33.8	45.8
	Fracture	0.3	0.35	18.7	27.6	26.8
100%	Gravel soil	1.0	0.40	17	32	45
	Fracture	0.2	0.35	19	27	26

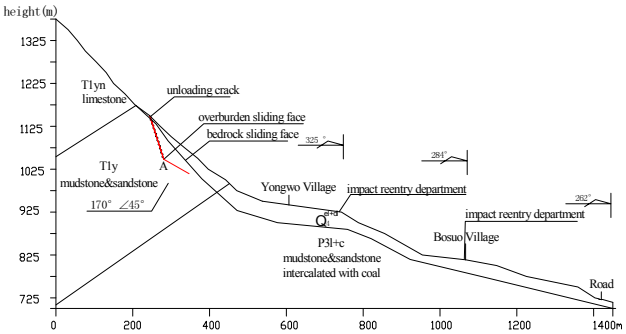


Fig.2 The engineering geological profile

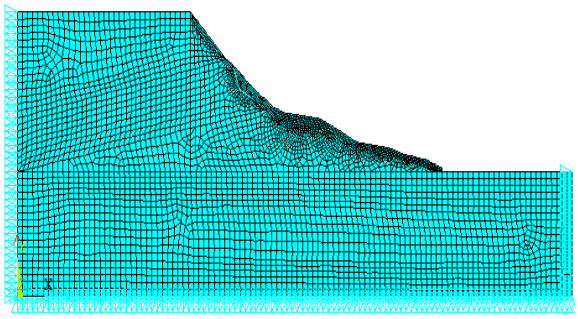


Fig.3 The FEM grid model of slope

The results and analysis. Calculation under different saturation had been conducted to solve, by defining the unit, to introduce the form criteria Mohr-Column, so that we can get the exact calculation of the maximum [9, 10].

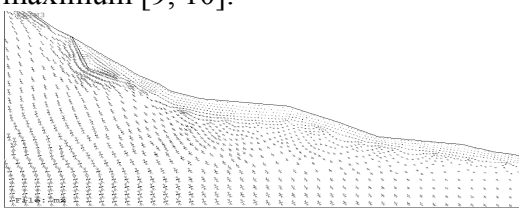


Fig.4 Initial principal stresses vectors

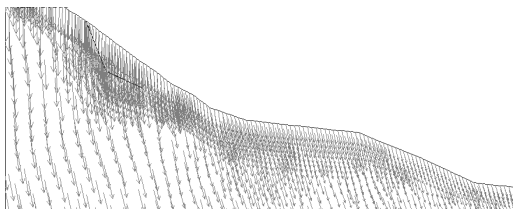


Fig.5 Initial displacement vectors

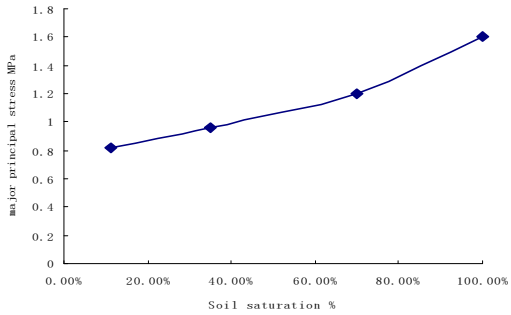


Fig.6 Rainfall and major principal stress curves of A

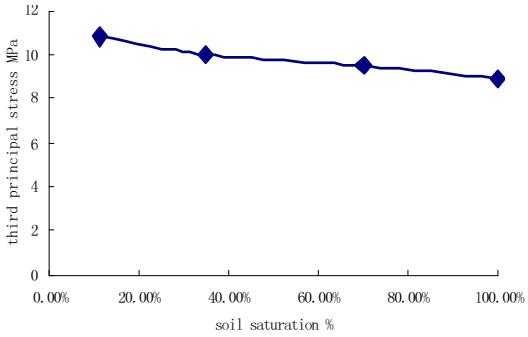


Fig.7 Rainfall and third principal stress curves of A

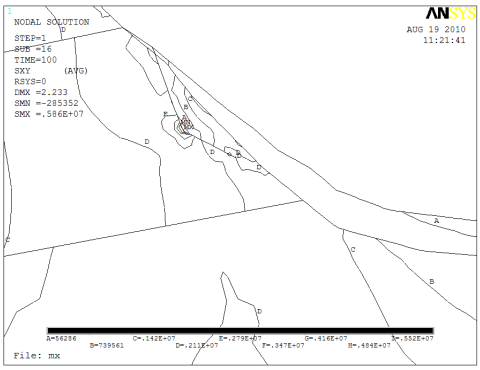


Fig.8 Shearing stress distribution after saturation /Pa

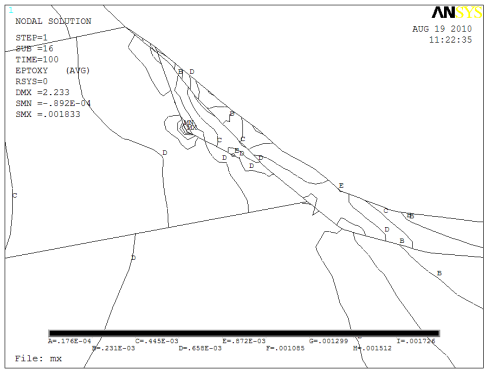


Fig.9 Shearing strain distribution after saturation

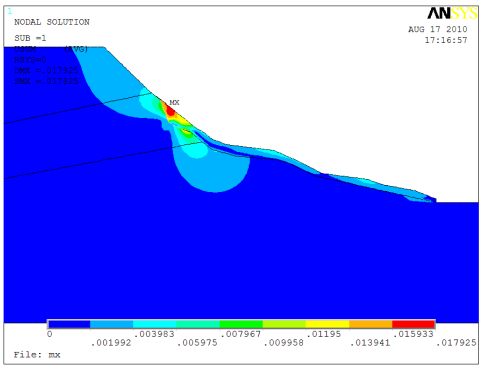


Fig.10 Displacement increment after saturation /m

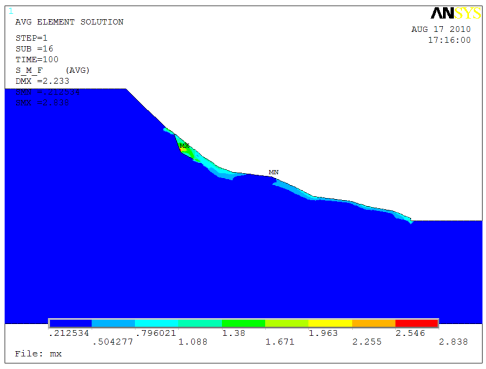


Fig.11 Distribution of plastic region after saturation /m

(1) Looking from figure 4, maximum principal stress in the slope body gradually changed from a vertical direction to be parallel with the slope. As the depth increased, the compressive stress value also increased, and the maximum shear stress in slope vivo was also increasing with depth. Maximum shear stress contours were parallel to the slope downward, and they didn't intersect with the slope. This stress distribution showed that because of loose soil on surface slope had no tensile strength, tensile stress concentration zone appeared only in the location of the posterior edge of the slope, where formed tensile cracks.

(2) Looking from figure 6 and figure 7, with the increasing rainfall, in the lock-fixed section the first principal stress increased, which increment was 0.78MPa, and the third principal stress decreased, which decrement was 1.9MPa. The increased accumulation period of the first principal stress exactly corresponded to the decreased period of the third principal stress, and that adjustment method was extremely unfavorable for the lock-fixed section.

(3) Looking from figure 5, because the broken soil in the back slope crept to the free face direction, slip body and slip zone would develop from back to front gradually, until they cut-through the slope body. Looking from figure 10, the displacement of upper slope increased about 1.8cm, 1.2cm displacement increments in the lower place, but only 1cm displacement increments in lock-fixed section. The upper displacement was blocked by lock-fixed section, it couldn't pass down freely, and so the lock-fixed section would gradually accumulate stress and strain. Looking from figure8,

figure9, the maximum shear stress in lock-fixed section reached 5.5MPa when the soil was saturated, maximum shear strain also reached 8.8e3. This made the lock-fixed section broken at last.

(4) Looking from figure 11, the largest distribution of plastic region in the locked section had been up to 2.8m, the damage of lock-fixed section would transfer the stress to the creeping section, and the *piles* were sheared at last.

Summary

We used finite element software ANSYS with the Mohr-Coulomb yield criterion to simulate the soil stress, strain and displacement, etc. The results showed that: slope in natural conditions was stable, but with the infiltration of rain, the soil saturation increased, the scope of slope failure gradually increased, and finally the slope developed from local failure to whole sliding failure.

Analysis told us that slope deformation and landslide were a complex geological mechanical process. With the rainfall increasing, the strength of unsaturated soil and weak interlayer decreased, that strengthened the role of slope deformation and failure. The lock-fixed section first happened brittle fracture, then that brought shear failure to the creeping section. The interaction between lock-fixed section and creeping section was the key of landslide failure.

Limited to the degree of current study, we didn't mention the decrease of sliding zone mechanical strength and the dynamic effect of landslide creep behavior under rainfall. Therefore, the results reflected in the intensity of landslide activity might be less than the actual state of landslide; rainfall intensity and infiltration might be greater than the actual situation. This model considered only the role of rain on the soil shear strength reduction, it was necessary to go further analysis on other factors that influenced and affected on the sliding body.

References

[1] J.H.Wang: Research the principle of water-soil interaction in soil-slope and analyze and predict its stability in condition of rain infiltration. Graduation Thesis (2006)

[2] X.Z.He: Italy landslide death toll statistics and compared with some countries. In journal: *The Chinese journal of geological hazard and control*. Vol. 14 (2003), p. 122-123.

[3] X.W.Liu, G.Liu: Analysis on Rainfall Influence on Slope Deformation and Failure. In journal: *Chinese Journal of Rock Mechanics and Engineering*. Vol. 22 (2003), p. 2715-2718.

[4] Fredlun D: Soil Mechanics for Unsaturated Soils. New York: *John Wiley&Sons publishing* (1993).

[5] J.M.Kong: Developmental Stage Characteristics of landslide and Observations. In journal: *Journal of Mountain Science*. Vol. 22 (2004), p. 725-729.

[6] R.Q.Huang: Mechanism of Large Scale Landslides in Western China. In journal: *Advance in Earth Sciences*. Vol. 19 (2004), p. 443-450.

[8] Y.R.Zheng, S.Y.Zhao： Application of strength reduction FEM in Soil and Rock Slope. In journal: *Chinese Journal of Rock Mechanics and Engineering*. Vol. 23 (2004), p. 3381-3388.

[9] W.Li: Ansys Civil Engineering Application. Beijing: *China waterpower press publishing* (2009).

[10] Dawson E M, Roth W H: Slope stability analysis by strength reduction. In journal: *Journal of Geotechnical and Geo-environmental Engineering ASCE*. Vol. 126 (2000), p. 75-80.

doi:10.4028/www.scientific.net/AMM.50-51.120

Impact and Control of Environmental Vibration on Precision Instruments

Zhengwei Gu[1, a], Tieyi Zhong [1, b], Mingbo Zhang[2, c] and Kun Zhang [1, d]

[1] School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China

[2] Central Research Institute of Building & Construction Co.,Ltd,MCC Group, Beijing 100088, China

[a]zhengweigu@163.com, [b]tyzhong2008@163.com, [c]m.zhang@foxmail.com, [d]jdakun@gmail.com

Keywords: environmental vibration, precision instrument, vibration control

Abstract. In this paper, the impacts of environmental vibration on precision instruments as well as the vibration allowances of precision instruments are introduced firstly. Then the integrated model of a high-rise building is established, and the vibration responses of the floor where a precise instrument is fixed are calculated by inputting micro-vibration loads. The passive control method is used by putting several lead rubber bearings at the bottom of instrument and at the bottom of building separately, and the isolated vibration responses are calculated and analyzed. Some valuable conclusions are obtained by comparing the vibration responses of isolation system and non-isolation system with acceptable values according to the environmental vibration codes for precise instruments.

Introduction

With the rapid development of modern industry and urbanization, impact of environmental vibration on production and on people's daily life has attracted widespread attention. Now vibration and noise has been listed as one of the seven hazards in the world.

The study on environmental vibration begins in the 1940s [1]. In the latest 30 years, Europe, the United States, Japan and some other developed countries have made some research on the production, dissemination, attenuation, refraction, focusing and resonance amplification of environmental vibration, and proposed standards of ground vibration isolation design method [2-6]. The studies also focused on ground vibration caused by railway transportation, especially caused by high-speed trains and heavy trains, and some calculation methods and calculation models are proposed [7-9]. In China, the researches on environmental vibration are also carried out, and some research methods and calculation models are obtained [10-12].

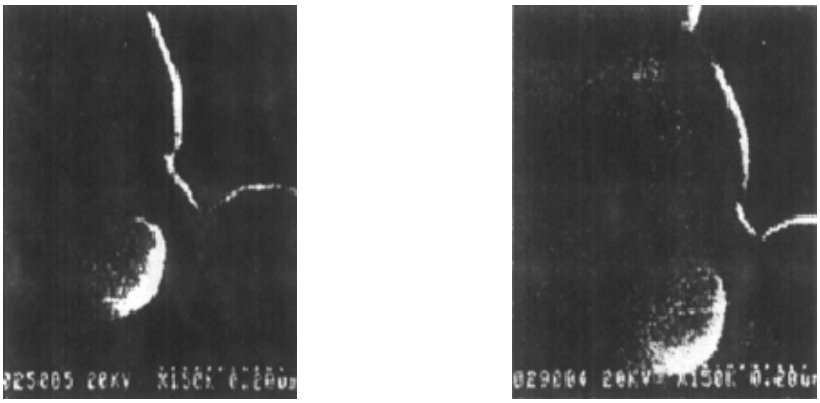
With the development of science and technology, human society has entered the "nanometer era". The great progress of industrial technology makes industrial products and instruments need high requirements on environmental vibration for production and operation. Urban rail transit line would inevitably run through factories, hospitals, laboratories and other areas where precision instruments are installed. When the vibration reaches some certain intensity, it will affect the normal use of the precision instruments. For instance, Beijing Metro Line 4 and Line 10 respectively pass closely to the Physics Laboratory of Peking University and a lot of sensitive instruments could be affected by metro trains. In order to ensure the normal teaching and scientific researching, the provision that environmental vibration frequencies which lower than 1Hz shall not exceed 5dB of ChengFuLu section is taken, at the same time the research on *Vibration and Noise Reduction Measures of Peking University Physics Laboratory* is carried out.

Environmental Vibration on Precision Instrument

Impact of Environmental Vibration on Precision Instrument. If the vibration velocity is larger than precision machining equipment allowances, the machines such as precision lathes, grinding machines, boring, milling and other precision equipments could reduce machining accuracy 1 to 2

levels, which will causes adverse effects on the smoothness, waviness, roundness, squareness and dimensional precision of products.

For precision measurement of physical and chemical analysis and the other testing instruments, the shaking or vibration of precision instruments occur when the vibration velocity exceeds vibration allowance. Thereby the vibration may cause measuring and testing system errors, as shown in Fig.1. It may reduce precision equipment service life, or even cause damage to the equipment directly.



(a) Normal image (b) Image impact of vibration

Fig. 1. Impact of vibration on electron microscope images

Precision Instrument Vibration Control Standards and Allowances. The allowed vibration levels of instruments are distinct according to their different characteristics. There are appropriate vibration allowances for different types of industries and instrument in China [13, 14]. The relevant documents of America and Japan point out that environmental vibration allowance of displacement is 1μm and the allowance of acceleration is 1gal for precision instruments. For large scale integrated circuit factory, the average value of vibratory displacement allowed is 0.1 ~ 1μm, and the average value of vibratory acceleration allowed is 0.1~1gal.

Environmental Vibration Responses and Vibration Control of Precision Instrument

Finite Element Model. The main body structure of the building is a reinforced concrete frame - tube structure, which has 18 layers with rectangular planes, and the height of each floor is 3m. A precise instrument is fixed on the 12th floor.

In order to make the calculation more close to the actual situation, the *M* method is used in this paper for calculating the coefficient of pile horizontal reaction, and the integrated model is established based on finite element software ANSYS. In this program, beam and column are simulated by three-dimensional beam element beam188, and shear wall and floor are simulated by shell element shell63. The FEA model is shown in Fig. 2. The coordinate system is set to the structure by defining x axis paralleling to the length direction and y axis paralleling to width direction as well as z axis paralleling to the height direction.

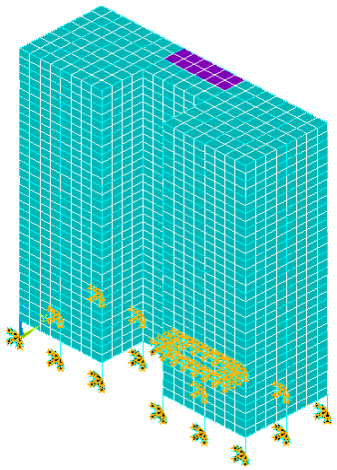


Fig. 2. FEA model of a high-rise building

A periodic micro-vibration load is taken as excitation in calculation, and the peak value of an actual environmental vibration wave is taken as peak load which the value is 0.025m/s^2. The propagation directions of the waves are assumed to parallel to x axis.

Elastic Foundation. The peak value of instrument responses are calculated, and the instrument acceleration is transformed to vibration level according to the standard of *ISO 2631*, which is shown as table1.

Table 1. Peak value of instrument responses

Label	X direction	Y direction
Velocity (mm/s)	0.168	0.373
Acceleration (m/s^2)	0.00559	0.00367
VL(dB) ($a_0=10^{-6}m/s^2$)	74.95	71.30

The vibration level of this instrument exceeded vibration allowable values defined in *ISO 2631/2-1989* and in *JBJ 16-2000*, so the following research will use the passive control method that put lead rubber bearing (LRB) at the bottom of instrument and at the bottom of building separately to meet the environmental vibration requirements of precise instruments.

Instrument Isolation. LRB is the most commonly used isolation device in passive control method. By using LRB in structure, the natural vibration period and damping of the structure increase, and energy imported into structure decreases. Moreover, because of rising of the damping, the structure's ability to consume energy is increased and the vibration response of the structure is decreased accordingly.

The responses of instrument with LRB or without LRB at the bottom of instrument are shown in Fig.3-Fig.4 below.

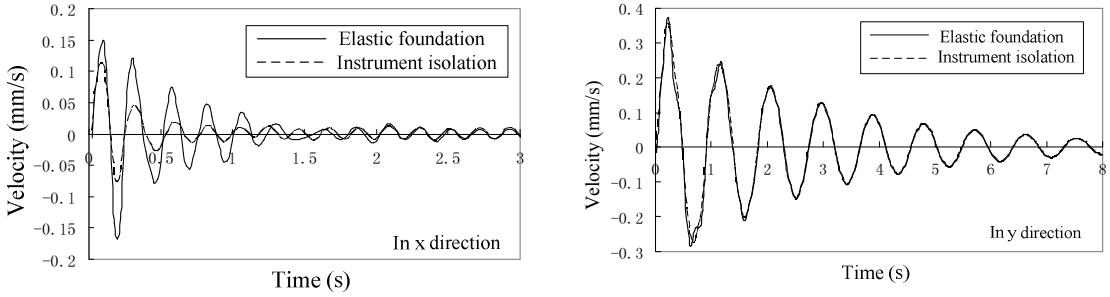


Fig. 3. Velocity responses with and without instrument bearing

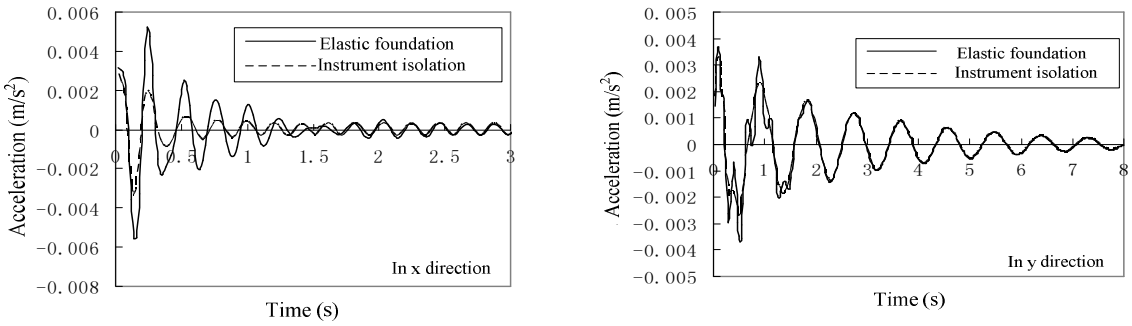


Fig. 4. Acceleration responses with and without instrument bearing

Fig.3-Fig.4 show that vibration responses change very slightly in y direction by taking LRB at the bottom of instrument, but the velocity responses and acceleration responses in x direction reduce obviously.

The acceleration responses can be effectively reduced by putting LRB at the bottom of instrument and the vibration level has reached the requirements in the codes.

Base Isolation. The responses of instrument with base isolation or without base isolation are shown as Fig.5-Fig.6.

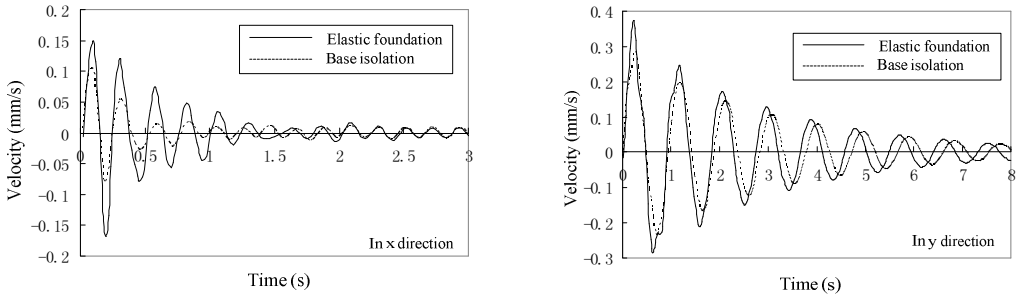


Fig. 5. Velocity responses with and without base isolation

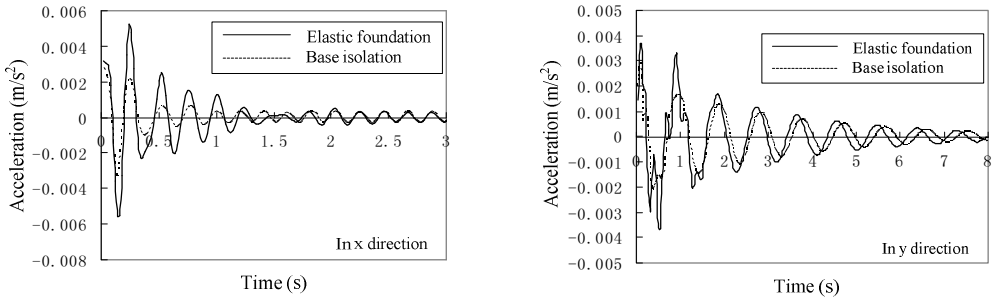


Fig. 6. Acceleration responses with and without basal bearing

It can be seen from Fig.5-Fig.6, velocity responses and acceleration responses of the instrument are reduced in x direction and in y direction. It is obvious that putting LRB at the bottom of building is an effective method to reduce the vibration responses of instrument.

The micro-vibration responses of precise instrument in the high-rise building are compared, and shown in Table2.

Table2. Vibration responses of precise instrument in a high-rise building

Maximum response		Connection mode			Allowance			
		Elastic foundation	Instrument isolation	Base isolation	ISO 2631/2	JBJ 16-2000	Chinese	American
VL(dB) ($a_0=10^{-6}m/s^2$)	X direction	74.95	69.24	68.90	71	71	——	——
	Y direction	71.30	70.35	69.51			——	——
Velocity [V] (mm/s)	X direction	0.168	0.113	0.106	——	——	V(0.2)	③ (0.2)
	Y direction	0.373	0.355	0.284	——	——	VII-VI(0.5-0.3)	② (0.4)

Table 2 shows that, vibration level of this precision instrument can meet the requirements defined in *ISO 2621/1* and in *JBJ 16-2002* both by putting LRB at the bottom of instrument and at the bottom of building. Besides, the base isolation method can reduce instrument speed more effectively; thereby vibration responses of this instrument can meet higher environmental vibration requirement according to Chinese cord, and the allowance level change from VII (0.5mm/s) to VI(0.3mm/s).

Conclusion

Although the responses caused by environmental vibration loads are tiny, the vibration acceleration and vibration velocity allowances for precise instrument are very small. Excessive vibration not only reduces precision machining accuracy or affects normal use of precision instruments, but also may reduce precision equipment service life or even cause damage to the equipment directly. Therefore carrying out research on analysis and control of precise instrument environmental vibration responses is very necessary.

The velocity and acceleration responses of a precision instrument in a high-rise building caused by a periodic micro-vibration load are obtained caused by a periodicity micro-vibration load. When without taking any vibration isolation measures, vibration acceleration level of instrument has exceeded the acceptable value of the codes, and velocity responses are larger than those of isolation structure. The acceleration responses can meet standards for building indoor vibration limit value, regardless of putting LRB at the bottom of instrument or at the bottom of building. The decrease of

velocity responses and acceleration responses is the most obvious in x direction, and the reduce rates are 32% and 36% respectively by putting LRB at the bottom of instrument, and 44% and 46% respectively by putting LRB at the bottom of building. In conclusion, the LRB can effectively reduce micro-vibration responses of precise instrument, which can improve the working performance of precise instrument largely.

Acknowledgments

This study is sponsored by the Natural Science Foundation of China (Grant No. 50538010), and the Belgium-China Bilateral Project (Grant No. BIL07/07).

References

[1] D.D. Barkan, Dynamic of bases and foundations. [M]. Mc Graw-Hill Bood Co., Inc., New York, N.Y., 1962.

[2] K.R. Massarch, Man made vibration and solution. [J]. Proceeding of 3rd international conference case histories geotechnical engineering, Vol.2, St. Louis, 1993.

[3] B.M. Zhao. T. Kagawa. M. Tsurugi. Study on long-period strong motion simulation for large-scale earthquakes. [J]. Bull. Seism. Soc. Am. 2007

[4] J. Avilés, Sanchez,-Sema, F.J. Foundation isolation from vibration using piles as barriers. [J]. Engineering Mechanics, Vol.144, No. EM11, Div ASCE, 1988.

[5] S. Ahmad, Al-Hussaini, T.M. Investigation on active isolation of machine foundation by open trenches. [J]. Geotechnical Engineering, Vol.122, No.GT6, Div ASCE, 1996.

[6] P. Jean, Villo, M. Study of the vibration power injected to a wall excited by a ground surface wave. [J]. Sound and Vibration. 2000.

[7] E. Kausel. Fundamental solutions in elastodynamices: a compendium. [M]. Cambridge University Press, New York. 2006

[8] C. Madshus, Kaynia, A.M. High-speed railway lines on ground: Dynamic behavior at critical train speed. [J]. Sound and Vibration, 2000.

[9] C. Jones, Sheng, X., Petyt, M. Simulations of ground vibration from a moving load on a railway track. [J]. Sound and Vibration, 2000.

[10] H. Xia, N. Zhang, Dynamic interaction of vehicles and structures. [M]. Science Press, Beijing, 2005.

[11] X.L. Du, M. Zhao, J. Y. Wang. A stress artificial boundary in FEA for near-filed wave problem. [J]. Chinese Journal of Theoretical and Applied Mechanics. 2006

[12] H. F. Zhou, J.Q. Jiang, T.Q. Zhang, Steady-state response of infinite beam on Kelvin foundation under moving loads. [J]. Journal of Zhejiang University. 2004

[13] GB50463-2008, Code for design of vibration isolation. [S]. 2008

[14] GB 50190-93, Code for design of anti-micovibration of multistory factor floor. [S]. 1993.

[15] ISO2631/2, Mechanical vibration and shock-evaluation of human exposure to whole body vibration-Part 2: continuous and shock induced vibration in buildings (1-80Hz) [S]. 1989

[16] JBJ 16-2000, Design code for environment protection of machinery industry [S]. 2000.

© 2010 Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.125

Assessment of the performance of Modal Based Decoupling Method in identifying structural joints characteristics

K. Jahani[1, a], A.S. Nobari[2, b]

[1]Mech. Eng. Dept. Tabriz Univ., Tabriz, Iran

[2]Aerospace .Eng. Dept. AUT, Tehran, Iran

[a]ka_jahani@tabrizu.ac.ir, [b]andishan.pars@gmail.coml

Keywords: Modal Based Decoupling Method, joint Identification, Damping, Stiffness

Abstract. In this paper, the performance of modal based decoupling method in structural joint identification is investigated. This method does not need any iteration and is independent from updating techniques. Because of these advantages, in present work it is tried to assess the effect of experimental and numerical issues of using modal data in this method. Issues such as effect of number of digits of precision of modal pairs, effects of lower hand higher eliminated modes, effect of any normalization of modal vectors and effect of possible noisy experimentally extracted modal parameters on identified joint characteristics are studied through some case studies.

Introduction

Identification of structural joints characteristics using updating methods, is depend on existing of analytical models for joint members and coupling of these models with an initial model for the joint region[1,2]. To get reasonable and accurate results, usually these methods need iterative procedure in other words the structural coupling repeated for several times that is a time consuming process.

In structural decoupling method that is independent from updating methods, the integrated structure is decomposed to its members and using the experimental or analytical modal parameters of the members and considering an analytical model for joint region, the parameters of joint are identified in one stage without any need to iteration and initial value to joint parameters [1].

The results of reference [1] show that despite of simplicity and not need to repetition, achieving accurate results using modal based decoupling method is difficult. In this paper, to do more investigation about this method, with some case studies it will be tried to study the effects of different computational and experimental issues on the performance of this method.

The governing equations of modal based decoupling method

The coupling equations for structure C that is constructed by coupling elastic sub-structures A and B through joint J, are as flow (more details are available in reference [1]) :

$$\begin{bmatrix}[\phi_{aki}] & [0]\\ [0] & [\phi_{bki}]\end{bmatrix}^T [Z_j]\left[\begin{bmatrix}[\phi_{aki}] & [0]\\ [0] & [\phi_{bki}]\end{bmatrix}-\begin{bmatrix}[R_{ai}] & [0]\\ [0] & [R_{bi}]\end{bmatrix}\begin{bmatrix}[\phi_{aki}] & [0]\\ [0] & [\phi_{bki}]\end{bmatrix}^{T+}[\Delta\lambda]\right]\begin{Bmatrix}\{p_a^c\}_{xr}\\ \{p_b^c\}_{xr}\end{Bmatrix}=[\Delta\lambda]\begin{Bmatrix}\{p_a^c\}_{xr}\\ \{p_b^c\}_{xr}\end{Bmatrix} \quad (1)$$

if the elastic substructure A is attached to the ground via joint J, equation (1) is simplified as:

$$[\phi_{aki}]^T_{k\times n_e}[Z_j]\left[[\phi_{aki}]_{n_e\times k}-[R_{ai}][\phi_{aki}]^{T+}_{n_e\times k}[\Delta\lambda]\right]\{p_a\}_{xr}=[\Delta\lambda]\{p_a\}_{xr} \quad (2)$$

$$\text{Where, } \Delta\lambda=[\lambda_{xr}]-[\omega^2_{ak}] \quad \text{and} \quad \{\phi_{xs}\}_{r_{nx}\times 1}=[\phi_{ak}]_{n_{xs}\times k}\{p_a\}_{xr_{k\times 1}} \quad (3)$$

In equation (1), $[R_{ai}]$ is the residue effect of eliminated modes which has two parts: residue due to rigid body modes, $[R_{ai}]_l$, and residue of eliminated high frequency modes, $[R_{ai}]_h$. i.e.:

$$[R_{ai}]=[R_{ai}]_l+[R_{ai}]_h \;;\; [R_{ai}]_h=[\phi_{aehi}][\omega_{ae}^2]^{-1}[\phi_{aehi}]^T \;;\; [R_{ai}]_l=\sum_{j=1}^{n}\{\phi_{aeli}\}_j\frac{1}{\omega^2}\{\phi_{aeli}\}_j^T \qquad (4)$$

Where, n is the number of rigid body modes and ω is resonance frequency of the mode in which the identification of joint characteristics is done.

In the studies of reference [1], damping has not been considered. However, here, to identify damping beside the dynamic stiffness of joint region, the complex modal vectors and eigen values of experimental modal tests will be used i.e. $\{p_a\}_{xr}$ and $[\lambda_{xr}]$ are complex. Therefore, it will be possible to compute complex $[Z_j]$ from equation (2).

Before applying the method, the correctness of extracted analytical complex modal pairs in this work is examined using a spring-dashpot system used by Adhikari [3] and illustrated in figure 1.The eigen value problem for a damped system is as follow:

$$s_k^2 M\psi_k + s_k C\psi_k + K\psi_k = 0 \qquad (5)$$

Where, s_k are eigen values and $\{\psi\}_k$ are modal vectors that satisfy the following orthogonality condition [ ]:

$$\psi_k^T[2s_k M + C]\psi_k = \theta_k \quad \forall k=1,...,2N \quad ; \quad \theta_k = 2s_k \quad \forall k. \qquad (6)$$

Equation (6) will be used in normalizing the modal vectors.

The extracted complex modal pairs using equations (5 and 6) in this work have been compared with the counterpart results presented by Adhikari in table 1. As it is seen form this table, the complex modal pairs extracted here, are the same presented by Adhikari. Therefore, the methodology of extracting analytical complex modal pairs in present work is faithful.

Table1. Comparison of extracted modal pairs in this study with that of h counterparts presented in ref. [3] for a spring-dashpot system (fig. 1)

	computed in present work		Adhikari
ψ	0.0115 + 0.0195i 0.0025 - 0.0204i -0.0066 + 0.0200i 0.0146 + 0.0153i		- -
s	-1.3786 +32.5838i -2.6214 +33.5570i		-1.3786 +32.5838i -2.6214 +33.5570i
ψ_n	0.8463 - 0.3349i -0.7589 - 0.4241i 0.7538 + 0.3877i 0.8296 - 0.3259i		0.8463 - 0.3349i -0.7589 - 0.4241i 0.7538 + 0.3877i 0.8296 - 0.3259i

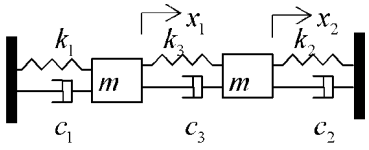


Fig 1. A two DOF spring-dashpot system with following parameters [3]:
k1= 1000 N/m; K2=1100 N/m; k3=50 N/m; m=1 kg; c1=3.0 N/m/s; c2=1.0N/m/s; c3=2.0 N/m/s

Studying the effect of various factors on performance of the modal based decoupling method in identifying joint characteristics

In this section, joint identification technique using modal based decoupling method is applied to the system that is presented in fig. 2 and the characteristics of joint i.e. k_j and c_j are identified($[Z_j]=k_j+Sc_j$). In present work, all 5 elastic modes will be used and the residue effect of rigid body modes is computed. Through some case studies, the performance of modal based

decoupling method under the following different conditions will be investigated:1- Effect of precision (decimal number of digits) of ψ *and* S; 2- Effect of normalizing of modal vectors; 3- Effect of $[R_{ai}]$; 4- Effect of adding noise to ψ *and* S

Case study 1- Effect of precision (decimal number of digits) of ψ *and* S**.** In this case study, modal pairs with 4 and 16 digits of precisions will be used in identification process. Samples of modal pairs with different digits of precisions are presented in Table 2. Identified values of joint parameters using each modal pair (for different sets of joint parameters) with considering 4 digits of precision for S and ψ in calculations are presented in Table 2 and compared with situation that16 digits of precision considered for the modal pairs. These results show that by using S and ψ with 4 digits of precision, the indentified values of joint damping ($c_{j_{identified}}$) are acceptable where the joint stiffness ($k_{j_{identified}}$) have big errors. Also, the correctness of identified values depends to the order of magnitudes of joint parameters. However, using S and ψ with 16 digits of precision, leads exactly identified k_j and c_j.

Table2.Samples of S and ψ with different number of digits of precisions

	4 digits	16 digits
s_2	-1.0900e+1+1.5139e+3i	-1.085547671125847e+001+1.513869780651153e+003i
$\psi(1,1)$	-0.6627 - 0.1976i	-0.66265133712611 - 0.19764103669266i

Table3. Comparison of identified joint parameters using 4 and 16 digits of precision for ψ *and* S in identification process with different desired values for joint parameters

Used Mode No.	Desired values		Identified values							
			accuracy of ψ *and* S : 4 digits				accuracy of ψ *and* S : 16 digits			
	k_j [N/m]	c_j [N/m/s]	$k_{j_{identified}}$ [N/m]	e [%]	$c_{j_{identified}}$ [N/m/s]	e [%]	$k_{j_{identified}}$ [N/m]	e [%]	$c_{j_{identified}}$ [N/m/s]	e [%]
1	300	30	301.60	0.5	30.0	0.00	300.00	0	30.00	0
2			432.34	44	29.98	-0.06	299.99	0	30.00	0
3			441.504	47	30.13	0.43	300.00	0	29.99	0
4			52.5891	-82	29.90	-0.35	299.99	0	30.00	0
5			-535.80	-279	30.21	0.07	299.99	0	30.00	0
1	3000	150	2999.6	-0.01	149.99	-0.00	3000.00	0	149.99	0
2			2824.4	-5.85	149.99	-0.00	3000.00	0	149.99	0
3			3078.3	2.61	150.00	0.00	2999.99	0	150.00	0
4			3145.8	4.86	149.91	-0.06	2999.99	0	150.00	0
5			4172.9	39.10	150.06	0.04	2999.99	0	150.00	0

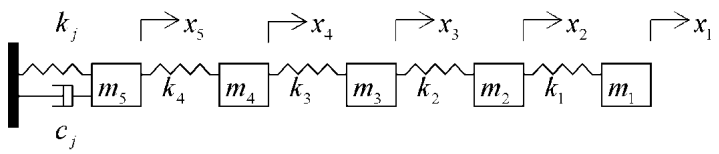


Fig 2. A 5 DOF spring-dashpot system with following parameters:
k1=k2=k3=k4 = 3e6 N/m; m1=m2-m4-m5-m6=0.5 kg; kj=300 N/m; cj=30N/m/s

Case study 2- Effect of normalizing of modal vectors. In previous case study, all of modal vectors before using in identification process were normalized using equation 6. However, in this case study for one set of desired c_j and k_j, un-normalized modal vectors with 16 digits of precision are implemented in identification process. The results that are presented in Table 4 show that there is no need to any normalization of modal vectors.

Table4. effect of normalizing of modal vectors on identified joint characteristics

Used Mode No.	Desired values		Identified values			
			Normalized		Un-normalized	
	k_j [N/m]	c_j [N/m/s]	$k_{j_{identified}}$ [N/m]	$c_{j_{identified}}$ [N/m/s]	$k_{j_{identified}}$ [N/m]	$c_{j_{identified}}$ [N/m/s]
1	3000	150	3000.000	149.999	3000.000	149.999
2			3000.000	149.999	3000.000	149.999
3			2999.999	150.000	3000.000	150.000
4			2999.999	150.000	2999.999	149.999
5			2999.999	150.000	3000.000	150.000

Case study 3- Effect of $[R_{ai}]$. In previous case studies, identification process were done considering residue effect ($[R_{ai}]$) in calculations. Here, to investigate the effect of considering $[R_{ai}]$on identified parameters, identification process for one set of desired joint parameters repeated without considering residue effect of eliminated modes. the results of considering and non-considering of $[R_{ai}]$in identification process are presented in Table 5.it can be concluded from the results of this table that, non-considering residue effect of eliminated modes will lead to big errors in identified characteristics of joint.

Table5. Effect of considering $[R_{ai}]$ on identified joint characteristics

Used Mode No.	Desired values		Identified values			
			With $[R_{ai}]$		Without $[R_{ai}]$	
	k_j [N/m]	c_j [N/m/s]	$k_{j_{identified}}$ [N/m]	$c_{j_{identified}}$ [N/m/s]	$k_{j_{identified}}$ [N/m]	$c_{j_{identified}}$ [N/m/s]
1	300	30	300.000	30.000	-2499889.998	10.999
2			299.999	30.000	-59.991	30.001
3			300.000	29.999	-59.999	30.001
4			299.999	30.000	-60.002	30.000
5			299.999	30.000	-60.002	30.000

Case study 4- Effect of noisy extracted modal pairs on identified results. Experimental modal data may contain some noise due to test chain and modal parameters extraction methods [4]. Here, to investigate the effect of the noisy modal data on identified joint characteristics, $\pm 5\ \%$ random noise is added on ψ and $\pm 1\ \%$ random noise added on S. With noisy modal pairs, identification process for previous case study is repeated here. The identified results are presented

in Table 6. These results show that with presence of noise, the identified values of $c_{j_{identified}}$ are acceptable whereas $k_{j_{identified}}$ are erroneous.

Table6. Effect of $\pm 5\ \%$ random noise of ψ and $\pm 1\%$ random noise of S on identified results

Used Mode No.	Desired values		Identified values			
	k_j [N/m]	c_j [N/m/s]	$k_{j_{identified}}$ [N/m]	e [%]	$c_{j_{identified}}$ [N/m/s]	e [%]
1	300	30	294.197	-2	29.707	-1
2			7456.766	2386	30.5104	1.7
3			8527.418	2743	29.085	-3
4			15231.192	4977	29.220	-2.6
5			2898.695	866	29.205	-2.7

Summary

In this paper, the performance of "modal based decoupling method" in structural joint identification investigated through some case studies. It can be concluded that this method is sensitive to the following items:

- The number of digits of precision of modal pairs (S and ψ) in this method is very crucial. It is advised to consider more than 8 digits of precision.
- None considering the residue effect of eliminated modes (rigid body modes and high frequency modes) will lead to erroneous identified joint characteristics.
- Normalizing of modal vectors does not affect identified joint characteristics.
- This method is sensitive to existing noise on modal parameters and noisy modal data lead to big errors in identified joint characteristics.

References

[1] A.S. Nobari: *Identification of the Dynamic Characteristic of Structural Joints* (Ph.D. Thesis, Imperial College, London, 1992)

[2] K. Jahani, A.S. Nobari: Exp. Mech., Vol. 48 (2008), p.599-611.

[3] S.Adhikari: Mechanical Systems and Signal Processing, Vol. 18 (2004), p. 1–27

[4] N.M.M. Maia, J.M.M. Silva: *Theoretical and Experimental Modal Analysis* (Research Studies Press LTD, John Wiley & Sons INC, 1997)

doi:10.4028/www.scientific.net/AMM.50-51.130

Construction Method of Fuzzy Attribute Information System Based on DEA and Attribute Reduction

Li Yankun

College of Light Industry, Hebei Polytechnic University, TangShan, Hebei, China

lyk_917@163.com

Keywords: fuzzy attribute information system; attribute reduction; data envelopment analysis; relative efficiency

Abstract: For the subjectivity of attribute value in the fuzzy attribute information system, this paper proposes a construction method of fuzzy attribute information system based on DEA (Data Envelopment Analysis). Then an attribute reduction method based on the dependability degree of attribute is given. At last, the proposed method is implemented successfully in the optimization of evaluation index system.

Introduction

In the classical rough sets theory [1], the condition attributes and decision attributes are both clear. Attributes correspond to the equivalence relations. We should make the continuous data discretely firstly before reducing. Traditional discrete technology may cause information loss because the membership degree of attribute values to the discrete values is not being considered. Because of these, rough sets theory has some limitations of practical application.

In the fuzzy rough sets theory, the condition attributes and decision attributes can be fuzzy due to the introduction of fuzzy ideas. Attributes correspond to the similarity relations. Before reducing, the discretization of attribute values is unnecessary [2]. But the determination of attribute values (membership function) has a certain degree of subjectivity in the fuzzy attribute information system. DEA method has prominent superiorities in avoiding the subjectivity, simplifying operation and decreasing error [3]. So this paper builds fuzzy attribute information system through integrating DEA with fuzzy rough sets theory. Then an attribute reduction algorithm is put forward and implemented in the optimization of evaluation index system.

Prior Knowledge

C^2R Model. Data Envelopment Analysis [4] evaluates the relative efficiency of decision making units (DMU) with multiple input and multiple output by mathematical programming model.

Suppose there are n DMU, and each one has m kinds of input and s kinds of output. Denote $\boldsymbol{X}_j = (\boldsymbol{x}_{1j}, \boldsymbol{x}_{2j}, \cdots, \boldsymbol{x}_{mj})^{\boldsymbol{T}}$, $Y_j = (y_{1j}, y_{2j}, \cdots, y_{sj})^T$, $j = 1, \cdots, n$ as the input and output of the jth DMU. Apply the C^2R model of DEA to built model of each DMU$_i$ $(i = 1, \cdots, n)$:

$$\begin{cases} \min \theta \\ \sum_{j=1}^{n} X_j \lambda_j \le \theta X_0 \\ \sum_{j=1}^{n} Y_j \lambda_j \le Y_0 \\ \lambda_j \ge 0 \end{cases} . \tag{1}$$

If the optimal solution $\theta = 1$, then the DMU$_i$ is DEA effective. Otherwise, it is DEA ineffective.

The dependability degree of decision attribute. Let $U = \{x_1, x_2, \cdots, x_n\}$ be the domain and $C = \{C_1, C_2, \cdots, C_m\}$ be a set of fuzzy condition attributes. Every attribute is generally expressed as a set of fuzzy language items $A(C_i) = (F_{ik} \mid k = 1,2,\cdots, p)$. $U/C = \{F_{ik} \mid i = 1,2,\cdots, m, k = 1,2,\cdots p\}$ is a fuzzy partition of U by fuzzy similarity relation R [5] on U. D is the decision attribute. $\forall x_i \in U$ is divided into class sets $A(D) = \{F_l \mid l = 1,2,\cdots, q\}$. F_l can be precision set or fuzzy set.

Given arbitrary fuzzy set A. $\mu_A(x) : U \to [0,1]$ is the membership function. For $\forall x \in U$, $F_{ik} \in U/P$, the tuple $(\mu_{\underline{A}}, \mu_{\overline{A}})$ constitute fuzzy rough sets. $\mu_{\underline{A}}$ and $\mu_{\overline{A}}$ are called lower and upper approximation membership function. They are defined as follows [6]:

$$\mu_{\underline{A}}(F_{ik}) = \inf_{x \in U} \max\{1 - \mu_{F_{ik}}(x), \mu_A(x)\} . \tag{2}$$

$$\mu_{\overline{A}}(F_{ik}) = \sup_{x \in U} \min\{\mu_{F_{ik}}(x), \mu_A(x)\} . \tag{3}$$

The membership degree of F_{ik} under fuzzy positive domain of D is defined as

$$\mu_{POS}(F_{ik}) = \sup_{F_l \in A(D)} \{\mu_{F_l}(F_{ik})\} . \tag{4}$$

The membership degree of $\forall x \in U$ for fuzzy positive domain is defined as

$$\mu_{POS}(x) = \sup_{F_{ik} \in A(C_i)} \min\{\mu_{F_{ik}}(x), \mu_{POS}(F_{ik})\} . \tag{5}$$

The dependability of decision attributes D for condition attributes C is defined as

$$\gamma_C(D) = \frac{\sum_{x \in U} \mu_{POS}(x)}{n} . \tag{6}$$

Construction model for system

Membership degree is the major component of fuzzy rough sets model. The determination of membership function has a certain degree of objective regularity and scientificity, but also exists the more dense subjectivity. This is one of factors which impact the objectivity of knowledge discovery in fuzzy rough sets theory.

For a fuzzy attribute, we can select the important factors according to the practical significance to structure input and output index. Then we can build the DEA model for the original data and evaluate the relative efficiency for elements. The relative efficiency of the elements for the fuzzy attribute could be regarded as membership degree. From this, the fuzzy attribute information system can be built.

For example, a is a fuzzy attribute. Select the factors $\{a_1, a_2, \cdots, a_6\}$ which have prominent effect for a and classify them into input indexes and output indexes according their nature. See Table 1.

Table 1 Input-output data of fuzzy index a

	input				output	
	a_1	a_2	a_3	a_4	a_5	a_6
x_1	21	2020	1616	602	24	293
x_2	32	1019	899	350	59	1617
x_3	69	1918	1542	319	198	1015
x_4	51	2217	1918	720	429	3100
x_5	36	1217	946	298	98	631

Calculate the relative efficiency with the Linear Programming (1997). Show the results in Table 2. (Select one significance digit after the decimal point.)

Table 2 Relative efficiency of object set on index a

object	x_1	x_2	x_3	x_4	x_5
efficiency	1	0.9	0.5	1	0.4

According to the result of efficiency, the membership degrees of $x_1, x_2, \cdots, x_5$ for the fuzzy attribute a follow by 1, 0.9, 0.5, 1, 0.4.

It can be seen that the subjectivity caused by determination of membership function can be effectively avoided through the construction method of fuzzy attribute information system based on DEA.

Fuzzy attributes reduction algorithm

Let $C = \{C_1, C_2, \cdots, C_m\}$ be the fuzzy condition attribute set, and D be the fuzzy decision attribute.

Let $P \subseteq C$, $|P| = i$ ($|\cdot|$ is denoted by the number of elements in a set).

If $\gamma_P(D) = \gamma_C(D)$, then $red(C) = P$. Suppose the decision table is consistent. Reduction algorithm is as follows:

For i=1 to m

If $\gamma_P(D) = \dfrac{\sum_{x \in U} \mu_{POS}(x)}{n} = 1$

Then Red=P
Else i=i+1
End if

Example

In the areas such as evaluation and decision-making, the design of index system should pay attention to the relevance which exists between indexes, when we seek the comprehensiveness. If we treat the index as the attribute, the reduction of index set can be obtained by the attribute reduction. This is known as the optimization of index system. For the fuzzy index, we can treat it as the fuzzy attribute, and then the optimization of fuzzy index system can be achieved through the above-mentioned construction method and attribute reduction algorithm.

Let $U = \{x_1, x_2, \cdots, x_5\}$ be the set of evaluation objects, $C = \{a, b, c, d, e\}$ be the set of fuzzy indexes and $D = \{h\}$ be the fuzzy decision index.

Step1: Build the model of DEA for each index by model (1) and evaluate efficiency of elements of U. Construct fuzzy attribute information system by relative efficiency instead of membership degree. See Table 3.

Table 3 Fuzzy information system

	a	b	c	d	e	h
x_1	1	0.8	1	0.6	0.9	1
x_2	0.9	0.6	0.8	0.7	1	0.8
x_3	0.5	1	0.6	0.9	0.7	0.6
x_4	1	0.8	0.8	1	1	1
x_5	0.4	0.5	0.4	0.7	0.3	0.5

Step2: Using the above-mentioned attribute reduction algorithm, we get a reduction $\{a, b\}$. It is the optimized result of original index set.

Conclusion

This paper proposes the construction method of fuzzy attribute information system based on DEA firstly. This method applies an effective way to avoid the subjectivity which is brought by the determination of membership degree. Then, a new reduction method based on the dependability degree of attribute is given. Finally, the method is applied to the optimization of evaluation index system.

References

[1] Zhang Wenxiu, Wu Weizhi, etc. Rough Set Theory and Method [M]. Beijing: Science Press, 2001.

[2] Nie Zuoxian, Liu Jiancheng. An Attribute-reducing Algorithm Based on GA and Fuzzy Rough Set Theory [J]. Computer Engineering, 2005, 31(6): 163-165.

[3] Hassan Y, Gibreel G. Evaluation of highway consistency and safety practical application [J]. Journal of Transportation Engineering, 2000, (3): 127-131.

[4] Wei Quanling. Data Envelopment Analysis [M]. Beijing: Science Press, 2004.

[5] RADZIKOW SKA AM, KERRE EE. A comparative study on fuzzy rough sets[J]. Fuzzy Sets and Systems, 2002, 126(2): 137-155.

[6] Wang Li, Feng Shan. Two attribute reduction algorithms based on fuzzy-rough set [J]. Computer Applications, 2006, 26(3): 635-637, 672.

© (2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.135

Study of Parameters Optimization on Seismic Isolated Railway Bridges

Tieyi Zhong [1, a], Chaoyi Xia[1, b], Fengli Yang[2, c]

[1] School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China

[2] Beijing Electric Power Construction Research Institute of SGCC, Beijing 100055, China

[a]tyzhong2008@163.com, [b]xiacy88@163.com,[c] yangfl@epcri.com

Keywords: seismic isolation, lead core bearing, optimal design, railway bridge.

Abstract. Based on optimization theories, considering soil-structure interaction and running safety, the optimal design model of the seismic isolation system with lead-rubber bearings (LRB) for a simply supported railway beam bridge is established by using the first order optimization method in ANSYS, which the parameters of the isolation bearing are taken as design variables and the maximum moments at the bottom of bridge piers are taken as objective functions. The optimal calculations are carried out under the excitation of three practical earthquake waves respectively. The research results show that the ratio of the stiffness after yielding to the stiffness before yielding has important effect on the structural seismic responses. Through the optimal analysis of isolated bridge system, the optimal design parameters of isolation bearing can be determined properly, and the seismic forces can be reduced maximally as meeting with the limits of relative displacement between pier top and beam, which provides efficient paths and beneficial references for dynamic optimization design of seismic isolated bridges.

Introduction

Vibration reduction and isolation technology provide a new way for bridge antiseismic and has been used in railway bridge engineering in China, but there still are many problems not solved in calculation analysis and design theory. In recent years, many researchers have investigated the seismic absorption and isolation design of bridges. Zhu Dongsheng et al. studied on the design parameters of the seismic isolation bridges, and the effects of initial period, ductility ratio as well as the ratio of the stiffness after yielding to the stiffness before yielding on the seismic isolation of bridges were investigated [1]. However, the seismic isolation design method considering all the parameters comprehensively weren't given. Li Jianzhong et al. researched on the optimum design of seismically isolated continuous bridges and presented the optimum design method for continuous bridges [2]. As for the application of isolation bearings, the beam displacements increase as soon as the forces at the bottom of the pier reduce on the seismic excitations. Kyu-Sik Park et al. solved the contradiction between the reduction of the maximum forces and the increment of the beam displacement by introducing the weight coefficients [3]. The soil-foundation interaction and the effects of the dynamic parameters of the lead core rubber bearing (LRB) on the Rayleigh damping coefficients weren't considered in the three studies above. The lead core rubber bearing between the beam and pier can reduce the seismic responses of bridges efficiently [4], and the optimum study on the LRB design parameters is done in this paper based on the single pier model considering the soil-foundation interaction. After the optimum design, the maximum forces of the bottom of the pier can be decreased at the most with the relative displacement between the beam and the top of the pier constrained in the allowable range. The hysteretic energy-dissipated characteristics of LRB can be determined by its dynamic controlling parameters which include the yield stress, the stiffness before yielding and the ratio of the stiffness after yielding to the stiffness before yielding. The optimum study on the dynamic controlling parameters of LRB provides important reference for the proper seismic isolation design of the railway simple supported beam bridge.

Analysis Model

Bilinear Model of the Lead Core Rubber Bearing. As shown in Fig. 1, u_B represents the efficient design displacement of the bearing and (u_y, Q_y) represents the yield point of the bearing [4]. The value of yield stress Q_y mainly depends on the weight of the beam structure and the producing techniques, and u_y is the yield displacement of the bearing. The stiffness after yielding K_2 can be calculated according to Fig. 1 by formula (2.1).

$$K_2 = \frac{F(u_B) - Q_y}{u_B - u_y} \tag{2.1}$$

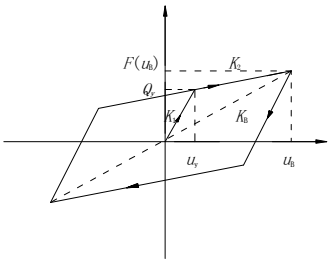


Fig. 1. Controlling parameters of LRB

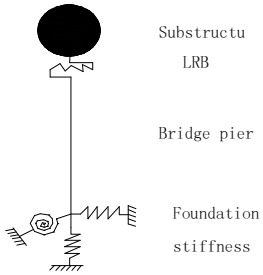


Fig. 2. FEA model

FEA Model for the Railway Simple Supported Beam Bridge. A typical pier of a super giant railway simple supported bridge is researched in this study, and the basic design parameters of the bridge can be seen in Table 1. According to the Code for Design on Subsoil and Foundation of Railway Bridge and Culvert [12], the foundation elastic stiffness can be calculated by M-method. The FEA model of the beam-bearing-pier-soil system is shown in Fig. 2. The general program ANSYS is used for the calculation analysis. In this model, the dynamic characteristics of LRB can be simulated by nonlinear element Combin39, and Mass21 element is applied for the substructure. The soil-foundation interaction is well simulated by an arbitrary element (Matrix27) whose geometry is undefined but whose elastic kinematics response can be specified by stiffness, damping, or mass coefficients. The foundation elastic stiffness values are listed in Table 2. The site which the bridge locates is I-site, and 3 earthquake waves of San Fernando, El Centro and Long Beach are selected as the excitations for the optimization of the LRB design parameters.

Table 1. General design parameters of the typical pier

Design parameters of the pier-section	Lumped mass on the top of pier (ton)	Height of the pier(m)
Solid circle-head shape, diameter of the pier top: 2.10m,width of the rectangular: 1.70m， slope:1/40	420	21

As seen in Table 2, the foundation stiffness values for the typical railway simple supported beam bridge are calculated by M-method when the m-value equals to 50000 KPa/m^2.

Table 2. Foundation elastic stiffness for m=50000 KPa/m^2

Foundation stiffness	$k_{11}(\times 10^9)$	$k_{12}=-k_{21}(\times 10^9)$	$k_{22}(\times 10^9)$	$k_{33}(\times 10^9)$
Values	1.6206 (N/m)	-3.3422 (N/rad)	9.859 (N·m/rad)	76.011 (N/m)

Optimization Program

Mathematic Description of the Design Parameters Optimization of LRB. The dynamic controlling parameters consist of the yielding strength Q_y, the stiffness before yielding K_1 and the ratio of the stiffness after yielding to the stiffness before yielding α. The seismic absorption and isolation can be realized by the application of LRB for its good hysteretic property. It is known that the application of LRB can lengthen the structural period and reduce the earthquake forces, but it may

increase the displacement between beam and top of pier [5]. This indicates that there exist optimum values of the dynamic design parameters for which the earthquake forces in the bridge are minimum. Therefore, Q_y, K_1 and α are taken as the design variables for the optimization of the whole structural system. Three types of intensities including minor earthquake, design earthquake and severe earthquake are considered in the optimal design for the whole isolated bridge system, so that the bridge structure can remain elastic when subjected to a minor earthquake, while the LRB can undergo inelastic deformation to dissipate the seismic energy under a severe earthquake.

The maximum moment of the bottom of the pier on design earthquake excitation is taken as the objective function. Because the application of LRB enlarges the relative displacement between beam and top of pier, the beam would collapse on the design seismic excitation if the displacement weren't constrained. Accordingly, as the state variable in the optimum process, the relative displacement between beam and top of pier should be constrained so as to ensure the maximum relative displacement in the permissible range. The running safety of vehicles can be ensured by the index of spectral intensity SI [10-11]. The optimization problem of the design parameters of LRB for railway simple supported bridge can be mathematically described as follows [6].

The objective function is

$$\min_{x} \max_{t \in T} \{abs(\mathrm{M}(x,t))\} \tag{3.1}$$

where $\mathrm{M}(x,t)$ is the moment of bottom of pier; and the constrains are

$$\mathrm{x}_i^{\mathrm{L}} \le \mathrm{x}_i \le \mathrm{x}_i^{\mathrm{U}} \quad (\mathrm{i}=1,\ 2,\ 3) \tag{3.2}$$

$$\max_{t \in T} abs(\mathrm{rd}_j(x,t)) \le [\mathrm{rd}] \quad (\mathrm{j}=1,\ 2,\ 3,\ \cdots,\mathrm{n}) \tag{3.3}$$

$$SI \le SI_{\mathrm{Lim}} \tag{3.4}$$

where the vector of design variables $\mathrm{X}^{\mathrm{T}} = [\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3]^{\mathrm{T}} = [Q_y, K_1, \alpha]$; T =the delay time of the seismic excitation and Δt =the record interval of the seismic excitation, $\mathrm{n} = \dfrac{T}{\Delta \mathrm{t}}$; $\mathrm{x}_i^{\mathrm{U}}$ =the upper limit value of the design variable and $\mathrm{x}_i^{\mathrm{L}}$ = the low limit value of the design variable; $X_{\mathrm{U}}^{\mathrm{T}}$, $X_{\mathrm{L}}^{\mathrm{T}}$ =the upper limit vector and the low limit vector of the design variables. According to the test datum of the lead core rubber bearings from China Academy of Railway Science, $\mathrm{X}_{\mathrm{L}}^{\mathrm{T}} = [2\times10^4\,\mathrm{N}, 1\times10^7\,\mathrm{N/m}, 0.1]$ and $\mathrm{X}_{\mathrm{U}}^{\mathrm{T}} = [2\times10^5\,\mathrm{N}, 1\times10^8\,\mathrm{N/m}, 1]$ [7]; rd_j =the relative displacement between beam and top of the pier at different time of the seismic excitation history; $[\mathrm{rd}]$ =the acceptable value of relative displacement between beam and top of the pier. SI is the spectral intensity corresponding to the fundamental period of the isolated bridge, and SILim is the limit value of SI. The value of SILim is selected as 4000mm according to reference [10].

Solution for the Parameters Optimization of LRB. The first order optimization method is adopted to solve the optimum problem. The fundamental theory of this method is transforming the common optimum problem to single unconstrained optimum problem by the mixed penalty function, and many analysis loops can be carried out in an optimum iteration [8, 9]. In the optimum process, the damping ratio of the structural system without isolation is determined as 5%. The first two frequencies(ω_1, ω_2) of the system and the Rayleigh damping coefficients corresponding to the change of LRB design parameters can be determined by calculation.

Optimized Results

Three dynamic controlling parameters of LRB are taken as design variables in the optimization of LRB, and the limit values of the relative displacement between beam and top of the pier corresponding to three intensities of earthquakes are set at 2cm, 5.5cm and 10cm respectively. Three intensities of earthquakes are minor earthquake, design earthquake and severe earthquake. The optimum solutions are shown in Table 3.

Table 3. Design parameters of LRB and the seismic responses

Items	Minor earthquake			Design earthquake			Severe earthquake		
	San Fernando	El Centro	Long Beach	San Fernando	El Centro	Long Beach	San Fernando	El Centro	Long Beach
$Q_y(10^5N)$	1.0098	1.4468	1.3484	1.1798	2.9530	3.3285	1.050	4.386	3.464
$K_1(10^7N/m)$	2.069	1.000	4.825	3.883	6.140	8.135	5.877	9.956	9.753
α	0.354	0.123	0.085	0.302	0.116	0.167	0.339	0.126	0.184
rd(cm)	1.94	1.77	1.93	5.47	5.37	4.81	9.99	9.92	9.26
SI (mm)	550.2	623.0	430.1	2024.2	1730.6	1523.5	4499.8	3538.6	3104.4
$M(10^7N\cdot m)$	0.6170	0.4051	0.7133	2.5606	2.0264	2.7638	6.2474	4.6170	5.6000

The iteration histories of the maximum relative displacement between beam and the top of the pier, and the iteration history of maximum moment at bottom of the pier as well as the time-history of the relative displacement between beam and top of the pier for design earthquake intensity of San Fernando earthquake excitation are shown in Fig. 3, Fig. 4 and Fig. 5 respectively.

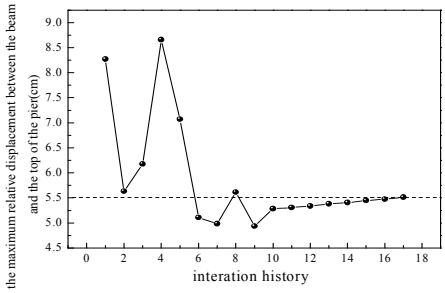


Fig. 3. Iteration history of the maximum relative displacement between beam and the top of the pier

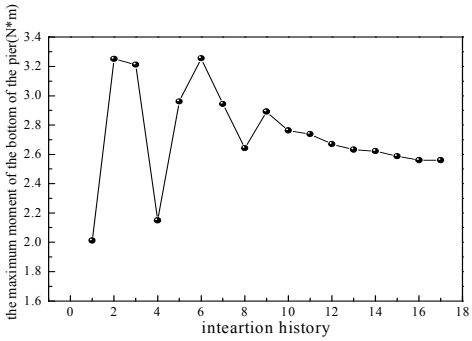


Fig. 4. Iteration history of the maximum moment at the bottom of the pier

As shown in Fig. 3 and Fig. 4，the optimum analysis can converge quickly and obtain optimum solution by using first order method. The relative displacement histories corresponding to initial design parameters and optimum design parameters for the design intensity earthquake are presented in Fig. 5, and the results indicate that the maximum relative displacement can be reduced efficiently through the optimum of the LRB design parameters.

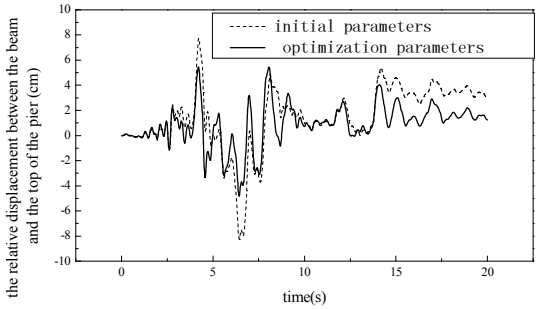


Fig. 5. Time-history of the relative displacement between beam and top of the pier

Conclusions

(1) The first order method is appropriate to solve large-scale dynamic optimum problem on seismic isolated railway bridges considering the soil-foundation interaction and running safety of vehicles, and it's easy to converge in the calculation analysis.

(2) The analysis results show that variation of the LRB dynamic controlling parameters have significant effects on the seismic response of the railway simple supported bridge. In order to acquire the reasonable parameters which meet both good vibration reduction effect and running safety requirement of vehicle, the dynamic optimization of the LRB dynamic controlling parameters is necessary and meaningful to the seismic isolation design for bridges.

(3) In order to enhance the optimal efficiency, the running safety can be considered out of the optimal process and the values of spectral intensity for the seismic input to the running vehicles can be checked after determining the optimal bearing parameters. Through the optimal analysis of isolated bridge system, the optimal design parameters of isolation bearing can be determined properly, and the seismic forces can be reduced maximally as meeting with the limits of relative displacement between pier top and beam

(4) Optimal calculation with three practical earthquake waves as input respectively is carried out. The results indicate that the optimal parameters take obvious difference, so it is not appropriate using the optimal parameters based on only one practical earthquake wave. Further optimal design study should be done considering multi-earthquake waves objective function as input to obtain the optimum solution and ensure the safety of bridge structure and vehicle in earthquakes.

Acknowledgment

This study is sponsored by the Science & Technology Research Development Project of China Railway Ministry (Grant No. 2008G010-A) and the Science & Technology Research Development Project of China Railway Ministry (Grant No. 2008G028-C).

References

[1] Zhu Dongsheng, Lao Yuanchang(2000). Research on Some Design Parameters of Seismically Isolated Bridges . China Civil Engineering Journal 33:5, 53-56.

[2] Li Jianzhong, Yuan Wancheng, Fan Lichu(1998). Optimal Design of Seismic Isolation System for Continuous Bridge. China Civil Engineering Journal 31:3, 47-54.

[3] Kyu-Sik Park, Hyung-Jo Jung, In-Won Lee(2002). A comparative study on aseismic performance of base isolation systems for multi-span continuous bridge . Engineering Structures 24, 1001-1013.

[4] Wang Li(2002). Performance Analysis of the Lead Rubber Bearing in Seismic Isolation of Bridge [D]. Northern Jiaotong University, Beijing, China.

[5] R.I.Skinner，W.H.Robinson，G.H.MeVerry，Translated by Xie Lili(1996). Engineering Isolation Theory, Earthquake Press, Beijing, China.

[6] Rong Jianhua, Zheng Jianlong, Xu Feihong(2002). Structural Dynamic Modification and Parameters Optimization . Communication Press, Beijing, China.

[7] Wu Bin(2003). Study on Mechanical Characteristic of Lead core rubber Bearing and the Response of Isolated Bridge with the Lead core rubber Bearing.China Academy of Railway Sciences, Beijing, China.

[8] ANSYS INC(2005).First Order Optimization Method. Ansys6.1 Help. American.

[9] Chen Zhiheng(2004). Study on Northern Channel Bridge Structure Optimization of Hangzhou Bay Bridge. China Railway Science 25:3, 76-79.

[10] Xia H, Zhang N (2005). Dynamic Interaction of Vehicles and Structures. Science Press, Beijing, China.

[11] Yang Fengli (2007). Study on Design Method and Parameters for Seismically Isolated Railway Bridges. Beijing Jiaotong University, Beijing, China.

[12] Ministry of Railway of the People's Republic of China(2000). Code for Design on Subsoil and Foundation of Railway Bridge and Culvert, TB10002.5-99. China Railway Press, Beijing, China.

doi:10.4028/www.scientific.net/AMM.50-51.140

Analysis of Fault Tree Base on Uncertain Optimistic Value

Jianjun Liu[1a],Yufu Ning[1]

[1]Department of Computer Dezhou University,Dezhou Shandong China

[a]ljj@dzu.edu.cn

Keywords: fault tree; uncertain variable; optimistic value; top event; bottom event

Abstract. Based on uncertainty theory, this paper proposes a method that constructs and analyzes fault tree. In this paper, it would be characterized as crisp number if fault rate of bottom event is obtained from reliable handbook, empirical data and so on; it would be characterized as uncertain variable if fault rate of bottom event has no statistical data but is obtained from expert's subjective judgment. The optimistic value of overall system's top event is calculated by using uncertain simulation technology. Finally feasibility and validity of this method is confirmed by taking fault tree of internal combustion engine as example.

1 Introduction

Fault tree is a special inverted treelike logic graph, it describes causality among all kinds of events in system with event symbol, logic gate symbol and transition symbol. Fault tree analysis method is to analyze all kinds of factors that make system invalid to construct fault tree. It can determine possibility of system failure with some method and improve reliability of system by taking corrective action.

The traditional fault tree analysis can effectively process massive failure events by using probability theory, but still has the following problems: (1) Fault tree analysis is carried on at early design stage of system, so it is difficult to obtain past failure data; (2) New components are usually used in practical application, default probability of them cannot be obtained. (3) In system involving persons, 20%-90% of system failures are caused by human factors, it is very difficult to be characterized with probability theory. Therefore the above problems are solved by using theory and method of fuzzy mathematics in many literature. Paper [1] depicts probability which event occurs with fuzzy number, carries on fuzzy fault tree analysis with fuzzy set theory and fuzzy operator, has overcome shortcoming that probability is difficult to be evaluated precisely in traditional fault tree analysis, it is a simple and effective fault tree analysis method. Paper [2] has redefined two suppositions in reliable theory: (1) fuzzy probability supposition: probability of event is expressed with fuzzy variable, namely breakdown behavior of system and composition unit is described with fuzzy probability; (2) two condition supposition: it is as same as traditional reliable theory. Paper [3],[4],[5] analyzed detailedly that how probability of top event is calculated by using triangle fuzzy number, trapezoidal fuzzy number, LR fuzzy number, normal fuzzy number as probability of bottom event. Paper [6] describes all factors with random fuzzy variable, calculates source opportunity that top event occurs, this method has widespread utility.

The above fuzzy fault tree analysis methods have following insufficient: (1) The construction process of fault tree has not been analyzed. (2) In the actual project, it is very difficult to describe occurrence opportunity of fault tree's bottom event. Therefore in this fault tree analysis, it would be expressed with the crisp number if the fault rate of bottom event is obtained from reliable handbook, empirical data and so on; it would be expressed with uncertain variable obtained from expert's subjective judgment if the fault rate has no statistical data. The optimistic value of overall system's top event is calculated by using uncertainty simulation technology. Finally feasibility and validity of this method is confirmed by taking the fault tree of internal combustion engine as example.

2 Preliminaries

Definition1[8]. Let Γ be a nonempty set. L is a σ-algebra over Γ,each element $\Lambda \in L$ is called event, it is necessary to assign to each event Λ a number $\mathrm{M}\{\Lambda\}$ which indicates the level that will occur. The set function $\mathrm{M}\{\Lambda\}$ is called uncertain measure if it satisfies the following four axioms.

(1) (Normality) $\mathrm{M}\{\Gamma\}=1$.

(2) (Monotonicity) $\mathrm{M}\{\Lambda_1\} \le \mathrm{M}\{\Lambda_2\}$ whenever $\Lambda_1 \subset \Lambda_2$.

(3) (Self-Duality) $\mathrm{M}\{\Lambda\}+\mathrm{M}\{\Lambda^c\}=1$ for any event Λ.

(4) (Countable Subadditivity) For every countable sequence of events $\{\Lambda_i\}$, we have $\mathrm{M}\{\bigcup_{i=1}^{\infty}\Lambda_i\} \le \sum_{i=1}^{\infty}\mathrm{M}\{\Lambda_i\}$.

Definition2[8]. Let Γ be a nonempty set, L a σ-algebra over Γ, and M an uncertain measure, then the triplet (Γ, L, M) is called an uncertain space.

Definition3[8]. An uncertain variable is a measurable function ξ is from an uncertain space (Γ, L, M) to the set of real numbers i.e,for any *Borel* set B of real number, the set $\{\xi \in B\}=\{\gamma \in \Gamma \mid \xi(\gamma) \in B\}$ ξ is an event

Definition4[8]. An uncertain variable is said to have a first identification function λ,if

(1) $\lambda(x)$ is a nonnegative function on R such that $\sup_{x \ne y}(\lambda(x)+\lambda(y))=1$;

(2) for any set B of real numbers, we have

$$\mathrm{M}\{\xi \in B\}=\begin{cases}\sup\limits_{x\in B}\lambda(x), & \text{if } \sup\limits_{x\in B}\lambda(x)<0.5\\ 1-\sup\limits_{x\in B}\lambda(x), & \text{if } \sup\limits_{x\in B}\lambda(x)\ge 0.5\end{cases}$$

Example1 By a triangular uncertain variable we mean the uncertain variable fully determined by the triplet (a,b,c) of crisp numbers with $a<b<c$, whose first identification function is

$$\lambda_2(x)=\begin{cases}\dfrac{x-a}{2(b-a)}, a\le x\le b\\ \dfrac{x-c}{2(b-c)}, b\le x\le c\end{cases}$$

Definition5[8]. Let ξ be an uncertain variable .and $\alpha \in [0,1]$.Then

$\xi_{\sup}(\alpha)=\sup\{r \mid M\{\xi \ge r\} \ge \alpha\}$

is called the α-optimistic value to ξ.

3 Analysis of fault tree base on uncertain optimistic value

The main analysis steps of fault tree base on uncertainty measure are listed as follows:

Step1 The top event is choosed, then fault tree is constructed based on actual situation.

Step2 The bottom events of fault tree are respectively characterized as crisp probability and uncertain variable, it would be characterized as crisp number if fault rate of bottom event is obtained from reliable handbook, empirical data and so on; it would be characterized as triangular or trapezoidal uncertainty variable obtained from expert's subjective judgment if fault rate of bottom event has no statistical data.

Step3 The expression of top event's uncertainty measure is expressed with "and" and "or" operator.

Step4 The optimistic value of top event is calculated By using uncertain simulation algorithm.

Step5 At last the result is analyzed.

4 Application of fault tree

Table1 Chance that bottom event occurs

code	Basic event	Chance or evaluated result
q_1	The fault rate of vaporizer is low	0.02
q_2	The last oil is exhausted	An triangular uncertain variable determined by triplet (0.06,0.08,0.1)
q_3	Fuel box isn't checked	An trapezoidal uncertain variable determined by quadruplet (0.02,0.03,0.04,0.05)
q_4	Tubing is blocked	0.01
q_5	Seal leakage	0.001
q_6	Piston can't move	0.042
q_7	Piston ring fault	0.01
q_8	Spark plug fault	An triangular uncertain variable determined by triplet (0.08,0.09,0.1)
q_9	Line fault	An triangular uncertain variable determined by triplet (0.04,0.05,0.06)
q_{10}	Magnet motor fault	0.02

This chapter describes analysis method of fault tree base on uncertainty measure with example in lecture [11].

The estimated value about chance that basic event of internal combustion engine's fault tree occurs is illustrated as Table 1,every bottom event in fault tree is described with probability and uncertain variable.

In order to educe objective function about chance that top event occurs, the following definition is introduced.

Definition 6("and"operator[2])

$$Q_{and} = q_1 \otimes q_2 \otimes \cdots q_n = \prod_{i=1}^{n} q_i \quad (1)$$

q_i is chance that event i occurs, is uncertain variable.

Definition 7 ("or"operator[2])

$$Q_{or} = q_1 \oplus q_2 \oplus \cdots q_n = 1 - \prod_{i=1}^{n} (1 - q_i) \quad (2)$$

q_i is chance that event i occurs, is uncertain variable.

The following expressions are obtained from Fig.1, if every basic event is absolute.

$$E = B \oplus C \oplus D \quad (3)$$

$$B = \text{event1} \oplus A \oplus \text{event4} \quad (4)$$

$$A = \text{event2} \otimes \text{event3} \quad (5)$$

$$C = \text{event5} \oplus \text{event6} \oplus \text{event7} \quad (6)$$

$$D = \text{event8} \oplus \text{event9} \oplus \text{event10} \quad (7)$$

The chance that the above bottom events occur are expressed with q_i, the chance value that actual event happens is introduced to formula (2) and (6)

$$\begin{aligned} C &= q_5 \oplus q_6 \oplus q_7 \\ &= 1 - (1 - q_5)(1 - q_6)(1 - q_7) \\ &= 0.053 \end{aligned} \quad (8)$$

It is obtained from formula (2) and (7)

$$\begin{aligned} D &= q_8 \oplus q_9 \oplus q_{10} \\ &= 1-(1-q_8)(1-q_9)(1-q_{10}) \\ &= 1-0.98(1-q_8)(1-q_9) \end{aligned} \quad (9)$$

It is obtained from formula (1) and (5)

$$A = q_2 \otimes q_3 = q_2 q_3 . \quad (10)$$

It is obtained from formula (2) and (4)

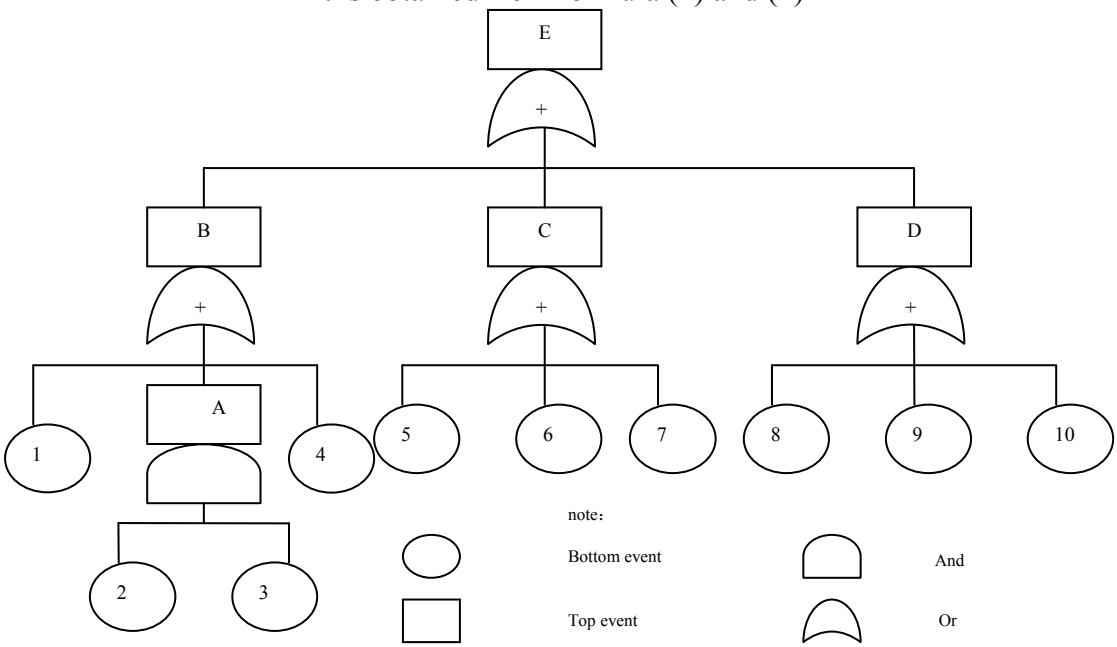


Fig. 1. Fault tree of internal combustion engine

$$\begin{aligned} B &= q_1 \oplus A \oplus q_4 \\ &= 1-(1-q_1)(1-A)(1-q_4) . \\ &= 1-0.97(1-q_2 q_3) \end{aligned} \quad (11)$$

The objective function of top event is obtained from (2)(3)(8)(9)(11) at last.[9],[10]

$$\begin{aligned} E &= B \oplus C \oplus D \\ &= 1-(1-B)(1-C)(1-D) \\ &= 1-0.91(1-q_2 q_3)(1-q_8)(1-q_9) \end{aligned} \quad (12)$$

The optimistic value of chance that top event of internal combustion engine's fault tree occurs is calculated according to uncertain simulation algorithm. When α changes from 0 to 1 by taking 0.1 as step, the relation between optimistic value of chance that top event occurs and α is shown in Figure 2. We can know that optimistic value of chance that top event occurs reduces with α increasing. We also discover that optimistic value of chance that top event occurs changes most quickly when F increases from 0 to 0.3, so it is the most effective to carry on the control at this stage. How to control ? Because the situation is very complex, we shall study and discuss about it in the future.

5 Conclusions

This paper proposes concept of uncertain fault tree and construction method. In this fault tree, it would be characterized as crisp number if fault rate of bottom event is obtained from reliable handbook, empirical data and so on; it would be characterized as uncertainty variable obtained from expert's subjective judgment if fault rate of bottom event has no statistical data. The optimistic value of overall system's top event is calculated by using uncertain simulation technology. Finally feasibility and validity of this method is confirmed by taking fault tree of internal combustion engine as example.

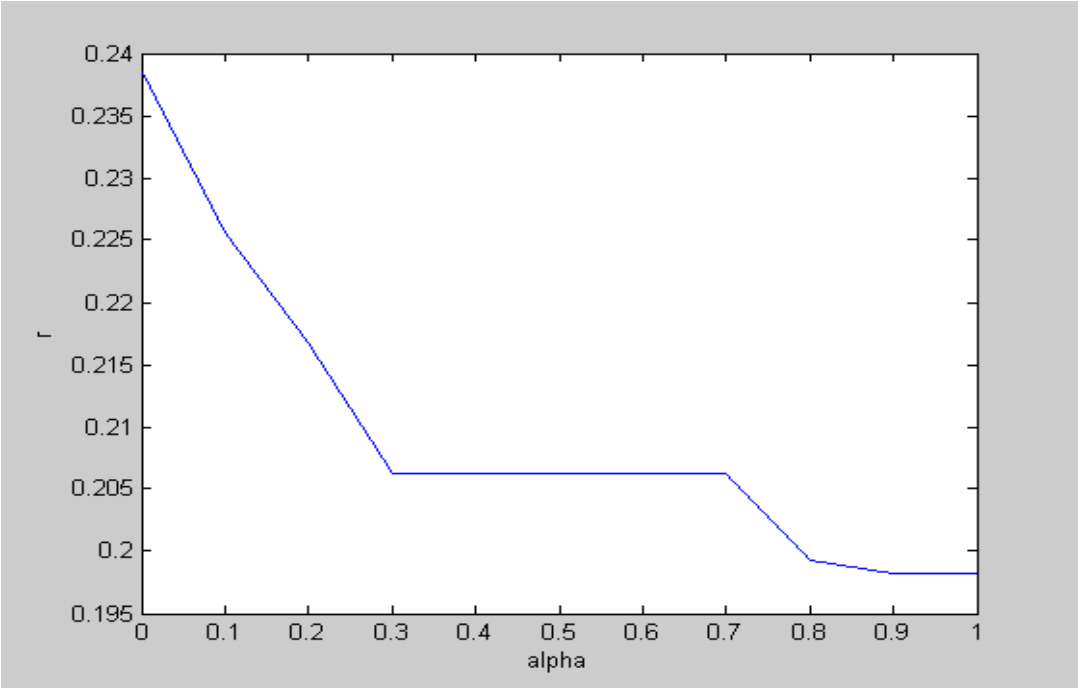


Fig. 2. Relation between top event's optimistic value and F

References

[1] Yuge Dong, Wenyu zhu, Xinzhao Chen: Analysis and application of fuzzy fault tree. submitted to Journal of HeFei Industrial University(1996)

[2] WuZhuang, ShiZhu, Xingui He:Analysis method and application base on fuzzy set. Submitted to Journal of Systems Engineering and Electronic Technology(2000)

[3] YongchuanWang, Wenxian Yu: Research of fuzzy tree analysis method based on fuzzy fault rate. Submitted to Journal of Systems Engineering Theory and Practice(2000)

[4] Xiaoyang Hua, Fuzzy fault tree analysis.(Shanghai Jiao Tong University press ,1995)

[5] Shizhu,Xingui He: Reliability analysis method based on arithmetic operation of fuzzy number: submitted to Journal of Systems Engineering and Electronic Technology(1999)

[6] Haiguang Chen,Xianyi Cheng: Analysis and application of random fuzzy fault tree: submitted to Journal of Computer Engineering and Application(2008)

[7] Liu Y K and Liu B：Random fuzzy programming with chance measures defined by fuzzy integrals: submitted to Journal of Mathematical and Computer Modelling(2002)

[8] B. Liu：Theory and Practice of Uncertain Programming, 2nd ed.(Springer-Verlag,Berlin, 2009)

[9] Baoding Liu, Ruiqing Zhao, WangGang: Uncertainty Programming and Application.(Tsinghua University press,China,2003)

[10] Zadeh L A: Fuzzy sets. Submitted to Journal of Information and Control(1965)

[11] Wenyu Zhu, Mechanism Reliability Design. (Shanghai JiaoTong University Press,China, 1992）

doi:10.4028/www.scientific.net/AMM.50-51.145

Research on E-commerce platform and modern logistics management system based on knowledge management platform

Cao Xu-hua[1], WANG Fu-zhong[2]

School of Economic and Management, Zhejiang University of science and technology,

Hangzhou 310023, China

[1]Morefast@126.com , [2]Wfz212126@126.com

Keywords: knowledge management; E-commerce platform; modern logistics management system

Abstract: For the stored data of E-commerce platform and modern logistics management system, which is very difficult to form knowledge's continuity and inheritance, about its researches on the related fields are insufficient, so the paper proposes an idea of constructing E-commerce platform and modern logistics management system based on knowledge management platform. Moreover, for a practical example, the paper introduces its application.

Introduction

Since Sveiby proposed the concept of knowledge management in 1986, the related researches about knowledge management have been growing in foreign and domestic. In the related researches, some researchers divided the life circle of knowledge management into four stages: knowledge creating, knowledge acquiring, knowledge organizing and knowledge spreading and sharing, every stage was also researched. Anjali Bakhru(2004) studied the effects that knowledge management putting on organization abilities[1]. Some researchers also studied the research progress, such as Miller, Ron. Econt (2005) summarized the research progress for knowledge management. Armour, Georgina (2005)reviewed and analyzed the research development trends[2]. Moreover, Liebowitz (1999), Fleenor, John (2005), Hou J.L(2005) also had studied the knowledge management for a long time, the research contents included many aspects existing in enterprises' knowledge management[3], [4].

The domestic researches about the knowledge management have been growing, the main research reasons are related to a great amount information and data. The knowledge management extracts the great amount data from large-sized database, by transferring and reorganizing them based on the decision topic to form a special data store mode, by further using datamining and knowledge discovery technology to manage and analyze the information, at last, the great economic benefits will be made.

Literatures reviewing and analyzing

General speaking, E-commerce is the producing, marketing, selling and circulating based on the electronic information network, which not only means the deal on the internet, bus also means the business activities by the information technology to solve the practical problems, such as: cutting down the cost, adding the value and creating the business opportunities. E-commerce includes a series of trade activities from the purchasing of raw material, the exhibiting of products, the ordering and storing of products and electronic paying and so on[5]. Moreover, Ma Qingguo(2000)

pointed out that the E-commerce based on the internet had a revolutionary significance, which was a main form of E-commerce[5]. Because E-commerce has a great competition advantage, more and more enterprises use it to rebuild the traditional methods and circulating channels, so the development of E-commerce produces great logistics demands and promotes the development of logistics industry(Dong Jiwen, 2004).

For the cognition of logistics, the scientific logistics theory began about several decades ago, after the development of many years, logistics has a great contribution to the society[6]. Nowadays, the batch data that the E-commerce platform and modern logistics management system store is independent, which is difficult to satisfy the demands of knowledge continuity and inheritance, but the knowledge management can make good use of its advantages to satisfy the demands. From the related research achievements, some scholars have done some jobs, which are shown as follows: Henry M. Kim(2000) modeled the sharing knowledge of an enterprise and studied the integration solving method of E-commerce[7], Martin, William J. (2001) studied the knowledge role content in E-commerce[8], Smirnov, Alexander el al.(2003)studied the knowledge source network configuring method in the knowledge logistics[9]. Cheung, C. F. el al.(2005)studied the inventory management based on knowledge in the production logistics[10]. From the foreign researches, there are many achievements related to E-commerce and knowledge management, logistics and knowledge management. However, the domestic research achievements are still scarce. According to the development of China's E-commerce and logistics industry, the domestic researches about them are very difficult to satisfy the demands of China's E-commerce and logistics industry.

Based on the above literatures reviewing and analyzing, this paper constructs the E-commerce platform and modern logistics management system based on knowledge management platform.

The research of E-commerce platform and modern logistics management system based on knowledge management platform

Research background

A cocoon and silk exchange market J has become the center of China's cocoon and silk exchange and price information, with the growing of exchange scale, some problems appeared and showed as follows: (1)Warehouses are insufficient. (2)The current information system of cocoon and silk exchange market J is difficult to satisfy the demands of international operation.

Based on these problems, the cocoon and silk exchange market J is urgent to use the advanced information technology to construct the international cocoon and silk E-commerce platform and its logistics system, whose aim is to enlarge the marketing and innovate the exchange functions, and improve the warehouses efficiency. Obviously, if the cocoon and silk exchange market J wants to have a good competition advantage, the market J must emphasize the core value of clients. In fact, there are many difficulties in front of market J: (1)Facing with more and more clients, how to guide them to communicate with the cocoon and silk exchange market J? (2)Because every client has a different business status, how to promote the business and risk management for it? (3)How to provide the best service to attract the clients and strengthen their loyalties? (4)How to cut down the cost under the condition of maintaining clients' loyalties?

In order to solve the above problems and difficulties, this paper proposes the construction plan of E-commerce based on knowledge management platform.

Knowledge management

The paper proposes a process figure about knowledge management, which is shown in figure 1:

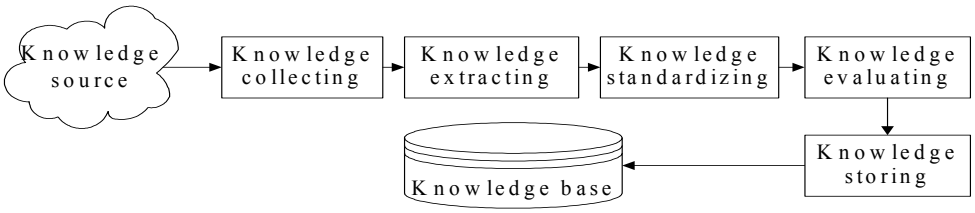


Figure 1. A process figure about knowledge management

In the process of knowledge management, the knowledge source means the information provided or released by competitors, clients, members, suppliers, after the knowledge collecting, knowledge should be extracted, and some useful information should be classified, coded, standardized and stored in the knowledge base at last. In figure 1, the knowledge evaluating includes the knowledge collecting and extracting performances and knowledge usefulness will be evaluated, which are very important for users who can select the useful knowledge.

The knowledge collecting, knowledge extracting, knowledge evaluating processes are shown in figure 2.

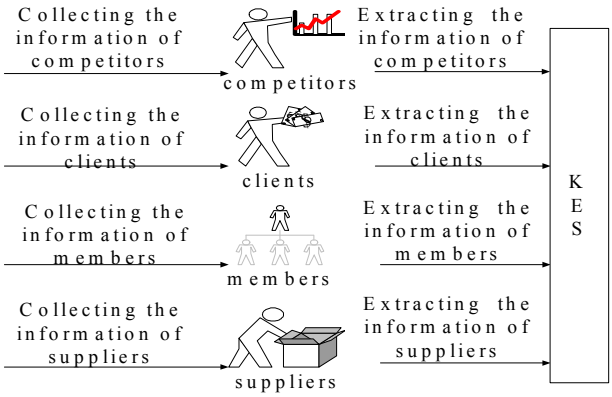


Figure 2. The knowledge collecting, knowledge extracting, knowledge evaluating processes

In figure 2, the KES means the knowledge evaluating system, after knowledge evaluating, some useful evaluation information can be produced.

The general structure of the E-commerce platform based on knowledge management platform

The paper constructs the general structure figure, which is shown in figure 3:

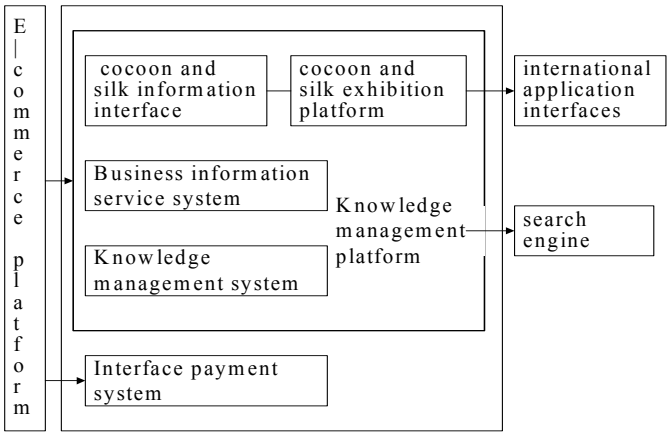


Figure 3. The E-commerce platform structure based on knowledge management platform

In order to realize the data transferring and sharing in the E-commerce platform, the design of the E-commerce platform uses the ebXML technology to satisfy the integration demands of

different data source. The core of the E-commerce platform based on ebXML is a knowledge management platform, who can satisfy the demands of knowledge continuity and inheritance, the international application interfaces can satisfy the demands of international business. The search engine can improve the searching speed of information, which can provide the convenience to users and improve the clients' satisfaction levels.

The system structure of E-commerce platform and modern logistics management system based on knowledge management platform

The system structure of E-commerce platform and modern logistics management system based on knowledge management platform is shown in figure 4.

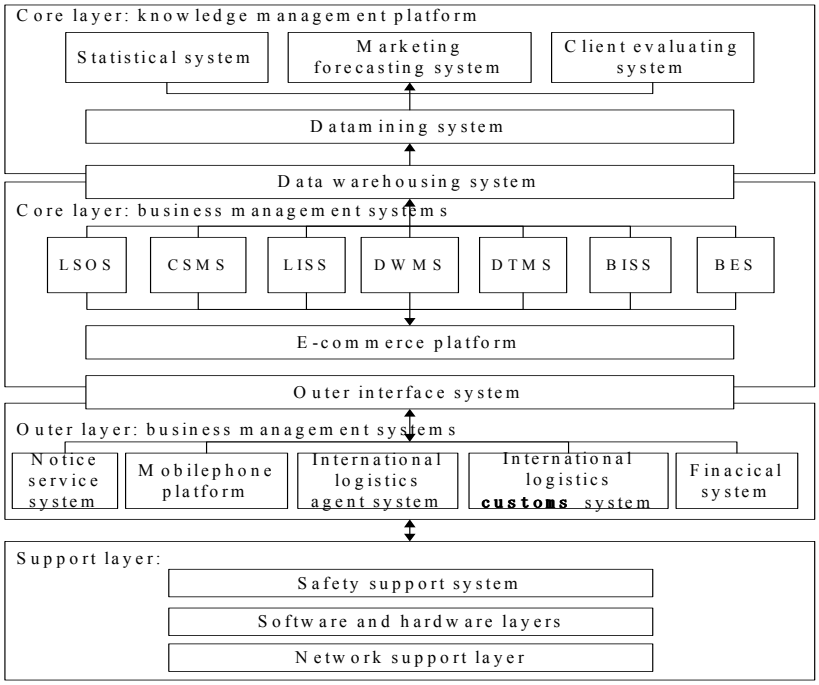


Figure 4. The system structure of E-commerce platform and modern logistics management system based on knowledge management platform

In the figure 4, the system structure includes the core layer, outer layer and support layer. The core layer includes knowledge management platform and business management systems, the outer layer includes business management systems and the support layer includes safety support system, software and hardware layers and network support layer. The knowledge management platform in the core layer includes a statistical system, marketing forecasting system and client evaluating system. The three systems run on a data warehousing system, by using the datamining technology to produce the statistical, forecasting and client evaluating information. The main functions of knowledge management platform are to provide the decision support for the core layer business management systems, the data warehousing system in the knowledge management platform plays a key role in the communications between the E-commerce platform and modern logistics management system. The business management systems in the core layer, which include a logistics scheduling optimization system(LSOS), client service management system(CSMS), logistics information service system(LISS), distributed warehousing management system(DWMS), distributed transportation management system(DTMS), business information service system(BISS) and business exchange system(BES). All systems in the business management system in the core layer run on the E-commerce platform. The business management systems in the outer layer, which includes a notice service system, mobilephone platform, international logistics freight agent system, international logistics custom system and financial system, these systems use the interface systems

to do the data exchanging and processing with the core layer business management systems. Especially for the mobilephone platform, the owners of vehicles and goods can query their transportation information on it.

Conclusion

For E-commerce platform and modern logistics management system, both of them have difficulties to form knowledge's continuity and inheritance for the great amount data, the paper is based on the advantages of knowledge management and uses it to construct the E-commerce platform and modern logistics management system. Firstly, the E-commerce platform structure based on knowledge management platform is constructed, secondly, the system structure of E-commerce platform and modern logistics management system based on knowledge management platform is also constructed. Especially for the system structure of E-commerce platform and modern logistics management system based on knowledge management platform, the paper divides it into three layers and analyzes them. According to the analysis for the system structure, obviously, the system structure has many advantages and can realize an effective coordination for the E-commerce platform and modern logistics management system.

The above construction idea has been applied to the cocoon and silk exchange market J, although many functional modules in the construction idea have been used in the market J, the practical economic effects still need time to verify.

Reference

[1]Anjali Bakhru. Managerial knowledge to organisational capability: new e-commerce businesses. Journal of Intellectual Capital,vol.5, no.2, pp.326-336, 2004

[2]Armour, Georgina. New Frontiers of Knowledge Management. KM Review, vol.8, no.3, pp.32-32, 2005.

[3]Liebowitz J.. Key ingredients to the success of an organization's management strategy. Knowledge and Process management,.vol.6, no.1, pp.37-40, 1999.

[4]Hou, J.L., Sun, M.T.,Chuo, H.C..An Intelligent Knowledge Management Model for Construction and Reuse of Automobile Manufacturing Intellectual Properties. International Journal of Advanced Manufacturing Technology, vol.26, no.1, pp.169-182, 2005.

[5]Ma Q.G. innovation of E-commerce and enterprise management theory. R&D Management.vol.12, no.5, pp.30-34, 2000.

[6]Wang Z.T. modern logistics management. Beijing: Publishing house of China's workers, 2001.

[7]Henry M. Kim. Enabling integrated decision making for electronic commerce by modelling an enterprise's sharable knowledge. Internet Research,vol.10,no.5,pp. 418-426, 2000.

[8]Martin, William J. The role of knowledge content in e-commerce. Journal of Information Science,vol.27, no.3, pp.180-185, 2001.

[9]Smirnov, Alexander; Pashkin, Mikhail; Chilov, Nikolai; Levashova, Tatiana; Haritatos, Fred. Knowledge source network configuration approach to knowledge logistics. International Journal of General Systems,vol.32, no.3, pp.251-270, 2003.

[10]Cheung, C. F.; Wang, W. M.; Kwok, S. K.,Knowledge-based inventory management in production logistics: a multi-agent approach. Proceedings of the Institution of Mechanical Engineers, Part B,vol.219, no.3, pp.299-307, 2005.

© (2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.150

Chance-Constrained Programming EOQ Model Based on Uncertain Measure

Lixia Rong[1, a]

[1]Computer Science and Technology Department of Dezhou University,

Shandong, Dezhou 253023, China

[a]lx_rong@163.com

Keywords: Uncertain variable, Uncertain distribution, Economic order quantity, Chance-constrained programming

Abstract. This paper provides chance-constrained programming EOQ (Economic Order Quantity) model for inventory based on uncertain measure. In the model, the ordering cost, shortage cost and holding cost per unit are assumed uncertain variables. When the parameters are some special uncertain variables with uncertainty distribution such as linear uncertain distribution and normal uncertain distribution, the model can be transformed into crisp equivalent programming. To illustrate the effectiveness of the model, some numerical examples are provided.

1 Introduction

Inventory theory is a branch of operational research. EOQ is one of fixed order quantity models, which determine the optimal selling period and order quantity. EOQ model was first proposed by Harris [1] in 1915, which gave the order quantity so as to meet customer service levels while minimizing the total inventory cost. There are many results about EOQ model from EOQ model was proposed.

A number of authors have studied the EOQ model when the parameters are random. Chen [2] developed an EOQ model under random demand. Liberatore [3] studied the EOQ model where the uncertainties in the lead time were represented stochastically.

In recent years, several researchers applied the fuzzy sets theory to EOQ models. Roy and Maiti [4] extended the classical EOQ model by introduced fuzziness both in the objective function and constraints of storage. Vujosevic [5] and Park [6] provided a fuzzy EOQ model by the fuzziness of ordering cost and holding cost. Chen and Wang [7] developed an EOQ model where the demand, ordering cost, holding cost and backorder cost are represented by trapezoidal fuzzy numbers. The fuzzy EOQ model with backorder was presented by Chang and Yao [8], where the backorder quantity was fuzzy. Ouyang and Yao [9] proposed a mixture inventory model with variable lead time, where demand is fuzzy variables. Production inventory problems were studied by Lee and Yao [10] where the demand and order quantity are represented by fuzzy sets. Yao and Chang [11] and Wang et al. [12] discussed inventory problems without backorder, where the order quantity and total demand quantity are represented by trapezoidal fuzzy numbers.

The above literatures dealt with the problem in stochastic or fuzzy environment. In many cases, the uncertainty behaves within the domain of neither randomness nor fuzziness. Most human decisions are made in the state of uncertainty. This fact provides a motivation to study the behavior of uncertain phenomena. In order to model uncertainty, Liu [13] founded an uncertainty theory that is a branch of mathematics based on normality, monotonicity, self-duality, countable subadditivity, and product measure axioms. In this paper, chance-constrained programming EOQ model for inventory is provided when the ordering cost, shortage cost and holding cost per unit are assumed uncertain variables. Then the optimum of model is solved when the parameters are some special uncertain variables with uncertainty distribution such as linear uncertain distribution and normal uncertain distribution.

2 Preliminaries

Let Γ be a nonempty set, and let A be a σ-algebra over Λ. Each $\Lambda \in \mathrm{A}$ is called an event. In order to provide an axiomatic definition of uncertain measure, it is necessary to assign to each event Λ a number $\mathrm{M}\{\Lambda\}$, which indicates the level that Λ will occur. In order to ensure that the number $\mathrm{M}\{\Lambda\}$ has uncertain mathematical properties, Liu [13] proposed the following four axioms:

Axiom 1 (Normality) $\mathrm{M}\{\Gamma\} = 1$.

Axiom 2 (Monotonicity) $\mathrm{M}\{\Lambda_2\} \le \mathrm{M}\{\Lambda_2\}$ whenever $\Lambda_1 \subset \Lambda_2$.

Axiom 3 (Self-Duality) $\mathrm{M}\{\Lambda\} + \mathrm{M}\{\Lambda^c\} = 1$ for any event Λ.

Axiom 4 (Countable Subadditivity) For every countable sequence of events Λ_i, we have

$$\mathrm{M}\{\bigcup_{i=1}^{\infty} \Lambda_i\} \le \sum_{i=1}^{\infty} \mathrm{M}\{\Lambda_i\} \tag{1}$$

Definition 1 (Liu [13]) The set function M is called an uncertain measure if it satisfies the normality, monotonicity, self-duality, and countable subadditivity axioms.

It is known that a random variable can be characterized by a probability density function and a fuzzy variable may be described by a membership function. Similarly, an uncertain variable may be described by uncertain distribution.

Definition 2 (Liu [13]) The uncertainty distribution Φ of an uncertain variable ξ is defined by $\Phi(x) = \mathrm{M}\{\xi \le x\}$ for any real number ξ.

Definition 3 (Liu [13]) An uncertain variable ξ is called linear if it has a linear uncertainty distribution

$$\Phi(x) = \begin{cases} 0, & if \quad x \le a \\ (x-a)/(b-a), & if \quad a \le x \le b \\ 1, & if \quad x \ge b \end{cases} \tag{2}$$

denoted by $L(a,b)$,where a and b are real numbers with $a < b$.

Definition 4 (Liu [13]) An uncertain variable ξ is called normal if it has a normal uncertainty distribution

$$\Phi(x) = (1 + \exp(\frac{\pi(e-x)}{\sqrt{3}\sigma}))^{-1}, x \in R \tag{3}$$

denoted by $\mathrm{N}(e,\sigma)$ where e and σ are real numbers with $\sigma > 0$.

Definition 5 (Liu [13]) Let ξ be an uncertain variable, and $\alpha \in (0,1)$. Then

$$\xi_{\inf}(\alpha) = \inf\{r \mid \mathrm{M}\{\xi \le r\} \ge \alpha\} \tag{4}$$

is called the α-pessimistic value to ξ.

3 Chance-Constraint Programming EOQ Model

In order to describe and analyze the problem, we suppose that the time of supplement is longer and shortage is permitted. First, the notations used in the paper are listed below.

R : demand per unit time;

P : product number per unit time;

t : selling period, a decision variable;

t_1: the time of product starting;
t_2: the time of shortage is zero;
t_3: the time of product ending;
ξ_1: holding cost per unit, a uncertain variable;
ξ_2: shortage cost per unit, a uncertain variable;
ξ_3: ordering cost per unit, a uncertain variable;
Φ_i: uncertainty distribution of ξ_i,i=1,2,3;
$f(\xi_1,\xi_2,\xi_3,t)$: the total cost per period.
Then the EOQ model is given by

$$f(\xi_1,\xi_2,\xi_3,t,t_2)=\frac{(P-R)R}{2P}[\xi_1(t_1-2t_2)+(\xi_1+\xi_2)\frac{t_2^2}{t}]+\frac{\xi_3}{t} \tag{5}$$

Since the ordering cost, shortage cost and holding cost per unit are uncertain variables, the total cost is also an uncertain variable. The target of EOQ model is to find the optimal selling period and order quantity such that the total cost is minimal. Assume that we believe the pessimistic value criterion. If we want to maximize the pessimistic value subject to some chance constraints, the chance-constrained programming is as follows:

$$\begin{cases}\min \bar{f} \\ st: \\ \mathrm{M}\{f(\xi_1,\xi_2,\xi_3,t,t_2)\le \bar{f}\}\ge\alpha \\ t>0\end{cases} \tag{6}$$

where α is specified confidence level, and $\min \bar{f}$ is the α-pessimistic return.

4 Model Solutions

The traditional solution methods require conversion of the objective function and the constraints to their respective deterministic equivalents. As we know, this process is usually hard to perform and only successful for some special cases. Let us consider the following forms of the uncertain ordering cost, shortage cost and holding cost per unit.

Case 1. Suppose that the holding cost, shortage cost and ordering cost per unit are uncertain variables $\xi_i \sim L(a_i,b_i), i=1,2,3$, then $f(\xi_1,\xi_2,\xi_3,t,t_2)$ is uncertain variable. In accordance with the properties of uncertain variable with linear uncertainty distribution, the deterministic equivalent of Eq.6 is as follows.

$$\begin{cases}\min \frac{(P-R)P}{4P}\{[(1-\alpha)a_1+\alpha b_1](t_1-2t_2)+[(1-\alpha)(a_1+a_2)+\alpha(b_1+b_2)]\frac{t_2^2}{t}\}+\frac{(1-\alpha)a_3+\alpha b_3}{t} \\ st: \\ t,t_2>0\end{cases} \tag{7}$$

Then the optimal selling period and order quantity are as follows.

$$t^*=\sqrt{\frac{2[(1-\alpha)a_3+\alpha b_3]}{R[(1-\alpha)a_1+\alpha b_1]}}\sqrt{\frac{2[(1-\alpha)(a_1+a_2)+\alpha(b_1+b_2)]}{(1-\alpha)a_2+\alpha b_2}}\sqrt{\frac{P}{P-R}} \tag{8}$$

$$Q^* = Rt^* = \sqrt{\frac{2R[(1-\alpha)a_3 + \alpha b_3]}{(1-\alpha)a_1 + \alpha b_1}} \sqrt{\frac{2[(1-\alpha)(a_1+a_2) + \alpha(b_1+b_2)]}{(1-\alpha)a_2 + \alpha b_2}} \sqrt{\frac{P}{P-R}} \tag{9}$$

Case 2. Suppose that the holding cost, shortage cost and ordering cost per unit are uncertain variables $\xi_i \sim N(e_i, \sigma_i), i = 1,2,3$, where e_i and σ_i are real number with $\sigma_i > 0$. Then we have

$$\begin{cases} \min \dfrac{(P-R)P}{4P}[(e_1 + \dfrac{\sigma_1\sqrt{3}}{\pi}\ln\dfrac{\alpha}{1-\alpha})(t_1 - 2t_2) + (e_1 + e_2 + \dfrac{(\sigma_1+\sigma_2)\sqrt{3}}{\pi}\ln\dfrac{\alpha}{1-\alpha})\dfrac{t_2^2}{t}] + \dfrac{e_3 + \dfrac{\sigma_3\sqrt{3}}{\pi}\ln\dfrac{\alpha}{1-\alpha}}{t} \\ st: \\ t, t_2 > 0 \end{cases} \tag{10}$$

Then the optimal selling period and order quantity are as follows.

$$t^* = \sqrt{\frac{2(e_3 + \dfrac{\sigma_3\sqrt{3}}{\pi}\ln\dfrac{\alpha}{1-\alpha})}{R(e_1 + \dfrac{\sigma_1\sqrt{3}}{\pi}\ln\dfrac{\alpha}{1-\alpha})}} \sqrt{\frac{e_1 + e_2 + \dfrac{(\sigma_1+\sigma_2)\sqrt{3}}{\pi}\ln\dfrac{\alpha}{1-\alpha}}{e_2 + \dfrac{\sigma_2\sqrt{3}}{\pi}\ln\dfrac{\alpha}{1-\alpha}}} \sqrt{\frac{P}{P-R}} \tag{11}$$

$$Q^* = Rt^* = \sqrt{\frac{2(e_3 + \dfrac{\sigma_3\sqrt{3}}{\pi}\ln\dfrac{\alpha}{1-\alpha})}{R(e_1 + \dfrac{\sigma_1\sqrt{3}}{\pi}\ln\dfrac{\alpha}{1-\alpha})}} \sqrt{\frac{e_1 + e_2 + \dfrac{(\sigma_1+\sigma_2)\sqrt{3}}{\pi}\ln\dfrac{\alpha}{1-\alpha}}{e_2 + \dfrac{\sigma_2\sqrt{3}}{\pi}\ln\dfrac{\alpha}{1-\alpha}}} \sqrt{\frac{P}{P-R}} \tag{12}$$

5 Numerical Examples

Assume that the demand per unit time is R=10, the product number per unit time is P=15, the specified confidence level is α =0.8.

Example 1. If the holding cost, shortage cost and ordering cost and unit per unit time are uncertain variables with linear uncertainty distribution. Let them be $\xi_1 \sim L(3,5)$, $\xi_2 \sim L(5,7)$, $\xi_3 \sim L(2,4)$. Then, we have

$$t^* = \sqrt{\frac{2(0.2\times 2 + 0.8\times 4)}{10(0.2\times 3 + 0.8\times 5]}} \sqrt{\frac{2(0.2\times 8 + 0.8\times 12)}{0.2\times 5 + 0.8\times 7}} \sqrt{\frac{15}{15-10}} = 1.26$$

$$Q^* = Rt^* = 10\times 1.26 = 12.6$$

Example 2. Assume that the holding cost, shortage cost and ordering cost per unit time are normal uncertain variables. Let them be $\xi_1 \sim N(4,6)$, $\xi_2 \sim N(5,7)$, $\xi_3 \sim N(3,5)$. Then, we have

$$t^* = \sqrt{\frac{2(3 + \dfrac{5\sqrt{3}}{\pi}\ln\dfrac{0.8}{1-0.8})}{10(4 + \dfrac{6\sqrt{3}}{\pi}\ln\dfrac{0.8}{1-0.8})}} \sqrt{\frac{4+5 + \dfrac{(6+7)\sqrt{3}}{\pi}\ln\dfrac{0.8}{1-0.8}}{5 + \dfrac{7\sqrt{3}}{\pi}\ln\dfrac{0.8}{1-0.8}}} \sqrt{\frac{15}{15-10}} = 0.94$$

$$Q^* = Rt^* = 10\times 0.94 = 9.4$$

6 Conclusions

In this paper, chance-constrained programming model for inventory is provided when the ordering cost shortage cost and holding cost are assumed to uncertain variables. In order to solve the proposed model by traditional methods, we studied the crisp equivalents when the ordering cost shortage cost and holding cost are some special uncertain variables. Finally, some numerical examples are given to verify the effectiveness of the model.

Acknowledgements

This paper was supported by Shandong Provincial Social Science Programming Research Project, No. 09CJGJ07, Shandong Provincial Scientific and Technological Research Plan Project, No. 2009GG20001029 and Scientific Research Development Plan Project of Shandong Provincial Education Department, No. J08LJ54.

The author would like to thank the anonymous reviewers and the editor-in-chief for their valuable comments, Dr. Yufu Ning for his help in the revision of the manuscript.

References

[1] Harris, F.W.: Operation and costs, The library of Factory Management, Vol. V.A.W shaw company, Chicago(1915), p. 47

[2] Chen Y.C.: A probabilistic approach for traditional EOQ model, Journal of Information Optimization Sciences, Vol. 24 (2003), p. 249

[3] Liberatore M.J.: The EOQ model under stochastic lead time, Operations Research, Vol. 27(1979), p. 391

[4] Roy T.K., Maiti M: A fuzzy EOQ model with demand dependent unit cost under limited storage capacity, European Journal of Operational Research, Vol. 99 (1997), p. 425

[5] Vujosevic M, Petrovic D, Petrovic R: EOQ formula when inventory cost is fuzzy. Inventory Journal of Production Economics, Vol. 45 (1996), p. 499

[6] Park K.S.: Fuzzy set theoretic interpretation of economic order quantity. IEEE Transactions on Systems, Man and Cybemieics, Vol. 17 (1987), p. 1080

[7] Chen S.H., Wang C.C.: Backorder fuzzy inventory model under functional principle. Information Sciences, Vol. 95 (1996), p. 71

[8] Chang S.C., Yao J.S., Lee H.M.: Economic reorder point for fuzzy backorder quantity. European Journal of Operational Research, Vol. 109 (1998), p. 183

[9] Ouyang L.Y., Yao J.S.: A minimal distribution free procedure for mixed inventory model involving variable lead time with fuzzy demand. Computers and Operations Research, Vol. 29 (2002), p. 471

[10] Lee H.M., Yao J.S.: Economic production quantity for fuzzy demand quantity and fuzzy production quantity. European Journal of Operational Research, Vol.109 (1998), p. 203

[11] Yao J.S., Chang S.C., Su J.S.: Fuzzy inventory without backorder for fuzzy order quantity and fuzzy total demand quantity. Computers and Operations Research, Vol. 27 (2000), p. 935

[12] Wang X.B., Tang W.S., Zhao R.Q.: Fuzzy economic order quantity inventory models without backordering, Tsinghua Science and Technology, Vol. 21 (2007), p. 91

[13] Liu B. Uncertainty theory, 3rd edition on http://orsc.edu.cn/lliu/ut.pdf.

doi:10.4028/www.scientific.net/AMM.50-51.155

Analysis on Ecological Principles in Macau Urban Design

WU Yao[1,a], Vizeu Pinheiro Francisco[2,b]

[1]School of Design, Jiangnan University, Wuxi, China

[2]Macao Civic and Municipal Affairs Bureau (IACM), Macau, China

[a] wuyaoseu@yahoo.com.cn , [b] francisco.vizeupinheiro@gmail.com

Keywords: Macau, ecology, principles

Abstract: Macao managed for over 400 years to develop Iberian characteristics in the city together with the Chinese characteristics; it has formed an organic urban fabric reflecting ecological principles original from different places, that could easily to be adopted under tropic circumstance, The study focus on the researches of how these elemental principles should be, and how to develop and maintain those ecological principles. Survey was done on existing data from three different sources, namely 1) graphic images and maps, 2) documents archives, 3) archeological findings. The paper contributed for the knowledge would not only helpful to Macau modern urban renewal, but also benefit to relevant projects on ecological city plan.

1. Preface

Macau is a small city located in the southern of China, which was found as Portuguese colony from the moment of 1557. Under such development background, Macau urban evolution was influenced by Sino-Latin culture. In 2005 Macau historical architecture group was listed into world cultural heritage by UNESCO, the precious architectural treasure reflects some ecological principles in construction methods and decoration details, it could be useful to study and restore carefully in Macau urban modern renewal to fulfill the sustainable development. (Refer with Fig.1.)

Fig.1. Macau urban imagine in 17 century. Picture from Macau Archive Office.

Macau urban fabric and architectural form not only reflect the influences by different cultures, but also show us some vernacular ecological principles which original from Chinese and Portuguese tradition. In general, the ecological principles are important rules in Macau city evolution throughout almost 500 years, it's still valuable in urban modern renewal process nowadays.

2. Design combined with nature

In the early time, Macau urban fabric have some characters of irregular and organic forms which are similar with medieval cities in European, especially in the evolution background, urban tissues, and the cities meanings; this urban fabric has been evolved and adapted for long time, because "comparing with regular square grid urban road planning, it's more efficient and reasonable." Some of considerable features in urban planning include geomorphology, economy, living condition, ideology, military defense, etc.

Macau old city center locates in sloped highland; obviously, the road network is influenced by the terrain. The main road is planed along with contour line and shore line, for examples: *Rua da Pedra --- Rua da Palmeria --- Calcada das Sortes*along with the *Sali mountain*; *Rua de Santo Antonio --- Rua de S.Paulo ---Rua da Palha --- Rua da S.Domingos*along with *Portaleza Do Monte,* furthermore, *Rua da Praia --- Rua da Barao --- Rua da Alfandega --- Rua das Mercadores --- Rua*

dos Ervanarios --- *Rua das Colonos* along with Macau western shore line to build. All those curved organic roads form the framework of urban transportation network.

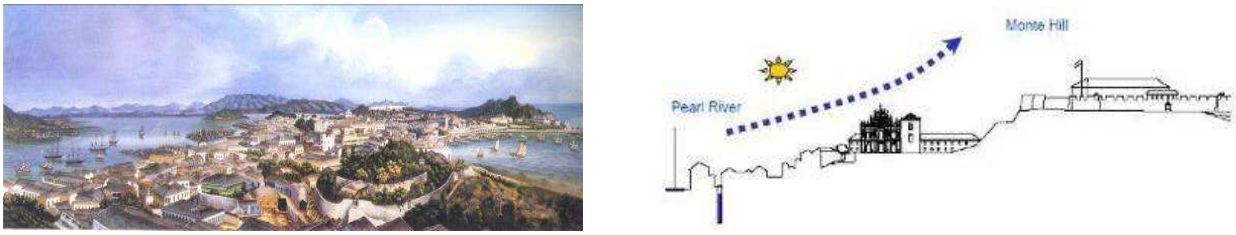


Fig.2. Macau city center has good relationship with nature.
Picture from Macau archive office and redrew by Pinheiro. F.V.

In the network pattern, the road paved along with terrain, connecting with some open spaces, then creating a good urban fabric and ecological environments. Those buildings located in different levels; therefore, there have good ventilation and sunshine individually. Furthermore, using local materials and traditional construction mythology is another good way to harmonized with environment. This is the "simple ecological principle" which defined by the famous modernism architect frank Wright[1], that is also emphasized by another famous architect Le Corbusier, namely "initial ecological principle"[2]. (Refer with Fig.2.)

3. Utilization of open space

Macau historical center is narrow sharp which includes complex function, for examples: residential building, market, administration building, religion building, and military construction. Those buildings arranged along with the "main street" to connect with different open spaces which could contain different urban function. Normally, buildings create a boundary of a community, the interior is backyard opposite against the street, if necessary, and the new extension will be built in this area. "At first, the residential building was two or three levels, continuously arranged along with circumambience of backyard, sometimes the interior contains a yard touching with main building…, this kind building continued along with whole community." "The building is not only for living but also for workshop, store, and counting house…, " this urban formation reflects outside activity intention in medieval ages, but also, the continuing building elevation provide integrated and beautiful façade of the street.

This is an efficient urban pattern in which of utilizing spatial resources in limited area, combining with different functions to create a harmony community. (Refer with Fig.3.)

1. Wright thought that architecture is a living organic, his design has intention to create a good relationship with nature, furthermore to integrate his designed building with nature properly. Wright insisted that design is dynamic process, no architecture is totally finished. It influences ambience and customers all the time. In his book *"The Natural House"* he emphasized that architectural integrity is so important, which should integrate building location, materials, and customer daily living.
2. In Le Corbusier book *"Towards a New Architecture"*, he loved outdoor living and sports, therefore, those two living features became characters in his design works, one is focus on simple living and freedom, the other is paying attention with nature. In 1922 he designed the *"real estate building",* it shows us the simple and thrift, every department has individual garden, it's easy to feel the nature. Look through the garden, the customer have good views of woods and greens, and touch carefully with nature.

Fig.3. the open space play important role in urban function, most of them have efficient multi-function.
Picture from IACM and redrew by WU Yao

4. Principles in architecture design

Basically, Macau local building could divide into Chinese style, western style, and hybridized style which parts original from china or/and other parts from western countries. Some architecture is good at spatial utilization, construction, details, all those methods could well suit with local climate, especially in hot summer, there often has good nature ventilation and lighting, and some principles still go into effect in modern architecture design.

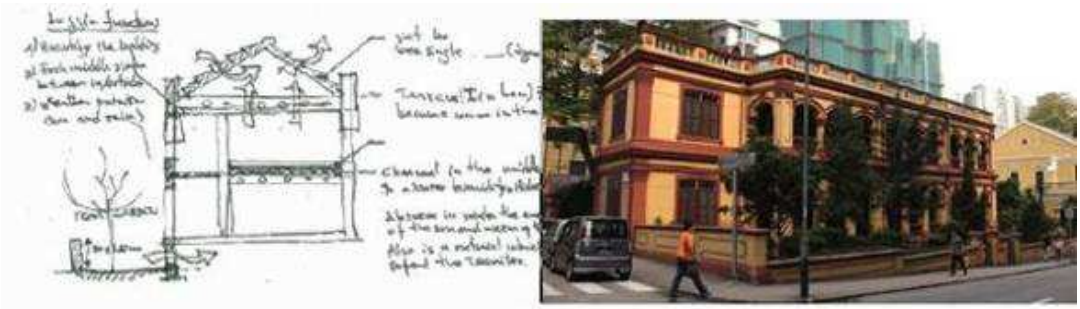


Fig.4. Some traditional ecological methods in western style building --- corridors, stilts, sun-shutter, small garden, etc.
Picture took and redrew by Pinheiro. F. V.

The western style villa located in *aveniad do conselheiro Ferreira de almeida*,which have adaptive changes from western building merging into local, and has been identified as "Macanese colonial style". (Refer with Fig.4.)

The building as above is classical Renaissance style, which have characters of good proportion, rich details, using arch and classical architectural order. In first and second level, there have corridors that are useful to protect sunshine in hot climate, meanwhile, it built on stilts to enhance the ventilation, the front garden provide visual and psychological buffer zone to connect with Noise Street, which created peaceful and cheerful environment around the whole building. Some building details and interior decoration were also adjusted by the local hot climate, for examples, Portuguese shutter windows, flower pattern pierced ceiling, etc.

Some Chinese traditional residential buildings are also good example for the ecological principles. For example *The Lu Mansion* located near the cathedral square3, using the atrium and lofty storey height for the ventilation, and all windows open to the atrium, the shadow can cool the fresh air then flow into the interior. (Refer with Fig.5.)

3. *The Lu Mansion* near cathedral square has southern Chinese style, which was built by wealth trader Lujiu in 1889, the building was divide into 20 parts in 1910, in 1960s Macau government restored this building and finished it in 2004, then be listed into world cultural heritage in 2005. After restoration it became workshop or exhibition hall of folk custom. Because of small volumes（336M2）, limited visitor could allowed come in at same time, at second floor 10 people allowed stay at same time.

Fig.5. Some traditional ecological methods in Chinese style building ---narrow atrium, lofty storey height, etc.
Picture took and drew by Pinheiro. F.V.

5. Some issues in Macau urban renewal

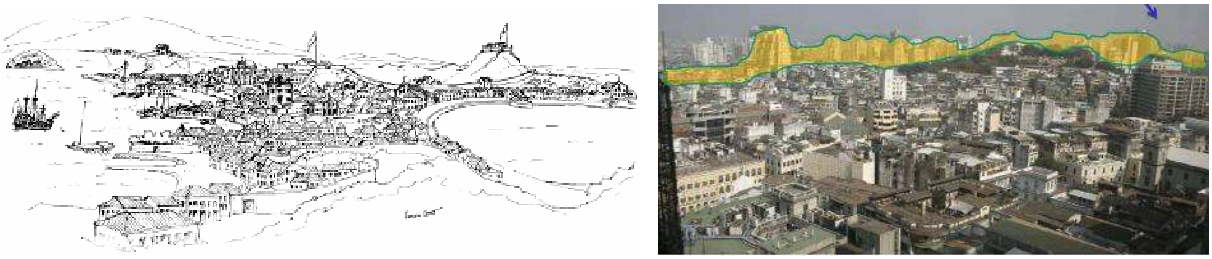


Fig.6. Restored imagine of Macau in the early of 20 century, nowadays the peninsular is around by high-rise building belt.
Picture took and redrew by Pinheiro, F.V

In the end 70s and 80s of last century, great many immigrant came into Macau, so real estate has tremendous development, therefore, Macau traditional urban fabric and construction model were totally changed. The land reclamation formed new out harbor, in this new area high-rise was built for office and residence. As consequence, it create a giant high-rise wall to block oceanity seasonal winds to flow into the interior of peninsular, the favorable ventilation corridor of peninsular is broken. It is difficult to eliminate pollution; the ecological urban fabric was changed. (Refer with Fig.6.)

Moreover, because of over-developed and lack of general planning, some ugly contemporary buildings without culture characters were constructed in the historical area; obviously it's harmful to the heritages cultural and ecological values. *The Mandarin House* [4] is just under such situation and need to be rehabilitated urgently. (Refer with Fig.7.)

Fig.7. The real model of mandarin house and ambience.
Picture took by Pinheiro, F.V.

4 *The mandarin house* was built in 1887, the owner is famous thinker Zheng Guanyin in Qing Dynasty, whose thoughts have deeply influenced Emperor Guangxu, Sun Zhongshan, Mao Zedong,. In his well-know book "warnings to a prosperous ages" (1893), he suggested to learn western countries advanced technology, ideology, science, economy, culture, foreign policy to prompt Chinese modernism. This building was built by his father, in 1893 Zheng Guanyin father died, then the building divided into 3 parts for three sons.

6. Ecological methods in urban renewal.

After consideration in Macau traditional fabric and construction method, there are some ecological principles which could adopted in the urban renewal as bellows.

1. Reducing the high-rise construction and keeping the low building density around the shore side belt to maintain the ventilation corridor throughout peninsular.

2. Taking carefully consideration on using of the vernacular ecological methods in the new construction to reduce the using of non-renewable resources, for example, using the natural ventilation, sun-shading, etc.

3. Urban green coverage is one of important features of ecological characters, because of big population and high density in the city center, the green area is not enough to the public. A potential way to increasing urban green coverage is rooftop virescence; owner should get some compensation by government for using the rooftop area.

ACKNOWLEDGMENT

1.Supported by" The Fundamental Research Funds for the Central Universities" No.JUSRP21015
2. Supported by "The Fund of Graduate Education Reform and Practice Project in Jiangsu Province in 2009"
3. Supported by "The Fund of Jiangnan University National Undergraduate Creative Practice Project in 2009".

REFERENCES

[1] Pinheiro, F.V., Yagi, K. And Korenaga, M., *"Urban De-fragmentation and Infrastructure Role of Open Public Spaces in the City Historical Center: Case Studies of Macao's Urban Rehabilitation"*, paper presented at the Third Symposium on Sustainable Development of Guangdong, Hong Kong and Macau, November 23-26 (2006), p.180.

[2] Lung, D.,*"Introduction: The Future of Macao Past"*, paper presented at the International Conference on the Conservation of Urban Heritage: Macao Vision, 10-12 September 2002, organized by the Cultural Institute of the Macao SAR Government, Xiii- XXi.

[3] Stephen Chen, *"The Culture Combination of Chinese and Western Countries in Macau",* Culture Magazine(Autumn of 2003), p.137-141.

[4] Clark, F. *"Planning for the Past: heritage services in local planning authorities in England"*, Cultural Trends, 43-44 (2001), p.63-94.

[5] Cohen, E. *"Authenticity and commoditization in tourism",* Annuals of Tourism Research,15 (1998), p.371-386.

[6] Wong Shiu Kwan. *"MACAO ARCHITECTURE ： AN INTEGRATE OF CHINESE AND PORTUGUESE",* Review of Architerture，Macau: Instituto Cultural of Macau (1998), p.30-39.

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.160

Analyzing the vibration system with time-varying mass

Yan Zhu [1,a], Shulin Wang [2,b]

[1]Department of Automobile Engineering, University of Military Transportation, TianJin 300161, China

[2] School of Materials Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

[a] zylzl@163.com , [b] wangshul@online.sh.cn

Keywords: vibration analysis; vibration with time–varying mass; parametric excited vibration; vibration design

Abstract: Considering that impact motion is not only a function of time but also dependent on phase-angle, we suppose the non-linear response of the vibro-impact system with time-varying mass is a function dependent not only on different time scale but also on the phase parameter. An approximate analytical solution of second order for the vibration is obtained by using Multi-Scales Method. The feasibility is verified by the numerical solution using Runge-Kutta algorithms. It is shown that the motion of the variable mass system has periodic behavior with the period of the changing mass, and the mass variation can only influence the system amplitude but not its cycle. However, the bigger the mass factor varies, the more intensive the response enlarges, and vice versa. The method and findings may be useful to analyze similar vibration systems with impact dampers or design the vibration control strategy.

Introduction

The differential equation of motion can be described as

$$\frac{d}{dt}[m(t)\frac{dx}{dt}]+c\frac{dx}{dt}+kx=F_0\sin\omega t \tag{1}$$

Where $m(t)$ is the mass of vibration system； c is damping parameter； k is stiffness coefficient； ω is excited frequency； F_0 is the amplitude of exciting force; x is the displacement of vibration system; t is time.

The mass of the vibration system is a time-varying function as discussed by Wang [1]. It can be written as:

$$m(t)=m_1[1+\varepsilon\sin(\omega_a t+\varphi)] \tag{2}$$

Where m_1 is the constant of mass, ε is the variation coefficient of time-varying mass and is a small parameter, φ is a constant, ω_a is the frequency of the time-varying mass, relating to the forcing frequency ω as $\omega_a=\omega/n$ with n is an integer.

Vibration system featuring time-varying mass can be found widely in industry such as chemical process, dispersive particle system, transportation, civil engineering, celestial mechanics, etc.[2]. Many approximate procedures have been developed to find solutions for second order nonlinear differential equations, such as the Multiple-Scales Method [3], the Krylov-Bogolubov method [4,5], and the method of harmonic balance [6], etc. These methods are extended for solving the second order nonlinear differential equation with complex function and constant parameters. But using these methods one may encounter difficulties of solving variable coefficient differential equations like Eq. 1.

To solve these problems, we suppose that the non-linear response of the vibro-impact system is a function dependent not only on different time scale but also on the phase parameter, considering that impact motion is not only a function of time but also of spatiality. By using Multi-Scales Method we obtained an approximate analytical solution of second order for the vibration. The feasibility is verified by the approach using Runge-Kutta numerical algorithms. The periodic behavior and the influence of mass changing on the system response are discussed in some detail according to the approximate solutions proposed.

1. Approach by The Multi-Scales Method

The differentiation of Eq.2 with respect to time gives $m'(t)=\varepsilon m'_0(t)$, Where $(')=d/dt$ and $m_0'(t) = m_1[\sin(\omega_a t + \varphi)]'$.

Letting $c/m(t)=2\varepsilon\zeta\omega_0(t)$, and $F_0=\varepsilon F$, we can re-written Eq.1 as:

$$\frac{d^2x}{dt^2} + \omega_0^2(t)x = -\varepsilon[2\zeta\omega_0(t) + \frac{m_0'(t)}{m(t)}]\frac{dx}{dt} + \varepsilon\frac{F}{m(t)}\sin\omega t \tag{3}$$

Eq. 3 can be simplified as the general form of parametric excitation:

$$\frac{d^2x}{dt^2} + \omega_0^2(t)x = \varepsilon f_0(\frac{dx}{dt}, \omega t) \tag{4}$$

Where $\omega_0^2(t) = k/m(t)$.

The analysis is based on the procedures of Multiple-Scales Method. However, we express the solution as a function of both different time scale and phase parameter, instead only introducing different time scale as usual. That is

$$x = x_0(\psi, T_1, T_2, \cdots) + \varepsilon x_1(\psi, T_1, T_2, \cdots) + \varepsilon^2 x_2(\psi, T_1, T_2, \cdots) + \cdots \tag{5}$$

Where T_n is a time variable, ψ is a phase variable, and $T_n=\varepsilon^n t$ for n=1,2,3…, $d\psi/dt=\omega_0(t)$. Because x is a function of phase and time with multi-scale, we have:

$$\frac{dx}{dt} = \omega_0\frac{\partial x_0}{\partial\psi} + \varepsilon(\frac{\partial x_0}{\partial T_1} + \omega_0\frac{\partial x_1}{\partial\psi}) + \varepsilon^2(\frac{\partial x_0}{\partial T_2} + \frac{\partial x_1}{\partial T_1} + \omega_0\frac{\partial x_2}{\partial\psi}) + \cdots \tag{6a}$$

$$\frac{d^2x}{dt^2} = \omega_0^2\frac{\partial^2 x_0}{\partial\psi^2} + \varepsilon[-\frac{\omega_0}{2}\cdot\frac{m_0'}{m}\cdot\frac{\partial x_0}{\partial\psi} + 2\omega_0\frac{\partial^2 x_0}{\partial\psi\partial T_1} + \omega_0^2\frac{\partial^2 x_1}{\partial\psi^2}] + \varepsilon^2[2\omega_0\frac{\partial^2 x_0}{\partial T_2\partial\psi} + \frac{\partial^2 x_0}{\partial T_1^2} - \frac{\omega_0}{2}\cdot\frac{m_0'}{m}\cdot\frac{\partial x_1}{\partial\psi} + 2\omega_0\frac{\partial^2 x_1}{\partial T_1\partial\psi} + \omega_0^2\frac{\partial^2 x_2}{\partial\psi^2}] + \cdots \tag{6b}$$

To simplify the calculation, we will find an approximate solution to second order. Substituting Eq.5, 6a and 6b into Eq.3 and respectively equating the coefficients of ε^0, ε^1 and ε^2 to zero, we obtain:

$$\varepsilon^0: \quad \omega_0^2\frac{\partial^2 x_0}{\partial\psi^2} + \omega_0^2 x_0 = 0 \tag{7a}$$

$$\varepsilon^1: \quad \omega_0^2\frac{\partial^2 x_1}{\partial\psi^2} + \omega_0^2 x_1 = (\frac{\omega_0}{2}\cdot\frac{m_0'}{m})\frac{\partial x_0}{\partial\psi} - 2\omega_0\cdot\frac{\partial^2 x_0}{\partial\psi\partial T_1} - (2\zeta\omega_0 + \frac{m_0'}{m})\omega_0\frac{\partial x_0}{\partial\psi} + \frac{F}{m}\cdot\sin(\omega t) \tag{7b}$$

$$\varepsilon^2: \quad \omega_0^2\frac{\partial^2 x_2}{\partial\psi^2} + \omega_0^2 x_2 = (\frac{\omega_0}{2}\cdot\frac{m_0'}{m})\frac{\partial x_1}{\partial\psi} - 2\omega_0\cdot\frac{\partial^2 x_0}{\partial\psi\partial T_2} - 2\omega_0\cdot\frac{\partial^2 x_1}{\partial\psi\partial T_1} - \frac{\partial^2 x_0}{\partial T_1^2} - (2\zeta\omega_0 + \frac{m_0'}{m})(\frac{\partial x_0}{\partial T_1} + \omega_0\frac{\partial x_1}{\partial\psi}) \tag{7c}$$

Letting $\theta = \omega t$, we get the solution of zero order approximation from Eq. 7a in the form:

$$x_0 = A(T_1, T_2)\exp(i\psi) + \bar{A}(T_1, T_2)\exp(-i\psi) \tag{8}$$

Where A is an unknown complex function of amplitude and $\bar{A}$ is the co njugate complex of A. Substituting Eq.8 into Eq.7band eliminating the terms that produces secular terms in x_1 yields

$$-(\frac{\omega_0}{2}\cdot\frac{m_0^{'}}{m})A+2\omega_0\frac{\partial A}{\partial T_1}+(2\zeta\omega_0+\frac{m_0^{'}}{m})\omega_0 A=0 \tag{9}$$

Because of $m'_0/m=-2\omega'_0/\varepsilon\omega_0$, from Eq. 9, we get:

$$\frac{\partial A}{\partial T_1}=-\frac{1}{2}(2\zeta\omega_0-\frac{1}{\varepsilon}\frac{\omega_0^{'}}{\omega_0})A \tag{10}$$

Let $$x_1=-i\cdot con\cdot\exp(i\theta) \tag{11}$$

Where con denotes a constant.
Substituting Eq. 11 into Eq. 9 yields

$$con=\frac{F\omega_0^2}{2k(\omega_0^2-\omega^2)} \tag{12}$$

then $$x_1=A(T_1,T_2)\exp(i\psi)-\frac{iF\omega_0^2}{2k(\omega_0^2-\omega^2)}\exp(i\theta)+cc \tag{13}$$

Substituting Eq. 8 and Eq.13 into Eq.7c and eliminating the secular term from x_2, we assume:

$$-(\frac{\omega_0}{2}\cdot\frac{m_0^{'}}{m})iA+2\omega_0 i\frac{\partial A}{\partial T_2}+\frac{\partial^2 A}{\partial T_1^2}+2\omega_0 i\frac{\partial A}{\partial T_1}+(2\zeta\omega_0+\frac{m_0^{'}}{m})(\frac{\partial A}{\partial T_1}+i\omega_0 A)=0 \tag{14}$$

The differentiation of Eq. 10 with respect to T_1 gives:

$$\frac{\partial^2 A}{\partial T_1^2}=\frac{A}{4}[4\zeta^2\omega_0^2-\frac{4}{\varepsilon}\zeta\omega_0^{'}+\frac{1}{\varepsilon^2}(\frac{\omega_0^{'}}{\omega_0})^2] \tag{15}$$

Substituting Eq. 10 and Eq. 15 into Eq. 14, we obtain:

$$\frac{\partial A}{\partial T_2}=-\frac{A}{8}i\cdot[4\zeta^2\omega_0+\frac{3}{\varepsilon^2}\frac{\omega_0^{'2}}{\omega_0^3}-\frac{8}{\varepsilon}\zeta\frac{\omega_0^{'}}{\omega_0}] \tag{16}$$

Writing A in the form:

$$A=A(T_1,T_2)=A(\varepsilon t,\varepsilon^2 t)=\frac{\alpha(t)}{2}\exp[i\beta(t)] \tag{17}$$

Where $\alpha(t)$ and $\beta(t)$ are real functions of T_1 and T_2. It follows that

$$\frac{dA}{dt}=\frac{1}{2}(\dot{\alpha}+i\alpha\dot{\beta})\exp(i\beta)=\varepsilon\frac{\partial A}{\partial T_1}+\varepsilon^2\frac{\partial A}{\partial T_2} \tag{18}$$

Substituting Eq. 10, Eq. 16 into Eq. 18, we separate the result into real and imaginary parts respectively and obtain:

$$\dot{\alpha}=-\frac{\varepsilon}{2}(2\zeta\omega_0-\frac{1}{\varepsilon}\cdot\frac{\omega_0^{'}}{\omega_0})\alpha \tag{19a}$$

$$\dot{\beta}=-\frac{\varepsilon^2}{8}[4\zeta^2\omega_0-\frac{8}{\varepsilon}\zeta\frac{\omega_0^{'}}{\omega_0}+\frac{3}{\varepsilon^2}\frac{\omega_0^{'2}}{\omega_0^3}] \tag{19b}$$

Substituting Eq. 8 and Eq.13 into Eq. 5 and using Eq.17，we obtain the second order approximation solution as follows:

$$x = x_0 + \varepsilon x_1 = (1+\varepsilon)\alpha(t)\cos(\psi + \beta(t)) + \frac{F_0\omega_0^2}{k(\omega_0^2 - \omega^2)}\sin(\omega t) \tag{20}$$

Where $\alpha(t)$ and $\beta(t)$ can be obtained from Eq.19a and Eq.19b.

2. Numerical Method Approach

Eq.1 can be rewritten as a system of three first order differential equations:

$$\begin{cases} \dot{x} = y \\ \dot{y} = -\dfrac{k}{m(z)}x - \dfrac{c + \varepsilon m_0'(z)}{m(z)}y + \dfrac{F_0\sin(\omega z)}{m(z)} \\ \dot{z} = 1 \end{cases} \tag{21}$$

Where $m(z)=m_1[1+\varepsilon\sin(\omega_a z+\varphi)]$ and $m'_0(z)=m_1\omega_a cos(\omega_a z+\varphi)$.
Runge-Kutta method of four-orders was introduced to get numerical solutions.

3. Comparison and discussions

Following parameters were used for the calculations to compare the solutions of the Multiple-Scales Method with that of the numerical one:
m_1=100kg, F_0=2.33kN, c=1.13$kN{\cdot}s/m$, k=39.2kN/m, ω=100.53rad/s, $\omega_a=\omega/6$, $\varphi=\pi/4$.[7]

As demonstrated in Fig.1, the responses of the two different methods show very good consistence, manifesting that the modified Multiple-Scales Method is valid.

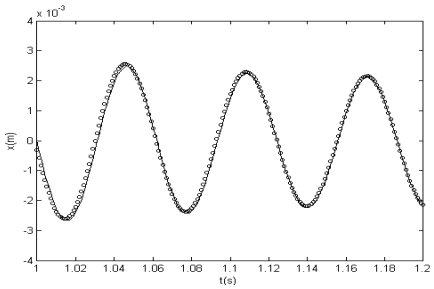
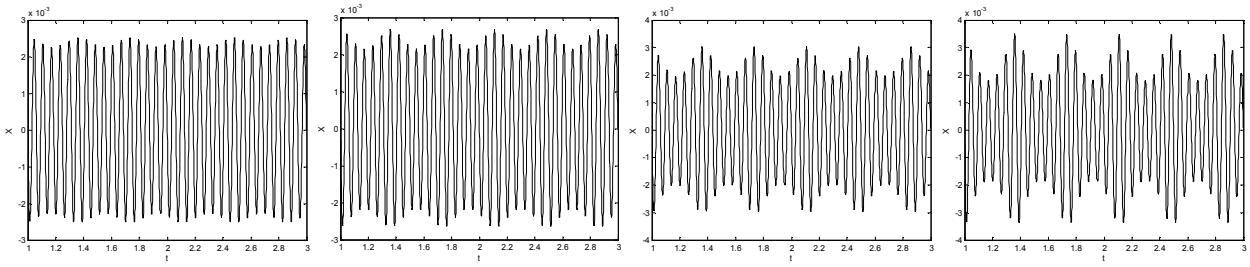


Fig.1 Response-time curves obtained analytical (A_a) and numerical (A_n) for ε=0.1
(——) A_a (○○○) A_n

3.1 Influence of the time-varying mass parameter ε on the system response

Fig.2 shows the response and time relation of the system obtained by Multiple-Scales Method for various values of the small parameter ε. One can see that the response features periodic behavior. It is important to recognize that ε is decisive to the amplitude magnification, although it did not influence the period. To further observe the contribution of ε to the motion, we represented the response determined by different ε in the same vibration period, as shown in Fig.3, and its local magnified pattern near the peak area is illustrated in Fig.4.

Obviously, the more the mass factor changes, that is, the bigger the coefficient ε, the more intensive the response enlarges, and vice versa. In other words, time-varying mass may be the main source causing the system non-stationary; this claim is consistent with [8].



(a) ε=0.05 (b) ε=0.1 (c) ε=0.2 (d) ε=0.3

Fig.2 Response-time curves determined by different ε

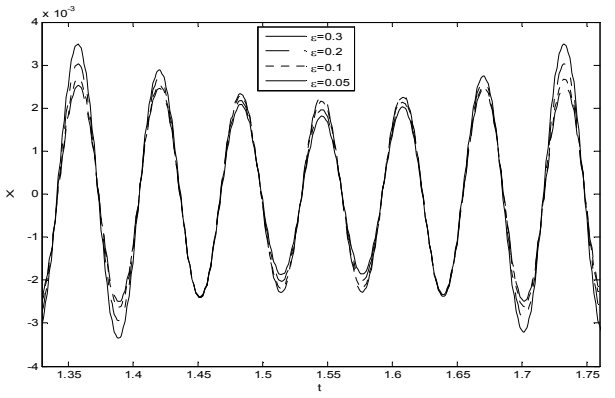


Fig.3 The response-time curves of the system with different ε

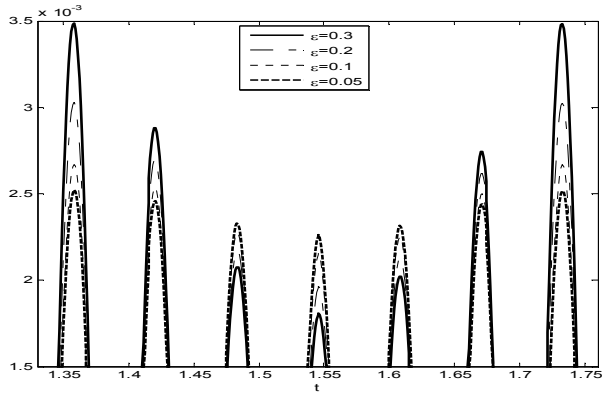
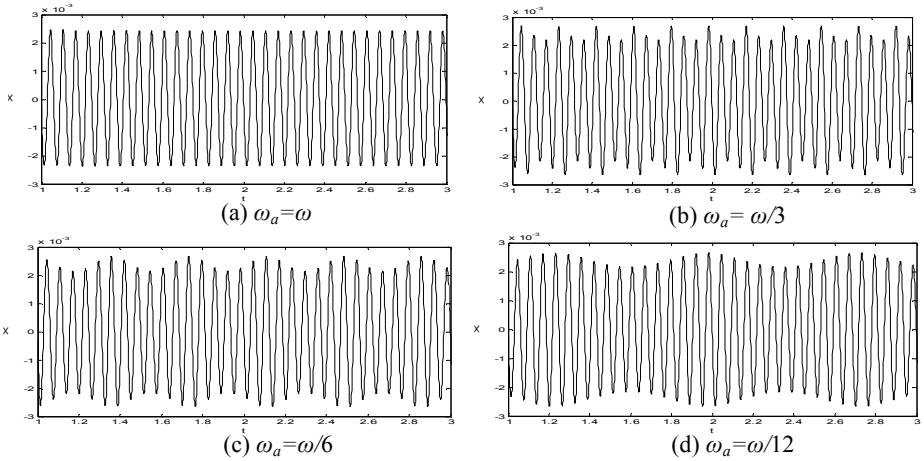


Fig.4 Local magnified pattern of response-time curves of the system with different ε

3.2 Influence of the time-varying mass frequency

Fig.5 is the influence of the time-varying mass frequency ω_a on the system response in time domain, where we assume the mass factor is constant of ε=0.1, while $\omega_a=\omega$, $\omega/3$, $\omega/6$ and $\omega/12$, respectively.

We can see that the period of response increases with the decreasing of ω_a. For quantitative analysis, we plotted the response-time curves of the system with different ω_a in the same time-history, as shown in Fig.6.



(a) $\omega_a=\omega$ (b) $\omega_a=\omega/3$

(c) $\omega_a=\omega/6$ (d) $\omega_a=\omega/12$

Fig.5 Responses-time curves of the system with different ω_a

Let T_0, T_1, T_2 and T_3 be the response periods, when$\omega_a=\omega$, $\omega/3$, $\omega/6$ and $\omega/12$, respectively. We can see that $T_1=3T_0$, $T_2=6T_0$ and $T_3=12T_0$. Thus, the response period is equal to that of the time-varying mass.

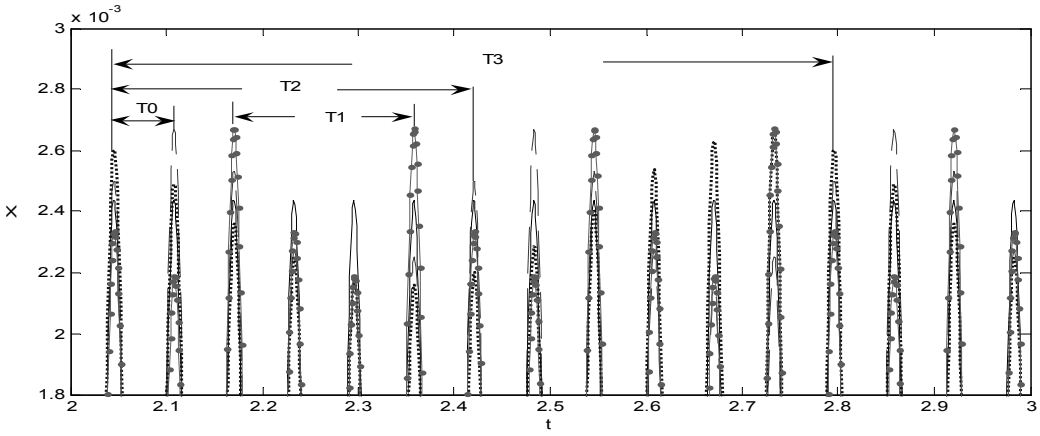


Fig.6 Local magnified pattern of response-time curves of the system with different ω_a

(——) $\omega_a = \omega$ (—•—) $\omega_a = \omega/3$ (— —) $\omega_a = \omega/6$ (-------) $\omega_a = \omega/12$

4. Conclusion

We used a modified Multiple Scales Method to solve the differential equation with time varying mass in that we express the solution as a function of both different time scale and phase parameter. By using Multi-Scales Method we obtained an approximate analytical solution of second order for the vibration. The feasibility is verified by the approach using Runge-Kutta numerical algorithms. We show that the motion of the variable mass system has periodic behavior with the period of the changing mass, and the mass variation can only influence the system amplitude but not its cycle. However, the more the mass factor varies, the more intensive the response enlarges, and vice versa. The method and findings may be useful to analyze similar vibration systems with impact dampers or design the vibration control strategy.

References

[1] S.L.Wang,M.Q.Liu,Y.Q.Hu,et al: Chinese Journal of Mechanical Engineering,Vol.33 (1997) , p.19-25.

[2] L.Cveticanin, in: Dynamics of Machines with variable mass, Gordon and Breach Science Publisher Australia(1998).

[3] A.H.Nayfeh, , Mook, D.T. in: Nonlinear oscillations. Wiley, New York(1995).

[4] S.Bravo Yuste: Int.J.Nonlinear Mechanics,Vol.26(1991),p.671-677.

[5] B.C.Wen: Analytical method of non-linear vibration theory and engineering application, Northeastern university press, Shenyang(2001)

[6] S. Hiamang, R.E.Mickens: Journal of Sound and Vibration,Vol.164(1993),p.179-181

[7] Y.Zhu, S.L.Wang: Journal of vibration and shock, Vol. 26 (2007) ,p.156-158

[8] S.L.Wang: Progress in Natural Science, Vol.12 (2002) , p.336-341

doi:10.4028/www.scientific.net/AMM.50-51.166

A small-world model of scale-free networks: features and verifications

Wenjun Xiao[1,a], Shizhong Jiang[1,2,b], Guanrong Chen[3]

[1]School of Software Engineering, South China University of Technology, Guangzhou 510641, P. R. China

[2]Department of Information Engineering, Gouangdong Pharmaceutical University, Guanggzhou 510006, P.R.China

[3]Department of Electronic Engineering, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong SAR, P. R. China

[a]wjxiao@scut.edu.cn, [b]jiang6499@126.com

Keywords: complex network, scale-free network, small-world network, computer network, software network.

Abstract. It is now well known that many large-sized complex networks obey a scale-free power-law vertex-degree distribution. Here, we show that when the vertex degrees of a large-sized network follow a scale-free power-law distribution with exponent $\gamma \geq 2$, the number of degree-1 vertices, if nonzero, is of order N and the average degree is of order lower than log N, where N is the size of the network. Furthermore, we show that the number of degree-1 vertices is divisible by the least common multiple of k_1^γ, k_2^γ, . . ., k_l^γ, and l is less than log N, where $1 = k_1 < k_2 < \cdots < k_l$ is the vertex-degree sequence of the network. The method we developed here relies only on a static condition, which can be easily verified, and we have verified it by a large number of real complex networks.

1. Introduction

Complex systems consisting of many components with intrinsic interactions are ubiquitous in nature and human society [1-5]. Components and interactions in such systems can be modeled by vertices and edges of networks, respectively, which in turn can be represented by graphs. A graph has a set of vertices and a set of edges, where each edge is defined by a pair of vertices [6].

Two typical models of complex networks have been studied extensively [1-9]: the scale-free model of Barabási and Albert [1] and the small-world model of Watts and Strogatz [2].

The small-world model features localized clusters that are connected by sparse long-range edges, leading to a short average distance between vertices which grows logarithmically with the network size. Watts and Strogatz [2] revealed the underlying mechanism as how a regular network can be transformed into a small-world network without significantly modifying the vertex-degree distribution.

Scale-free networks, on the other hand, show heterogeneous vertex connectivity, in which a fraction of vertices are highly connected. Barabási and Albert [1] demonstrated that the scale-free power-law distribution of vertex degrees in many large-sized networks is a direct consequence of two generic mechanisms that govern the network formation: (i) network expansion over time through addition of new vertices; (ii) preferential attachment of new vertex to those existing ones that are already highly connected.

In this paper, we focus on scale-free networks. We first establish some characteristic conditions for scale-free networks and show that these conditions are easy to verify and can be satisfied by many natural and man-made scale-free networks. We then show that the number of degree-1 vertices is divisible by the least common multiple of k_1^γ, k_2^γ, . . ., k_l^γ, with l less than log N, where $1 = k_1 < k_2 < \cdots < k_l$ is the degree sequence of the network. We consequently propose a way to reorganize a scale-free network so as to equip it with some small-world features. We finally verify the model by a large number of real network examples.

2. Some new features of scale-free networks

Assume, throughout, that all the networks under consideration are connected; unconnected networks may be similarly studied therefore are not discussed.

Let N and M be the numbers of vertices and edges, respectively, and $P(k)$ be the probability distribution of the number of vertices of degree k in the network. Let A denote the average vertex degree and n_k be the number of degree-k vertices. Clearly, $2M = NA$ is a good approximation and $n_k = NP(k)$. The following identities follow directly from the above definitions:

$$\sum_{k=1}^{N-1} n_k = N \tag{1a}$$

$$\sum_{k=1}^{N-1} k n_k = 2M \tag{1b}$$

For scale-free networks, typically $n_1 \neq 0$, so one has

$$n_k = n_1 P(k) / P(1) \tag{2}$$

and

$$\sum_{k=1}^{N-1} [P(k) / P(1)] = N / n_1 \tag{3a}$$

$$\sum_{k=1}^{N-1} [kP(k) / P(1)] = 2M / n_1 \tag{3b}$$

For scale-free networks, $P(k) = P(1)\, k^{-\gamma}$, which and (3a) together give

$$\sum_{k=1}^{N-1} k^{-\gamma} = N / n_1 \tag{4}$$

Therefore, assuming $\gamma \geq 2$, as is in many scale-free networks [3, 4], one obtains

$$N / n_1 \leq \sum_{k=1}^{\infty} k^{-\gamma} \leq \sum_{k=1}^{\infty} k^{-2} = \pi^2 / 6 \tag{5}$$

For large-sized scale-free networks, typically n_1 is very large, so the above implies that $N \approx n_1$.

Now, let $f(k) = P(k)/P(1)$. Then, when $N \approx n_1$, equations (3) yield

$$\sum_{k=1}^{N-1} f(k) \approx 1 \tag{6a}$$

$$\sum_{k=1}^{N-1} f(k) \approx A \tag{6a}$$

For $\gamma \geq 2$, one has

$$A \approx \sum_{k=1}^{N-1} k f(k) = \sum_{k=1}^{N-1} k \cdot k^{-\gamma} \leq \sum_{k=1}^{N-1} k^{-1} \approx \log_2 N \tag{7}$$

Therefore,

$$1 < A \leq \log_2 N \tag{8}$$

Thus, we have shown that for scale-free networks with power-law exponent $\gamma \geq 2$, the number of degree-1 vertices, if nonzero, is of order N and that the average degree is of order less than $\log_2 N$.

The above analysis leads to a general approach to reorganizing scale-free networks, satisfying $k^\gamma = n_1/n_k$. Assuming that this equality is actually exact and noting that its right-hand side is a rational number, one can show that k^γ must be an integer dividing n_1. Thus, n_1 must be divisible by the least common multiple of k_1^γ, k_2^γ, . . . , k_l^γ, denoted as $[k_1^\gamma, k_2^\gamma, ..., k_l^\gamma]$, where $1 = k_1 < k_2 < \ldots < k_l$ is the degree sequence of the network. Let $n_1 = c[k_1^\gamma, k_2^\gamma, ..., k_l^\gamma] \neq 0$ for some constant c. In the present case of $\gamma \geq 2$, one can easily obtain the estimate $n_1 = c[k_1^\gamma, k_2^\gamma, ..., k_l^\gamma] \geq 2^\gamma \times 2^\gamma \times \cdots \times 2^\gamma \geq 2^{\gamma(l-2)}$, which gives an estimate $n_1 \geq 2^{\gamma(l-2)}$. Consequently, it follows from $\log_2 N > \log_2 n_1 \geq \gamma(l-2)\log_2 2 \geq l-1$ another estimate that $l \leq \log_2 N$. Note that the above is a theoretical estimate. Numerically, due to measurement errors existing in raw data sets, we could only verify some slightly different bounds, namely $l \leq (\log_2 N)^2$, for many real scale-free networks are not exact, as can be checked from the real data given in the Appendix. On the other hand, the conclusion will have many important applications such as maximal degree search in networks [9] and small world feature [5].

3. Reorganizing scale-free networks

In this section, we propose a method to reorganize a given scale-free network, so that it has both scale-free and small-world features, if it originally was not a small-world network (it is well known that not all scale-free networks are also small-world networks).

We start with an arbitrary scale-free network, and do the following grouping.

Let $1 = k_1 < k_2 < \cdots < k_l$ be the vertex-degree sequence of the given network, and group all vertices of degree k together, labeled it as B_k, so the whole set of vertices is $V = \bigcup_{i=1}^{l} B_{k_i}$.

We then classify the whole set of edges according to two types of their connections: (1) the intra connections in each B_{k_i} and (2) the inter connections between B_{k_i} and B_{k_j}, $i \neq j$. Since the network is connected and B_1 is the set of vertices of degree 1, it is clear that the vertices in B_1 must be connected to other B_{k_i} with $i \neq 1$. For $i \geq 2$, the spanning subgraph of B_{k_i} in the network may be disconnected. But, when all the connected components of the spanning subgraph of B_{k_i} are small-world subnetworks for $i \geq 2$ and are connected to B_{k_j} with $i \neq j$, the entire network has small-world features, since l and A are in the same order and the latter was shown in (8) in the order of $\log_2 N$, which is a typical small-world property.

Thus, the resulting network maintains the original scale-free features, since nothing has been changed structurally, and yet possesses additional small-world properties as a whole [5].

4. Verification by some real networks

We have verified our method of remodeling by a large number of diverse real networks including social, biological and artifact networks, as summarized in Tables 1 and 2.

Table 1. Number of vertices (N), number of edges (M), average degree (A), characteristic exponent (γ) in some real networks.

Network	***N***	***M***	***A***	γ
Internet	10,687	31,992	5.98	2.5
Film actors	449,913	25,516,482	113.43	2.3
Metabolic network	765	3686	9.64	2.2
Protein interactions	2115	2240	2.12	2.4

Table 1 lists the parameters N, M, A, and γ for several real scale-free networks [3, 4]. It can be easily verified that $A \approx$ ½ $\log_2 N$ (5 $\log_2 N$, $\log_2 N$, and $\log_2 \log_2 N$) for the Internet (film actors,

metabolic, and protein interaction networks), respectively. Table 1 shows some real examples from diverse domains of technological networks, social networks, and biological networks, therefore is representative.

Various types of software graphs are shown in Table 2. The data presented in Table 2 further verify our remodeling method and results obtained above. On the other hand, the Ythan network, for example, has $\gamma = 1.04$, which is much below the lower bound 2 used to establish our results above.

Table 2 N vs n_1 and A vs $\log_2 N$ for some real complex networks.

Networks	N	M	n_1	A	n_1/N	$A/\log_2 N$
XMMS[s]	971	1802	300	3.711637	0.309	0.374
AbiWord[s]	1035	1719	447	3.321739	0.432	0.332
MySQL[s]	1480	4190	258	5.662162	0.174	0.538
Linux[s]	5285	11352	1200	4.295932	0.227	0.347
NCSTRL	6396	15872	894	4.963102	0.140	0.393
Ythan[b]	135	596	7	8.829630	0.052	1.248
Elegans[b]	314	363	203	2.312102	0.646	0.279

* *For software networks superscribed with s, see www.tc.cornell.edu/~myers/Data/SoftwareGraphs/index.html, for the data on biological networks superscribed with b, see www.cosin.org/extra/data, and the data of NCSTRL, a scientific collaboration network, is provided by Mark E. J. Newman [7-8].*

5. Conclusion

In this paper, we have shown that when the vertex degrees of a large-sized network follow a scale-free power-law distribution with exponent $\gamma \geq 2$, the number of degree-1 vertices, if nonzero, is of order N and the average degree is of order lower than $\log N$, where N is the size of the network. Furthermore, we have shown that the number of degree-1 vertices is divisible by the least common multiple of k_1^γ, k_2^γ, . . ., k_l^γ, with l less than $\log N$, where $1 = k_1 < k_2 < \cdots < k_l$ is the vertex-degree sequence of the network. Consequently, we have proposed a remodeling method to equip a scale-free network with small-world features. We have finally verified the method and results by a large number of real complex networks.

In the future, it is worth further investigating scale-free networks with power-law exponent $\gamma < 2$ and search in networks.

References

[1] Barabási, A.-L.,Albert, R:Emergence of scaling in random networks, Science, vol. 286(1999), p.509-512.

[2] Watts, D. J., Strogatz, S. H:Collective dynamics of 'small-world' networks", Nature, vol. 393(1998) , p.440-442.

[3] Albert, R., Barabási, A.-L:Statistical mechanics of complex networks, Rev. Mod. Phys., vol. 74(2002), p.47-97.

[4] Newman, M. E. J:The structure and function of complex networks, SIAM Rev., vol. 45(2003), p.167-256.

[5] Xiao, W. J., Parhami, B:Cayley graphs as models of deterministic small-world networks, Information Processing Letters, vol. 97(2006), p.115-117.

[6] Biggs, N:Algebraic Graph Theory (Cambridge University Press, 1993).

[7] Newman, M. E. J:Scientific collaboration networks, I: Network construction and fundamental results, Phys. Rev. E, vol. 64(2001), 016131.

[8] Myers, C. R:Software systems as complex networks: Structure, function and evolvability of software collaboration graphs, Phys. Rev. E, vol. 68(2003), 046116.

[9] Adamic, L. A, et al:Search in power-law networks, Phys. Rev. E, vol. 64(2001), 046135.

Appendix

Data of 7 complex networks used in this paper

	Network	N	M	l	$\log_2 N$	n_1	n_4	k_1	n_1
1	Xmms	971	1802	28	9.923	300	95	36	1
2	abiword	1035	1719	29	10.015	447	64	89	1
3	mysql2	1480	4190	43	10.531	258	140	220	1
4	Linux	5285	11352	51	12.368	1200	615	1058	1
5	ncstrlwg2	6396	15872	42	12.643	894	741	79	1
6	2004Internet	18408	33963	121	14.168	6505	667	2303	1
7	2009Internet	31164	63226	172	14.928	11325	1299	2283	1

doi:10.4028/www.scientific.net/AMM.50-51.171

The Design of E-Yuan Home Multimedia System

Weichang Feng

School of Computer and Communication Engineering, Weifang University,Weifang 261061, China

fweichang@163.com

Keyword: Linux, Embedded system, Web 2.0, RSS, TV, Remote Control

Abstract: E-Yuan Home multimedia system is used to help user to enjoy, search and integrate the plenty of multimedia resource on the internet. It used TV as the output device and user can watch the multimedia resource at real-time. It's fully controlled by a remote controller and also has a very friendly user interface. This article introduces its infrastructure, main technical ideas and you can also see some details about the remote control driver and how to solve the garbled characters problem.

Overview

Currently, Supporting RSS becomes more popular on the Web2.0 sites, RSS has become the most successful XML application in Web2.0. With the high speed of Internet bandwidth, the rapid growth of the Internet and the rapid increasing quantity of bandwidth user, the internet real-tine media industry should have a bright feature.

E-Yuan Home multimedia system is used to help user to enjoy, search and integrate the plenty of multimedia resource on the internet. It used TV as the output device and user can watch the multimedia resource at real-time. It's fully controlled by a remote controller and also has a very friendly user interface. In addition, the system also supports to play video, audio files or browse the images in the removable storage media.

System Introduction

E-Yuan Home multimedia system based on embedded systems , because it wants to minimize the costs.

Embedded systems is undoubtedly the most popular areas of IT applications, especially with the intelligent household electrical appliance. As we are usually familiar with mobile phones, electronic dictionaries, video phone, MP3 players, digital video (DV), high-definition television (HDTV), game consoles, smart toys are typical of embedded systems.

E-Yuan Home multimedia system separates into two parts (Hardware part and software part).

System Hardware Structure System hardware includes: motherboard, power supply, memory, hard drive, infrared remote control and the system enclosure.

The major part of system hardware is an industrial motherboards. Usually an Industrial motherboard designs for an fixed use and not easy to expend its function. With the market requirements of the board become more complex, it requires industrial motherboard has a strong expansion capabilities to meet different customer needs.

This system picks up J7F2WE1G5D embedded high-performance motherboard. The following is the main advantage of this board:

VIA CN700 chipset supports VIA C7 ® family of processors, using high bandwidth V4 bus,

support for DDR2 memory mode, integrated VIA UniChrome ™ Pro IGP core, using ChromotionCE video display engine, hardware MPEG-2 playback, support for a variety of digital video formats, support for a resolution of 1080p, HDTV, DVI or TV encoder compatible, supports 2GB bandwidth DDR2 533 or DDR400 memory, high-speed V-Link interface VIA VT8237A South Bridge chip, transmission speeds of up to 533MB / s, support SATA, SATA II and V-RAID device, provides two COM interfaces and eight USB2.0 interface, supports RAID 0, RAID 1 and RAID +1 arrays, support for VIA Vinyl HD Audio, VIA Velocity controller supports high throughput amount of Gigabit Ethernet.

System Software Structure System software includes the embedded Linux operating system and its applications. Linux operating system is the most successful applications in the embedded field. After the hardware design is complete, we need software to achieve every kind of functions. The value-added of embedded system largely depends on the level of the embedded software application. That is the intelligence level of products is determined by the software.

The system software design has two major works: (1) Drivers for all hardware on linux. (2) The search algorithm based on the internet video and audio resources. Software system includes the base subsystem, multimedia subsystem, and human-computer interaction subsystems.

Functional description of each subsystem are as follows:

Base subsystem: Contains the basic library and some base functions. Other subsystems depends on this subsystem.

Multimedia Subsystem: The core subsystem, including all media-related input and output functions.

HCI subsystem: allows the users to quickly and easyly interact with the system.

System data flow control principle shown in Fig.1. (MMS means Multimedia Subsystem)

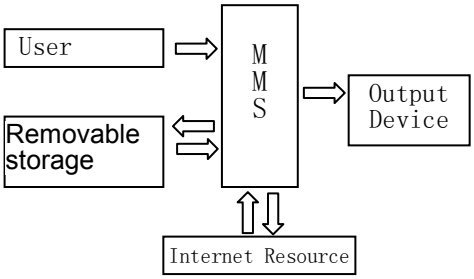


Fig.1 System data flow control principle

IMS is the core of the system design. The principle framework shown in Fig.2.

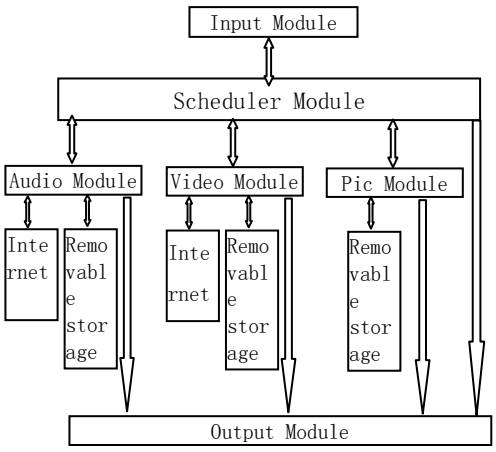


Fig.2 The principle of MMS

Main technical characteristics

Adjusting instruction architecture according to the cpu on the motherboard Adjusting the linux kernel hardware instruction architecture according to the cpu on the motherboard to let the software system and hardware system work together more smoothly, as shown in Fig.3.

The design and implementation of internet resources integration, optimization functions One of the two major characteristics of this system is that users do not need computers. With a broadband connection, user could search and real-time view online video through a TV.

In order to automatically search for the internet video and audio resources, the most important is to design algorithms to classify the video and audio resources and output the result as channels. By using technology in WEB2.0 RSS subscription feature, system could gather the video and audio resources to every channel and instantly create the topic part of the search results view, so users can choose to play or download to the local storage. RSS-based work flow as shown in Fig.4.

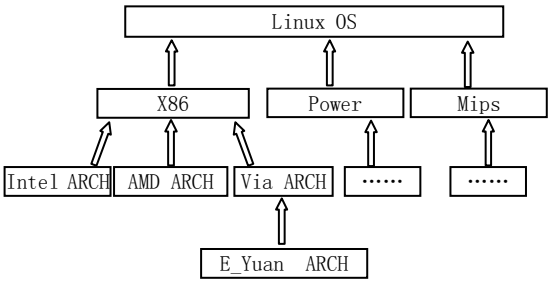


Fig.3 Adjusting the linux kernel hardware instruction architecture according to the cpu on the motherboard

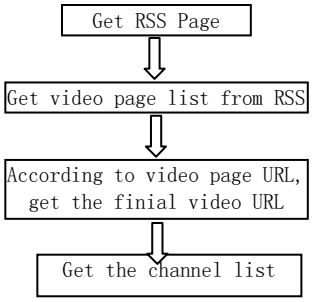


Fig.4 RSS data work flow

Development of remote control drivers The another feature of this system is: easy to use. The system uses infrared remote control, young and old, are available through a simple remote control to watch video, music and photos.

No computer, do not use the keyboard and mouse, all operations are replaced by an infrared remote control, so the remote control driver is most important. The system relies on linux input subsystem framework and try the best to make remote control more powerful and stable.

Register hardware device By calling function Input_register_device and Input_Unregister_device, driver layer register or remove from the input devices subsystem. Here is some sample code:

```
struct input_dev remote_dev;
static int __init remote_init(void)
{    memset(&remote_dev, 0, sizeof(struct input_dev));
     init_input_dev(&remote_dev);
     ……
     input_register_device(&remote_dev);
     return 0;
}
```

Registration Input Handler Before creating the device nodes, Event Handler layer need to register for a class of input device event handlers and related interfaces. Here is some sample code.

```
static struct input_handler remotedev_handler = {      event:  remotedev_event,
 connect:  remotedev_connect,
 disconnect: remotedev_disconnect,
 fops:  &remotedev_fops,
 minor:  REMOTEDEV_MINOR_BASE,
  ……};
……
```

Improve the system code to solve the garbled character problem To support all the languages in the world, needs a strategy with 8-byte character encoding character encoding system, its characters should be more than ASCII (a way to use unsigned byte extended version) 2 ^ 8 = 256 characters. Unicode is a character coding system so that it has the 128-D group (With a character encoding scheme by a number of ways to support the 94,140 characters) consisting of four-dimensional code space, the most popular in the Linux character encoding program is that the Unicode transforms to format UTF-8.

Linux system does not support the transformation between the UTF-8 and the popular Chinese character coding systems, therefore, in the data exchange will be a lot of issues (such as problems in the Chinese character display), need special functions to resolve the differences, about the transformation between UTF-8 character set and the popular character sets. The following is a sample function structure:

```
size_t wcsrtombs (char *dest, const wchar_t **src, size_t len, mbstate_t *ps);
size_t mbstowcs(wchar_t *pwcs, const char *s,size_t n);
int wcwidth(wchar_t wc);
……
int Uni2UTF(wchar_t wchar, char *utf8)
{……
    return len;}
```

Conclusion

Innovation of this article: Using the RSS subscription of WEB2.0 technology, designed a new algorithm, integrating the internet video and audio resources for the user and user could enjoy the resources directly on the TV at real-time. With the high speed of Internet bandwidth, the rapid growth of the Internet and the rapid increasing quantity of bandwidth user, the internet real-tine media industry should have a bright feature and I believe this system has a wide marketing prospect. Thanks.

References

[1] C. Teng, B. Huang, and X. Ma. Research and application of file system in embedded Linux. *Microcomputer Information,* Vol. 11-2 (2008), p. 88-90.

[2] G.. Xiao, J. Li, and W. Chang, Z. Ding. The realization of USB driving program in embedded Linux. *Microcomputer Information,* Vol. 9-2 (2007), p. 86-88.

[3] M. Liu. Embedded system interface design and linux driver development. Beijing: *Press of Beihang University* (2006)

[4] H. Wei. Embedded systems designer tutorial. Beijing: *Tsinghua University Press* (2006)

[5] J. C. O'Reilly, R. Alessandro and K.H. Greg. Linux Device Drivers (3rd). *San Francisco: Oreilly & Associates* Inc (2005)

[6] J. Lombardo. Embedded Linux. Indianapolis: *SAMS Technical Publishing* (2001)

[7] D. O'Reilly, P. Bovet, and C. Marco. *Understanding the Linux Kernel. San Francisco: Oreilly & Associates* Inc (2007)

©(2014) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.176

Strongly Ding Projective, Injective and Flat Modules

Jianmin Xing

College of Mathematic and Physics, Qingdao University of Science and Technology,

Qingdao 266061,China

Xingjianmin1978@163.com

Keywords: Strongly Ding projective; Strongly Ding flat; commutative ring; FP-injective

Abstract. Some properties of strong Ding projective ,injective and flat modules are studied and some connections between strong Ding projective ,injective and flat modules are discussed . At last, these properties under change of rings are considered.

Introduction

Unless stated otherwise, throughout this paper all rings are associative with identity and all modules are unitary modules. Let R be a ring. We denote by R-Mod (Mod-R)the category of left(right) R-modules respectively. For any R-module M , $\mathrm{pd}_R(M)$,$\mathrm{id}_R(M)$ and $\mathrm{fd}_R(M)$ denotes the projective ,injective and flat dimension of M .

When R is two-sided noetherian, Auslander and Bridger [1] introduced the G-dimension, $G\text{-dim}_R(M)$ for every finitely generated R-module M . Several decades later, Enochs and Jenda [2.3] extended the ideas of Auslander and Bridger and introduced three homological dimensions, called the Gorenstein projective, injective and flat dimensions. These have been studies extensively by their founders and by Avramov, Christensen, Foxby, Frankild, Holm, Martsinkovsky, and Xu among others over arbitrary associative rings. They proved that these dimensions are similar to the classical homological dimensions; i.e., projective,injective and flat dimensions respectively.

D.Bennis and N.Mahdou [4] studied a particular case of Gorenstein projective, injective and flat modules, which they call respectively, strongly Gorenstein projective, injective and flat modules. They proved that every Gorenstein projective (resp.Gorenstein injective, Gorenstein flat) module is a direct summand of a strongly Gorenstein projective (resp. strongly Gorenstein injective, strongly Gorenstein flat) module. Yang and Liu proved in [5] that a module M is strongly Gorenstein projective(resp. injective, flat)if and only if so is $M \oplus H$ for any projective(resp.injective,flat) module H and they continue the study of the properties under change of rings. Specially, they consider completions of rings, Morita equivalences, polynomial extensions and localizations.

In a recent paper [6], for any $n>1$, Bennis and Mahdou introduced the notions of n-strongly Gorenstein projective and injective modules, Then they proved that an n-strongly Gorenstein projective module is projective if and only if it has finite flat dimension. They also gave some equivalent characterizations of n-strongly Gorenstein projective modules in terms of the vanishing of some homological groups.

J.Gillespie in [7] introduced the definition and basic properties of Ding injective ,projective and flat module ,which play the role of fibrant and cofibrant objects over Ding-Chen rings.

In this paper, based on the results mentioned above, we study a particular case of Ding projective, injective and flat modules, which they call respectively, strongly Ding projective, injective and flat modules. and investigate the relation among them. At last, we study these properties under change of rings.

The rest of the text is organized as follows： some definitions and theorems will be introduced in section 2; the main results of strongly Ding projective, injective and flat modules are described in section 3;

Preliminaries and Notations

In this section, we introduce and study some definitions and basic properties which are defined as follows.

Definition 1.([7]) A right R-module E is a called FP-injective if $\mathrm{Ext}^1{}_R(F,E)=0$ for all finitely presented modules F. More generally, the FP-injective dimension of a right R-module N is defined to be the least integer $n \geq 0$ such that $\mathrm{Ext}^{n+1}{}_R(F,E)=0$ for all finitely presented right R-modules F. The FP-injective dimension of N is denoted FP-id(N) and equals ∞ if no such n above exists.

We now introduce and look at basic properties of Ding injective modules. They also appear in [8] as "Gorenstein FP-injective modules". We refer the reader to [7] for more results, including results concerning existence of covers and envelopes.

Definition 2.([7]) We call a right R-module N Ding injective (for short SD-injective) if there exists an exact sequence of injective modules $\cdots \to I_1 \longrightarrow I_0 \longrightarrow I^0 \longrightarrow I^1 \to \cdots$ with $N=\ker(I_0 \to I^0)$ and which remains exact after applying $\mathrm{Hom}_R(E,-)$ for any FP-injective module E. We denote the class of Ding injective modules by DI;

We call a right R-module M strongly Ding projective(for short SD-projective) if there exists an exact sequence sequence of projective modules $\cdots \to P_1 \to P_0 \to P^0 \to P^1 \to \cdots$ with $M=\ker(P_0 \to P^0)$ and which remains exact after applying $\mathrm{Hom}_R(-,F)$ for any flat module F. We denote the class of Ding projective modules by DP;

We call a right R-module M Ding flat if there exists an exact sequence of flat modules $\cdots \to F_1 \to F_0 \to F^0 \to F^1 \to \cdots$ with $M=\ker(F_0 \to F^0)$ and which remains exact after applying $E \otimes_R -$ for any FP-injective module E. We denote the class of Ding flat modules by DF.

*Remark 3. ([7])*By definition, Ding injective modules are Gorenstein injective. When R is Noetherian the two notions coincide; Ding projective modules are Gorenstein projective.

The Main Results

Definition 4. We call a right R-module N strongly Ding injective (for short SD-injective) if there exists an exact sequence $\cdots \to I \xrightarrow{f} I \xrightarrow{f} I \xrightarrow{f} I \to \cdots$ with I injective and $N=\ker f$ and which remains exact after applying $\mathrm{Hom}_R(E,-)$ for any FP-injective module E. We denote the class of strongly Ding injective modules by SDI;

We call a right R-module M strongly Ding projective(for short SD-projective) if there exists an exact sequence $\cdots \to P \xrightarrow{f} P \xrightarrow{f} P \to \cdots$ with P projective and $M=\ker f$ and which remains exact after applying $\mathrm{Hom}_R(-,F)$ for any flat module F. We denote the class of strongly Ding projective modules by SDP.

Remark 5. By definition, strongly Ding injective modules are Ding injective ; strongly Ding projective modules are Ding projective.

By definition, Ding injective modules are Gorenstein injective. When R is Noetherian , strongly Ding injective modules are strongly Gorenstein injective modules; When R is perfect , strongly Ding projective modules are strongly Gorenstein projective modules.

Using the definitions, we immediately get the following results.

Proposition 6. (1).If $(M_i)_{i\in I}$ is a family of SD-projective modules, then $\oplus M_i$ is SD-projective.

(2).If $(N_i)_{i\in I}$ is a family of SD-injective modules, then $\oplus N_i$ is SG-injective.

The next result gives a simple characterization of the SD-projective modules.

Proposition 7. For any $M \in$ mod- R ,the following statements are equivalent:

(1) M is SD-projective;

(2) There exists an exact sequence: $0 \to M \to P \to M \to 0$ in mod- R with P projective , and $\mathrm{Ext}^{i}{}_{R}(M,F)=0$ for any $i \geq 1$ and any flat module F ;

(3) There exists an exact sequence: $0 \to M \to P \to M \to 0$ in mod- R with P projective , and $\mathrm{Ext}^{i}{}_{R}(M,F)=0$ for any $i \geq 1$ and any module F with finite flat dimension .

Proof. Using standard arguments, this follows immediately from the definition of SD-modules.

Dually, we give a simple characterization of the SD-injective modules.

Proposition 8. For any $N \in$ mod- R ,the following statements are equivalent:

(1) N is SD-injective;

(2)There exists an exact sequence: $0 \to N \to I \to N \to 0$ in mod- R with I injective , and $\mathrm{Ext}^{i}{}_{R}(E,N)=0$ for any $i \geq 1$ and any FP-injective module E ;

(3)There exists an exact sequence: $0 \to N \to I \to N \to 0$ in mod- R with I injective , and $\mathrm{Ext}^{i}{}_{R}(M,F)=0$ for any $i \geq 1$ and any module E with finite FP-injective dimension .

Proposition 9. Let R be a commutative ring and Q a projective R -module. If M is an SD-projective R -module, then $M \otimes_R Q$ is an SD-projective R -module.

Proof. There is an exact sequence $0 \to M \to P \to M \to 0$ R with P projective. Then $0 \to M \otimes_R Q \to P \otimes_R Q \to M \otimes_R Q \to 0$ is exact and $P \otimes_R Q$ is a projective R -module by [9, Ch. 2, 1 Theorem 3]. Let F be any flat R -module. Then $\mathrm{Ext}^{i}{}_{R}(M \otimes_R Q, F) = \mathrm{Hom}_{R}(Q, Ext^{i}_{R}(M, F) = 0$ by [10, p. 258, 9.20] for all $i \geq 1$. Hence $M \otimes_R Q$ is an SD- projective R -module. by Proposition 8.

Definitions 10. We call a right R -module M strongly Ding flat if there exists an exact sequence $\cdots \to F \xrightarrow{f} F \xrightarrow{f} F \xrightarrow{f} F \to \cdots$ with F flat and $M = \ker f$ and which remains exact after applying $E \otimes_R -$ for any FP-injective module E . We denote the class of strongly Ding flat modules by SDF.

Also, similarly to Proposition 7 and 8, we have the following characterization of the SD-flat modules.

Proposition 11. For any $M \in$ mod- R ,the following statements are equivalent:

(1) M is SD-flat;

(2) There exists an exact sequence: $0 \to M \to F \to M \to 0$ in mod- R with F flat, and $\mathrm{Tor}_{i}^{R}(I, M) = 0$ for any $i \geq 1$ and any FP-injective module I ;

(3) There exists an exact sequence: $0 \to M \to F \to M \to 0$ in mod- R with F flat, and $\mathrm{Tor}_{i}^{R}(I, M) = 0$ for any $i \geq 1$ and any module F with finite FP-injective dimension .

Proposition 12. Let R be a left coherent ring. Then any SD-projective right R -module M is SD-flat.

Proof. Let M be an SD-projective right R -module. By definition there exists an exact sequence $\cdots \to P \xrightarrow{f} P \xrightarrow{f} P \to \cdots$ with P projective , $M = \ker f$
and which remains exact after applying $\mathrm{Hom}_{R}(-,F)$ for any flat module F . Let E be an FP-injective left R -module. Since R is left coherent, E^{+} is a flat right R -module by Theorem 2.2 of [11]. So applying $\mathrm{Hom}_{R}(-,E^{+})$ we get the exact sequence
$0 \to Hom_{R}(P, E^{+}) \to\to Hom_{R}(P, E^{+}) \to Hom_{R}(P, E^{+}) \to 0$.

But this is naturally isomorphic to $0 \to (P \otimes_R E)^{+} \to (P \otimes_R E)^{+} \to (P \otimes_R E)^{+} \to 0$.

Therefore $0 \to P \otimes_R E \to P \otimes_R E \to P \otimes_R E \to 0$ is exact, proving M is a SD-flat right R-module.

Theorem 13. Let R be right coherent and M is an SD-flat left R-module ,then M^{+} is an SD-injective right R -module.

Proof. There exists an exact sequence $0 \to M \to F \to M \to 0$ in R -Mod with F flat. Then $0 \to M^{+} \to F^{+} \to M^{+} \to 0$ is exact in Mod- R and F^{+} is injective. Let I be an FP-injective

right R-module. Then $\mathrm{Ext}^i_R(I, M^+) = \mathrm{Tor}^i_R(I, M)^+ = 0$ for all $i \geq 1$, and hence M is an SD-injective right R-module.

Proposition 14. Let R be a commutative ring and M a flat R-module. If M is an SD-flat, then $M \otimes_R Q$ is an SD-flat R-module.

Proof. There is an exact sequence $0 \to M \to F \to M \to 0$ in R-Mod with F flat.. Then $0 \to M \otimes_R Q \to F \otimes_R Q \to M \otimes_R Q \to 0$ is exact and $F \otimes_R Q$ is flat R-modules. Let I be any FP-injective R-module and $\mathcal{F}$ be a flat resolution of I, Then

$$\mathrm{Tor}^i_R(M \otimes_R Q, I) = H_i(M \otimes_R Q) \otimes_R I \cong \mathrm{Tor}^i_R(M, Q \otimes_R I) = 0$$

for all $i \geq 1$, Since $Q \otimes_R I$ is an FP-injective R-module. Hence $M \otimes_R Q$ is an SD-flat R-module.

Proposition 15. Let R be a commutative ring. If M is an SD-projective R-module, then $M[x]$ is an SD-projective $R[x]$-module.

Proof. There is an exact sequence $0 \to M \to P \to M \to 0$ where P is projective. So $0 \to M[x] \to P[x] \to M[x] \to 0$ is exact in $R[x]$-Mod and $P[x]$ is a projective $R[x]$-module. Let Q be any flat $R[x]$-module. Then $Q[x] \cong R[x] \otimes_R Q \cong R^{(N)} \otimes_R Q \cong Q^{(N)}$. Hence $Q[x]$ is a flat $R[x]$-module, and so Q is a flat R-module by [11, Proposition 5.11]. Thus $\mathrm{Ext}^i_{R[x]}(M[x], Q) \cong \mathrm{Ext}^i_R(M, Q) = 0$ by [12, p. 258, 9.21] for all $i \geq 1$, and hence $M[x]$ is an SD-projective $R[x]$-module.

References

[1] M.Auslander,M.Bridger: Stable module theory. Mem. Amer.Math. Soc. 94(1969).

[2] E.E.Enochs, O.M.G.Jenda, Gorenstein injective and projective modules. Math.Z. Vol. 220(4), p.611-633 (1995)

[3] E.E.Enochs, O.M.G.Jenda:On Gorenstein injective modules. Comm.Algebra.Vol.21(10)，p.3489-3501(1993)

[4] D.Bennis, N.Mahdou: Strongly Gorenstein projective, injective and flat modules.

J. Pure Appl. Algebra Vol. 210(2), p.437-445 (2007)

[5] X.Yang and Z.Liu:Strongly Gorenstein projective, injective and flat modules. J. Algebra Vol.320(7), p.2659-2674(2008)

[6] J.Gillespie : Model Structures on Modules over Ding-Chen Rings. Homology, Homotopy and Applications, Vol.12(1), p.61–73(2010)

[7] Information on http://arxiv.org/abs/0812.2566v1

[8] N.Ding and L.Mao: Gorenstein FP-injective and Gorenstein flat modules. J. Algebra Appl. Vol.7(4), p. 491–506 (2008)

[9] W.T. Tong:An Introduction to Homological Algebra.Higher Education(1996).

[10] J.J. Rotman, An Introductions to Homological Algebra, Academic Press, New York (1979)

[11] D.J. Fieldhouse, Character modules, dimension and purity. Glasgow Math. J. Vol.13(1) , p.144–146 (1972)

[12] M.S. Osborne: Basic Homological Algebra, Grad. Texts in Math.Springer-Verlag, New York, Berlin(2003)

doi:10.4028/www.scientific.net/AMM.50-51.180

Incomplete Concept Lattice Data Analytical Method Research Based on Rough Set Theory

Jinpeng Wang [1], Baoxiang Liu[1], Zhendong Li [2], Lichao Feng [1]
[1] College of Science, Hebei Polytechnic University, Tangshan, Hebei, 063000, China
[2] Basic Education Department, Tangshan College, Tangshan, Hebei, 063000, China
wangjinpeng1010@gmail.com

Keywords: incomplete concept lattice; rough set; discernibility matrix; attribute reduction algorithm;

Abstract. Concept lattice and rough set are powerful tools for data analyzing and processing, has been successfully applied to many fields. However, the decision information is incomplete in many information systems. In this paper, the definition of incomplete concept lattice has been proposed, and some relation established between imperfect concept lattice and rough set. As is very important that the paper gives a new attributes reduction algorithm about incomplete concept lattice aims at the matter of the inefficient of reduction strategy based on discernibility matrix. Comparing with the attributes reduction for incomplete concept lattice which based on discernibility matrix, this reduction algorithm, reduces the spatial-temporal complexity.

Introduction

Concept lattice theory, also known as formal concept analysis (FCA)[1,2], proposed by R.Wille in1982, while the rough set theory (RST)[3,4] is proposed by Z. Pawlak in the same year. As two different decision-making methods, FCA established the hierarchy of concept according to the binary relation between object and attributes. The process of generated concept lattice from data set is a process of conceptual clustering [5]; In addition, the concept lattice by Hasse diagram lively and simple reflects the relationship between concept of generalization and specialization [6].At present, the concept lattice has been successfully used in many machine learning tasks, and many concept lattice construction algorithms has been established, and widely used in information retrieval, digital library, software engineering and knowledge discovery [7]. RST is a mathematical tool used in deal with fuzzy and uncertain knowledge, its basic idea is summarized the concepts and rules through knowledge base categories, and theoretical basis is equivalence relations [6]. One of the main advantages of rough set is that it does not require any preparation or additional information about the data. For example, the probability distribution in statistics, the basic probability assignment of DempsterShafer theory, the membership of fuzzy set theory [8]. Concept lattice as a powerful tool for data analysis and dealing has a rapid growth in recent years [7]. Many previous studies on the concept lattice are only based on the complete concept lattice [5-7, 9-11]. However, under restrictions of experimental conditions and level there was a problem in the process of building intelligent expert system in practical problems, this problem is many systems' information is incomplete, that's some attributes' value of certain objects may be default. This will affect the range of data analysis and problem-solving accuracy that using of the concept lattice method [15]. Therefore, is very important study on concept lattice based on incomplete information. In addition, there were lots of similarities between the concept lattice and rough set, and attracted many researchers concerned the relationship between them [5-7, 9-11]. Discussed the relationship between RST and FCA has a great significance. On the one hand, enriched and developed the theory of RST and FCA. On the other hand, we can better understand the data analysis and data rules through established interoperability between the two theories [10]. This paper combines the basic definition of concept lattice and incomplete of information systems, put forward the definition of incomplete concept context、incomplete concept and incomplete concept lattice. Then explore the links between incomplete concept lattice and rough set, established the links between incomplete concept lattice and rough set theory in incomplete information systems.

In particular, attribute reduction is a very important issue in the concept lattice theory and rough set theory [3, 11, 16-18]. The problem can be defined as found attributes number is N in the initial attributes, its number is M (N not more than M), used to describe the original data set. The improved discernibility matrix used in concept lattice of attribute reduction algorithm has been established [11].Whether discernibility matrix idea used in attribute reduction of incomplete concept lattice? In addition, construct the discernibility matrix very large and computing capacity, and reduction based on discernibility matrix strategy inefficient [19]. This paper attempts to establish a new attribute reduction algorithm of incomplete concept lattice, that scientific answer to these questions.

Incomplete concept lattice

Definition 1: (Incomplete context) $(U, D, Sim(B))$is a incomplete context.where, $U = \{x_1, x_2, \cdots x_n\}$is a set of objects. $D = \{d_1, d_2, \cdots d_n\}$is a set of attributes. $Sim(B)$is binary relation between U and D. $Sim(B) \subseteq U \times D$. x has attributes of d and whether x has attributes of d that we can not get from the process of observation indicats by $(x,d) \in Sim(B)$). $(x,d) \notin Sim(B)$ indicats x hasn't attributes of d.

Definition 2: (Incomplete concept) $(U, D, Sim(B))$ is a incomplete context. $\forall X \subseteq P(U)$ and $\forall Y \subseteq P(D)$ is a binary group (X,Y) satisfy two conditions $f^*(X) = Y$ and $g^*(Y) = X$. We called (X,Y) is a incomplete concept. X is the extension of incomplete concept. Y is the connotation of incomplete concept.where, f^* and g^* are two maps between $P(U)$ and $P(D)$:

$$\forall X \in P(U), f(X) = \{y \in D \mid \forall x \in X, xSim(B)y\},$$
$$\forall Y \in P(D), g(Y) = \{x \in U \mid \forall y \in Y, xSim(B)y\},$$

Definition 3: $L^*(k)$ indicats a set of all concepts in context, and all elements in $L^*(k)$ can establish a partial order. If $L^*(k)$has two elements $H_1^* = (X_1, Y_1)$and $H_2^* = (X_2, Y_2)$.

Stipulate: $H_1^* \le H_2^* \Leftrightarrow Y_2 \subseteq Y_1 \Leftrightarrow X_1 \subseteq X_2$

" $\le$ " is a partial order in $L^*(k)$ can be verified facilely, and it induces a lattice structure on $L^*(k)$. Now we called $(U, D, Sim(B))$ is a incomplete concept lattice.

Interoperability between incomplete concept lattice and rough set

Theorem 1: Each node of the incomplete concept lattice was a similar class of rough set, and each node is a set with the largest collection of objects that has common attribute.

Proof: Because the context is incomplete in the incomplete concept lattice, and the nodes of the lattice can be divided into two parts: One part is all objects of this incomplete concept $x_i (i = 1,2,\cdots)$, The other part is the attribute values of objects that can not distinguish under the attribute of d_i. So the set of objects that came from node of lattice can not distinguish under the attribute of d_i ; Similar class $S_B(x)$ can not distinguish under the attributes also,and $S_B(x)$is the largest collection of objects Indistinguishable from x in terms of $S_B(x)$.

Theorem 2: $SIM(B) = \bigcup \sup\{X \mid B \subseteq Y, (X,Y) \in L^*(k)\}$ if $(U, D, Sim(B))$is a incomplete context, $B \subseteq D$, and $SIM(B)$is a similar relationship family determined by B (a attribute set). $\sup(X)$ indicats least upper bound of all concept in X.

Proof: Because the similar relationship family $SIM(B)$ can be indicated that:

$$SIM(B) = \bigcap_{r \in B} r$$

Therefore theorem is right.

Theorem 3: Upper and lower approximation of rough set can be indicated by node of incomplete concept lattice as follows:

$$\overline{Sim(B)}(X)=\bigcup\{Z\mid Z=\sup\{X\mid B\subseteq Y,(X,Y)\in L^*(k)\}\cap X\neq\Phi\}$$
$$\underline{Sim(B)}(X)=\bigcup\{Z\mid Z=\sup\{X\mid B\subseteq Y,(X,Y)\in L^*(k)\}\subseteq X\}$$

Proof: Because the similar relationship family $SIM(B)$ can be indicated that:

$$SIM(B)=\bigcup\sup\{X\mid B\subseteq Y,(X,Y)\in L^*(k)\}$$

According to the definition of upper and lower approximation:

$$\underline{Sim(B)}(X)=\{x\in U\mid S_B(x)\subseteq X\}=\bigcup\{S_B(x)\mid\subseteq X\}$$

If $x_i\in\underline{Sim(B)}(X)$, then $S_B(x)\subseteq X$. According to theorem 2 know that:

$$x_i\in\bigcup\{S_B(x)\mid S_B(x)\subseteq X\}\in\sup\{X\mid B\subseteq Y,(X,Y)\in L^*(k)\}\subseteq X$$

On the contrary may also permit.

If $x_i\in\sup\{X\mid B\subseteq Y,(X,Y)\in L^*(k)\}$, then $x_i\in Sim(B)(X)$,then

$$\underline{Sim(B)}(X)=\bigcup\{Z\mid Z=\sup\{X\mid B\subseteq Y,(X,Y)\in L^*(k)\}\subseteq X\}$$

Similarly can be proved on the approximation.

Incomplete Concept Lattice Reduction Algorithm of Rough

Definition 3.1.2[19] (Attribute reduction) If $x\in U$, $S_R(x)=S_R(x)$, and $\forall B\subset R$ $S_B(x)\neq S_A(x)$, then $R\subseteq A$ is a reductionthat came from S about x.

Definition 3.1.2 (Attribute function) If the objects number is n ,and attributes number is m in S (Information Sheet), and attribute values

$$V=(V_1,V_2,\cdots,V_P(V_P="*"))$$

then the number of vector of v (attribute values)of $j(j=1,2,\cdots,m)$ (attribute) is $x_i=(x_{j1},x_{j2},\cdots,x_{jp})$. Attribute function Arf is:

$$Arf=x_{j1}\times x_{j2}+x_{j1}\times x_{j3}+\cdots+x_{j(p-2)}\times x_{j(p-1)}$$

ArM is Attribute matrix.The element of $ArM(a_j)$ is all different pairs of objects of a_j (attribute).

To simplify information system with the equivalence class of concept before request the discernibility matrix in context,so that reducing the computation and storage space. Attribute reduction strategy based on discernibility matrix should be compared every two objects on all of the attributes,so computation is very big when the data is large.Here,we defined a attribute function for reduce the computational.

Concrete steps are as follows:

(1) Simplify information system with the equivalence class of concept before request the discernibility matrix in context,so that reducing the computation and storage space.

$$S_B(x)=\{y\in U\mid(x,y)\in Sim(B)\};$$

(2) $\forall a_j\in T$, Calculate Arf : $Arf(a_k)=Max(a_j)$;

(3) $R\leftarrow R\cup\{a_k\}$;

(4) $\forall a_j\in T$, $B(a_j)=ArM(a_j)\cap ArM(a_k)$, $H(a_j)=num(B(a_j))$

(5) $\forall a_j\in T$, $ArM(a_j)\leftarrow ArM(a_j)-B(a_j)$, $Arf(a_j)\leftarrow Arf(a_j)-H(a_j)$

If $Arf(a_j)=0$, then $T=T-a_j$;

(6) $T=T-\{a_k\}$, $ArM=ArM-ArM(a_k)$;

(7) Export R.

Applications

An instance introducted for verifying the proposed reduction algorithm.

Objects $U=\{x_1,x_2,\cdots x_8\}$, the set of attributes $A=(a,b,c,d,e)$, $Sim(B)$ is a binary relation between U and A. Incomplete context(Table 1):

Table 1 Incomplete context

U	a	b	c	d	e
x1	a1	*	c1	d2	e1
x2	*	b2	*	d2	e1
x3	a2	b1	*	d1	*
x4	a3	b1	c2	d1	*
x5	a1	*	c1	d2	e1
x6	a2	b1	*	d1	*
x7	a1	*	c1	d2	e1
x8	*	b2	*	d2	e1

(1)Simplify information system with the equivalence class of concept before request the discernibility matrix in context,so that reducing the computation and storage space.

$L_1=S_B(x_1)=S_B(x_5)=S_B(x_7)$; $L_2=S_B(x_2)=S_B(x_8)$; $L_3=S_B(x_3)=S_B(x_6)$; $L_4=S_B(x_4)$

(2)The simplified of incomplete concept lattice(Figure1)and incomplete context(Table 2):

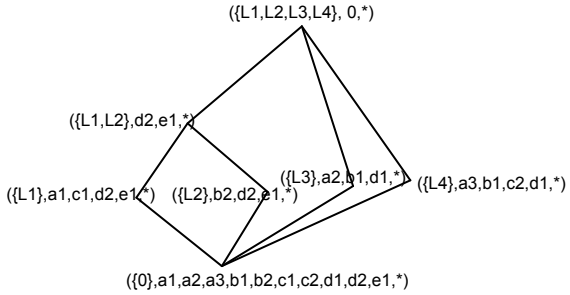


Table 2 The simplified incomplete context.

U/Sim(B)	a	b	c	d	e
L1	a1	*	c1	d2	e1
L2	*	b2	*	d2	e1
L3	a2	b1	*	d1	*
L4	a3	b1	c2	d1	*

Figure 1 Incomplete concept lattice after data compression.

(3)Established attribute matrix ArM（Table 3a）:

Table 3a The attribute values of simplified incomplete context.

attribute	a	b	c	d	e
Arf	3	2	1	4	0

(4)Delete attribute the attributes of the function is 0(Table 3b):

Table 3b The attribute values of simplified incomplete context.

attribute	a	b	c	d
Arf	3	2	1	4

(5)Recurrence (2)-(3), Ultimatly $R=\{d,a\}$.

(6)Export $R=\{d,a\}$.

References

[1] R. Will Restructuring lattice theory: an approach based on hierarchies of concepts. In: RivalI, editor, Ordered sets, Reidel, Dordrecht-Boston (1982), p.445-470.

[2] B. Ganter, and R. Wille: Formal Concept Analysis. Mathematical Foundations. Berlin: Springer, (1999)

[3] Z. Pawlak: Rough sets. International Journal of Computer and Information Sciences, vol.11 (1982), p. 341- 356

[4] Z. Pawlak: Rough sets. Theoretical Aspects of Reasoning about Data. Dordrecht: Kluwer Academic Publisher (1991)

[5] J.P.Wang and J.Y.Liang: Concept Lattice and Rough Set. Shanxi University (Natural Science) , vol.26 (2003), p.307-310

[6] K.Y.Hu: The data mining method based Concept lattice and rough set. Beijing: Tsinghua University (2001)

[7] K.Y.Hu and Y.C. Lu: Concept lattice and its application. Tsinghua University (Natural Science) vol.40 (2000), p.77-81.

[8] Z.Z. Shi: Knowledge Discovery. Beijing: Tsinghua University Press (2002)

[9] K.Y.Hu and Y.C. Lu: Rough Set Theory and Its Applications. Tsinghua University (Natural Science) vol.41 (2001),p.64-68.

[10] L. Wei and J.J. Qi: Concept Lattice and Rough Set Relations. Computer Science vol.33 (2006), p.18-21.

[11] S.P. Dai and T. He: Concept Lattice and Rough Set Data Analysis. Computer Engineering and Design vol.29 (2008), p.1423-1425.

[12] Y.Q. Wang and F.C. Lu: Synthetic fault diagnosis method of power transformer based on rough set theory and Bayesian network. Proceedings of the CSEE vol.26 (2006), p. 137-141.

[13] X.Y. Lu and J.L. Zhang: A method of risk rule mining in IT project based on rough set and bayes theory .Computer Engineering and Applications vol.22 (2006), p. 12-15.

[14] Q. Wen and H. Peng: Virus detection based on rough set and bayes classifier. Journal of Southwest Jiaotong University vol.40 (2005), p. 659-662.

[15] Y.L. Zhu and L.Z. Wu: Synthesized diagnosis on transformer faults based on bayesian classifier and rough set. Proceedings of the CSEE vol.25 (2005), p.159-165.

[16] Z. Pawlak: Rough sets. Theoretical aspects of reasoning about data. Kluwer Academic Publishers, Boston, London, Dordrecht (1991)

[17] B. Yu and B.P. Yan: On rough set attribute reduction and application of evolutionary algorithms. Microelectronics and Computer vol.22 (2005), p.189-194

[18] W.X. Zhang: Uncertain decision-making based on rough set. Beijing: Tsinghua University Press (2005)

[19] B. Shao and L.L. Wu: One rough set reduct for incomplete information system vol.24 (2007), p.204-207.

© (2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.185

Positive Solutions for the nth-order Delay Differential System with Multi-parameter

Qiuying Lu[1,a] , Weipeng Zhang[2,b]

[1]Department of Mathematics
Zhejiang Sci-Tech University, Xiasha Economic Development Area
Hangzhou City, Zhejiang Province, 310018, CHINA
[2]School of Mathematics and Statistics
Northeast Normal University
Changchun, Jilin, 130024, CHINA

[a]qiuyinglu@163.com [b]zhangwp996@nenu.edu.cn

Keywords: Positive solutions; Nonlinear nth-order delay differential systems; Cone fixed-point theorem; Boundary value problems;

Abstract. In this paper, we are concerned with the existence of positive solutions for the nonlinear eigenvalue problem of the nth-order delay differential system. The main results in this paper generalize some of the existing results in the literature. Our proofs are based on the well-known Guo-Krasnoselskii fixed-point theorem. Three main results are given out, the first two of which refer to the existence while the last one not only guarantees to its existence but also is pertinent to its multiplicity.

Introduction

This paper deals with the existence of positive solutions for the following nth-order delay differential system:

$$\begin{cases} x_1^{(n)}(t) + \lambda_1 g_1(t, x_1(t-\tau_1), x_2(t-\tau_2), \cdots, x_n(t-\tau_n)) = 0, \\ x_2^{(n)}(t) + \lambda_2 g_2(t, x_1(t-\tau_1), x_2(t-\tau_2), \cdots, x_n(t-\tau_n)) = 0, \\ \vdots \\ x_n^{(n)}(t) + \lambda_n g_n(t, x_1(t-\tau_1), x_2(t-\tau_2), \cdots, x_n(t-\tau_n)) = 0, \quad 0 < t < 1, \tau_i \geq 0, i = 1, 2 \cdots n, \end{cases} \tag{1.1}$$

satisfying

$$\begin{cases} x_i^{(k)}(t) = 0, \\ x_i(1) = 0, \qquad -\tau_i \leq t \leq 0,\ 0 \leq k \leq n-2,\ i = 1, 2 \cdots, n. \end{cases} \tag{1.2}$$

System (1.1)-(1.2) is a generalization of the single form boundary value problem (BVP, for short) of the nth-order delay differential equation :

$$u^{(n)}(t) + \lambda g(t, u(t-\tau_1)) = 0,\ 0 < t < 1, \tau_1 \geq 0, \tag{1.3}$$

$$u^{(k)}(t) = 0,\ -\tau_1 \leq t \leq 0,\ 0 \leq k \leq n-2,$$

$$u(1) = 0. \tag{1.4}$$

For the case of $\lambda = 1, \tau_1 = 0$, (1.3) and (1.4) are related to the nth-order BVP of ordinary differential equation. For the case of second order delay differential equation, (1.3) and (1.4)

describe many phenomena in the applied mathematical science such as nonlinear diffusion generated by nonlinear sources, thermal ignition of gases, and concentration in chemical or biological problems, where positive solutions are meaningful; see, for example [2-4].

In this paper, we devote section 2 to giving some important inequalities of the Green's function for the nth-order BVP thanks to the procedure provided by Eloe P.W. and Henderson J. [1]. These inequalities will take an important part in establishing the existence of solutions of (1.1) and (1.2). In section 3, we use these Green's function inequalities in defining an operator which is positive with respect to a cone. Then, positive solutions of (1.1) and (1.2) arise as fixed points of this operator which results from applications of the well-known Guo-Krasnoselskii fixed-point theorem [5]. We give not only the theorems which guarantee their existences but also one result about their multiplicity.

Useful Inequalities

In this section, we shall give several lemmas which are very useful for the proof of our main results.

Lemma 2.1[6] Let $n \geq 2$ and let $h \in C^{(n)}[0,1]$ satisfy

$$h^{(n)}(t) \leq 0,\ t \in [0,1],$$
$$h^{(k)}(0) \geq 0,\ 0 \leq k \leq n-2,$$
$$h(1) \geq 0.$$

Then,

$$h(t) \geq \| h \| \min\{1-t, t^{n-1}\},\ t \in [0,1], \tag{2.1}$$

where $\| h \| = \sup_{0 \leq t \leq 1} | h(t) |$.

Lemma 2.2[1] Let $n \geq 2$ and let $h \in C^{(n)}[0,1]$ satisfy

$$h^{(n)}(t) \leq 0,\ t \in [0,1],$$
$$h^{(k)}(0) \geq 0,\ 0 \leq k \leq n-2,$$
$$h(1) \geq 0.$$

Then,

$$h(t) \geq \frac{\| h \|}{4^{n-1}},\ \frac{1}{4} \leq t \leq \frac{3}{4}. \tag{2.2}$$

Let $G(t,s)$ be the Green's function for the BVP

$$-u^{(n)}(t) = 0,\ 0 \leq t \leq 1,$$

satisfying (1.4). It is well known [7] that

$$G(t,s) > 0 \text{ on } (0,1) \times (0,1).$$

For each $0 < s < 1$, let $\tau(s) \in [0,1]$ be defined by

$$G(\tau(s), s) = \sup_{0 \le t \le 1} G(t, s).$$

Lemma 2.3[6]

$$G(t,s) \ge G(\tau(s), s) \min\{1-t, t^{n-1}\},\ t \in [0,1]. \tag{2.3}$$

Lemma 2.4[1]

$$G(t,s) \ge \frac{1}{4^{n-1}} G(\tau(s), s),\ t \in [\frac{1}{4}, \frac{3}{4}],\ s \in [0,1]. \tag{2.4}$$

Eigenvalue problem with singularity

In this section, we consider the existence of positive solutions for the nonlinear eigenvalue problem (1.1) and (1.2) and their multiplicity.

Let

$$X_i = \{x \in C^{(n)}[-\tau_i, 1] : x^{(k)}(t) = 0,\ \forall\ t \in [-\tau_i, 0],\ 0 \le k \le n-2;\ x(1) = 0\}$$

with norm $\|\cdot\|_i$ given by $\| x \|_i = \sup\{| x(t) |: -\tau_i \le t \le 1\}$ for $i = 1, 2, \cdots, n$. Then $(X_i, \|\cdot\|_i)$ is a Banach space. Obviously, $\|\cdot\|_{[0,1]} = \|\cdot\|_i$ for $x \in X_i$, where $\|\cdot\|_{[0,1]}$ stands for the sup-norm of $C[0,1]$. So we can express $\|\cdot\|_i,\ (i = 1, 2, \cdots, n)$ uniformly as $\|\cdot\|$. Define $E = X_1 \times X_2 \times \cdots \times X_n$ with norm $\| (\cdot, \cdot, \cdots, \cdot) \|$ given by $\| (x_1, x_2, \cdots, x_n) \| = \sum_{i=1}^{n} \| x_i \|$. It is easy to see that E is a Banach space. Suppose that $(x_1(t), x_2(t), \cdots, x_n(t))$ is a solution of BVP(1.1) and (1.2), then it can be expressed as

$$(x_1(t), x_2(t), \cdots, x_n(t)) = (A_1(x_1, x_2, \cdots, x_n)(t), A_2(x_1, x_2, \cdots, x_n)(t), \cdots, A_n(x_1, x_2, \cdots, x_n)(t)),$$

where for $i = 1, 2, \cdots, n$,

$$A_i(x_1, x_2, \cdots, x_n)(t) = \begin{cases} 0, & -\tau_i \le t \le 0, \\ \lambda_i \int_0^1 G(t,s) g_i(s, x_1(s-\tau_1), x_2(s-\tau_2), \cdots, x_n(s-\tau_n))\, \mathrm{d}s, & 0 \le t \le 1. \end{cases}$$

Define a cone $K \subset E$ by $K = K_1 \times K_2 \times \cdots \times K_n$ with

$$K_i = \{x \in X_i :\ x(t) \ge 0,\ \forall t \in [0,1],\ \min_{t \in [\frac{1}{4}, \frac{3}{4}]} x(t) \ge \frac{1}{4^{n-1}} \| x \|\}.$$

Let

$$f_0^j = \liminf_{\sum_{i=1}^n x_i \to 0^+} \min_{s \in [0,1]} \frac{f_j(s, x_1, x_2, \cdots, x_n)}{\sum_{i=1}^n x_i}, \qquad f_j^0 = \limsup_{\sum_{i=1}^n x_i \to 0^+} \max_{s \in [0,1]} \frac{f_j(s, x_1, x_2, \cdots, x_n)}{\sum_{i=1}^n x_i},$$

$$f_\infty^j = \liminf_{\sum_{i=1}^n x_i \to \infty} \min_{s \in [0,1]} \frac{f_j(s, x_1, x_2, \cdots, x_n)}{\sum_{i=1}^n x_i}, \qquad f_j^\infty = \limsup_{\sum_{i=1}^n x_i \to \infty} \max_{s \in [0,1]} \frac{f_j(s, x_1, x_2, \cdots, x_n)}{\sum_{i=1}^n x_i}.$$

Define an operator $T : E \longrightarrow E$

$$T(x_1, x_2, \cdots, x_n)(t) = (A_1(x_1, x_2, \cdots, x_n)(t), A_2(x_1, x_2, \cdots, x_n)(t), \cdots, A_n(x_1, x_2, \cdots, x_n)(t))$$

with $T_i(x_1, x_2, \cdots, x_n)(t) = A_i(x_1, x_2, \cdots, x_n)(t),\ i = 1, 2, \cdots, n$. Then, one can easily see that $(x_1, x_2, \cdots, x_n)$ is a solution of (1.1) and (1.2) if and only if $(x_1, x_2, \cdots, x_n)(t) = T(x_1, x_2, \cdots, x_n)(t)$.

Let $p_i > 1,\ i = 1, 2, \cdots, n$ such that $\sum_{i=1}^{n} \frac{1}{p_i} = 1$. And

$$\tau_M = \max\{\tau_i, i = 1, 2, \cdots, n\},\ \tau_m = \min\{\tau_i, i = 1, 2, \cdots, n\}.$$

For $j = 1, 2, \cdots, n$, denote by

$$C_{j1} = \frac{4^{n-1}}{f_\infty^j \cdot \sup\limits_{t\in[0,1]} \cdot \int_{\frac{1}{4}+\tau_j}^{\frac{3}{4}+\tau_j} G(t,s)a_j(s)\,\mathrm{d}s},\ C_{j2} = \frac{1}{p_j \cdot f_j^0 \cdot \int_{\tau_m}^{1} G(\tau(s),s)a_j(s)\,\mathrm{d}s},$$

$$D_{j1} = \frac{4^{n-1}}{f_0^j \cdot \sup\limits_{t\in[0,1]} \cdot \int_{\frac{1}{4}+\tau_j}^{\frac{3}{4}+\tau_j} G(t,s)a_j(s)\,\mathrm{d}s},\ D_{j2} = \frac{1}{p_j \cdot f_j^\infty \cdot \int_0^1 G(\tau(s),s)a_j(s)\,\mathrm{d}s}.$$

Theorem 3.1 *Assume that*
(H_1) $0 \le \tau_M < \frac{1}{4}$;
(H_2) $g_j(t, x_1, x_2, \cdots, x_n) = a_j(t)f_j(t, x_1, x_2, \cdots, x_n),\ a_j$: $(0, 1) \longrightarrow [0, \infty)$ is continuous, and $f_j : [0, 1] \times [0, \infty) \longrightarrow [0, \infty),\ j = 1, 2, \cdots, n$ is continuous;
(H_3) $\int_0^1 a_j(s)\,\mathrm{d}s < \infty,\ 0 < \int_0^1 G(\tau(s), s)a_j(s)\,\mathrm{d}s < \infty,\ j = 1, 2, \cdots, n$;
(H_4) $f_j^0 < \infty,\ f_\infty^j > 0,\ j = 1, 2, \cdots, n$.
Then there exists at least one positive solution to BVP (1.1) and (1.2) for

$$(\lambda_1, \lambda_2, \cdots, \lambda_n) \in \prod_{j=1}^{n} (C_{j1}, C_{j2}). \tag{3.1}$$

Theorem 3.2 *Suppose that the conditions* $(H_1) - (H_3)$ *still hold and*
(H_5) $f_0^j > 0,\ f_j^\infty < \infty,\ j = 1, 2, \cdots, n$.
Then there exists at least one positive solution to BVP (1.1) and (1.2) for

$$(\lambda_1, \lambda_2, \cdots, \lambda_n) \in \prod_{j=1}^{n} (D_{j1}, D_{j2}). \tag{3.2}$$

Let

$$\lambda_{j1} = \frac{4^{n-1}}{f_0^j \cdot \sup\limits_{t\in[0,1]} \int_{\frac{1}{4}+\tau_j}^{\frac{3}{4}+\tau_j} G(t,s)a_j(s)\,\mathrm{d}s},\ \lambda_{j2} = \frac{4^{n-1}}{f_\infty^j \cdot \sup\limits_{t\in[0,1]} \int_{\frac{1}{4}+\tau_j}^{\frac{3}{4}+\tau_j} G(t,s)a_j(s)\,\mathrm{d}s},$$

$$\lambda_{j3} = \frac{1}{p_j \cdot M_j \int_0^1 G(\tau(s),s)a_j(s)\,\mathrm{d}s},\ \text{where } M_j = \max_{\substack{t\in[0,1] \\ \|(x_1,x_2,\cdots,x_n)\|\le 1}} f_j(t, x_1, x_2, \cdots, x_n).$$

Then, we derive a result on the multiplicity of the positive solutions for the BVP (1.1) and (1.2) furthermore in the following theorem.

Theorem 3.3 *Suppose that* $(H_1)-(H_3)$ and

(H_6) $\max\{\lambda_{j1},\lambda_{j2}\}\leq\lambda_{j3},\ j=1,2,\cdots,n;$

(H_7) $f_0^j>0,\ f_\infty^j>0$ are valid.

Then there exist at least two positive solutions to BVP (1.1) and (1.2) for

$$(\lambda_1,\lambda_2,\cdots,\lambda_n)\in\prod_{j=1}^{n}\left(\max\{\lambda_{j1},\lambda_{j2}\},\lambda_{j3}\right), \tag{3.3}$$

and at least one positive solution to BVP (1.1) and (1.2) for

$$(\lambda_1,\lambda_2,\cdots,\lambda_n)\in\prod_{j=1}^{n}\left(\min\{\lambda_{j1},\lambda_{j2}\},\max\{\lambda_{j1},\lambda_{j2}\}\right). \tag{3.4}$$

This work is supported by Science Foundation of Zhejiang Sci-Tech University (ZSTU) under Grant No.11432732611046, National Natural Science Foundation of P.R. China under Grant No.11001041 and No.10926105, and SRFDP under Grant No. 200802001008.

Author for correspondence.

References

[1] P. W. Eloe and J. Henderson, Positive solutions for (n-1,n) conjugate boundary value problems, Nonl. Anal. TMA. 28 (1997), 1669-1680.

[2] D. G. De Figueiredo, P. L. Lions and R. D. Nussbaum, A priori estimates and existence of positive solutions of semilinear elliptic equations, J. Math. Pure. Appl. 61 (1982), 41-63.

[3] H. J. Kuiper, On positive solutions of nonlinear elliptic eigenvalue problems, Rc. Mat. Circ. Palermo II (II) (1979), 113-138.

[4] L. Sanchez, Positive solutions for a class of semilinear two-point boundary value problems, Bull. Aust. Math. Soc. 45 (1992), 439-451.

[5] D. Guo, V. Lakshmikantham, Nonlinear Problems in Abstract Cones, New York: Academic Press, 1988.

[6] Q. Lu, D. Zhu, Positive solutions for an nth-order delay differential equation, Journal of East China Normal University (Nat. Sci. Edi.), 5 (2007), 20-33.

[7] W. Coppel, Disconjugacy, Lecture Notes in Mathematics, Springer: Berlin, Vol. 220, 1971.

doi:10.4028/www.scientific.net/AMM.50-51.190

Least squares skew bisymmetric solution for a kind of quaternion matrix equation *

Shifang Yuan and Handong Cao

School of mathematics and computational science, Wuyi University, Jiangmen 529020, P.R.China

ysf301@yahoo.com.cn (S.F. Yuan)

Keywords: Matrix equation; Least squares solution; Quaternion matrices

Abstract. In this paper, by using the Kronecker product of matrices and the complex representation of quaternion matrices, we discuss the special structure of quaternion skew bisymmetric matrices, and derive the expression of the least squares skew bisymmetric solution of the quaternion matrix equation AXB =C with the least norm.

Introduction

Let $R^{m\times n}$, $C^{m\times n}$, $Q^{m\times n}$, $BR^{n\times n}$, $SBR^{n\times n}$ be the sets of all $m\times n$ real matrices, all $m\times n$ complex matrices, all $m\times n$ quaternion matrices, all $n\times n$ real bisymmetric matrices, and all $n\times n$ real skew bisymmetric matrices, respectively. For $A\in C^{m\times n}$, $\mathrm{Re}(A)$ and $\mathrm{Im}(A)$ denote the real part and the imagine part of the matrix A, respectively. For $A\in Q^{m\times n}$, A^T, A^H and A^+ denote transpose matrix, conjugate transpose matrix, and Moore-Penrose generalized inverse matrix of the matrix A, respectively. Let $S_n=(e_n,e_{n-1},\cdots,e_1)$, where e_i represents the ith column of the identity matrix. $A\in Q^{n\times n}$ is called a skew bisymmetric matrix if $A^H=-A,(S_nA)^H=-S_nA$ [1]. The set of all $n\times n$ quaternion skew bisymmetric matrices is denoted by $SBQ^{n\times n}$. The Frobenius norm of $A\in Q^{m\times n}$ is denoted by $\|A\|$. For matrix $A\in Q^{m\times n}$, let $a_i=(a_{1i},a_{2i},\ldots,a_{mi})$, $(i=1,2,\ldots,n)$, and denote $vec(A)=(a_1,a_2,\ldots,a_n)^T$.

Many results have been achieved about the quaternion matrix equations. See, for example, [1-8] for more details. Matrix equations are important in many applications. For example, the well known Lyapunov equation $AX-XA^H=C$ and Stein equation $X-AXA^H=C$ play important roles in the theory and applications of stability and control. The matrix equation $AXA^H=D$ is very useful in inverse problems in vibration [9]. In this paper, we will consider the following problems.

Problem I. Given $A\in Q^{m\times n}, B\in Q^{n\times s}, C\in Q^{m\times s}$, let

$$S_L=\left\{X\mid X\in SBQ^{n\times n},\|AXB-C\|=\min_{\tilde{X}\in SBQ^{n\times n}}\left\|A\tilde{X}B-C\right\|\right\}. \tag{1}$$

Find $\hat{X}\in S_L$ such that

$$\|\hat{X}\|=\min_{X\in S_L}\|X\|. \tag{2}$$

The solution $\hat{X}$ of Problem I is called the least squares skew bisymmetric solution of the quaternion matrix equation $AXB=C$ with the least norm.

*This work is supported by Natural Science Fund of China (No.61070150), Guangdong Natural Science Fund of China (No.10452902001005845), and Program for Guangdong Excellent Talents in University, Guangdong Education Ministry, China (LYM08097) .

The structure of $SBQ^{n\times n}$

We begin with the structure of quaternion skew bisymmetric matrices.

A quaternion a can be uniquely expressed as $a=a_0+a_1i+a_2j+a_3k$ with real coefficients a_0, a_1,a_2,a_3, and $i^2=j^2=k^2=ijk=-1$, and a can be uniquely expressed as $a=c_1+c_2j$, where c_1 and c_2 are complex numbers. For any $A\in Q^{m\times n}$, A can be uniquely expressed as $A=A_1+A_2j$, where A_1, $A_2\in C^{m\times n}$. The complex representation matrix of $A\in Q^{m\times n}$ is denote by

$$f(A)=\begin{pmatrix} A_1 & A_2 \\ -\overline{A}_2 & \overline{A}_1 \end{pmatrix}\in C^{2m\times 2n}, \tag{3}$$

and $f(A)$ is uniquely determined by A. For $A\in Q^{m\times n}, B\in Q^{n\times s}$, we have $f(AB)=f(A)f(B)$. For any $B=B_0+B_1i+B_2j+B_3k\in SBQ^{m\times n}$, we have $B_0\in SBR^{n\times n}$, $B_i\in BR^{n\times n}$, $(i=1,2,3)$.

Definition 1. For $A\in R^{n\times n}$, let $a_1=(A(1,1),\sqrt{2}A(2:n-1,1)^T,A(n,1))$,

$a_2=(A(2,2),\sqrt{2}A(3:n-2,2)^T,A(n-1,2))$, …, $a_k=A(k:n-k+1,k)^T$, $a_{k+1}=A(k+1,k+1)$,

denote by the following vector containing the entries of matrix A:

$$vec_B(A)=\sqrt{2}(a_1,a_2,\ldots,a_k)^T, \quad \text{when } n=2k, \tag{4}$$

$$vec_B(A)=\sqrt{2}(a_1,a_2,\ldots,a_k,a_{k+1})^T, \quad \text{when } n=2k+1. \tag{5}$$

Definition 2. For $A\in R^{n\times n}$, denote by the following vector containing the entries of matrix A:

$$vec_A(A)=2[A(2:n-1,1)^T,A(3:n-2,2),\ldots,A(k:n-k+1,k-1)]^T, \quad \text{when } n=2k, \tag{6}$$

$$vec_A(A)=2[A(2:n-1,1)^T,A(3:n-2,2),\ldots,A(k:n-k+1,k-1),A(k+1,k)], \quad \text{when } n=2k+1. \tag{7}$$

Lemma 1. Suppose $A\in R^{n\times n}$, and $vec_B(A)$ is in the form of (Eq. 4) and (Eq. 5). Then

$$A\in BR^{n\times n}\Leftrightarrow vec(A)=B_n vec_B(A), \tag{8}$$

where

$B_{2k}=$

$$\frac{1}{2}\begin{pmatrix}
\sqrt{2}e_1 & \cdots & e_k & e_{k+1} & \cdots & \sqrt{2}e_n & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \\
0 & \cdots & 0 & 0 & \cdots & 0 & \sqrt{2}e_2 & \cdots & e_k & e_{k+1} & \cdots & \sqrt{2}e_{n-1} & \cdots & 0 & 0 \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & \vdots \\
0 & \cdots & e_1 & e_n & \cdots & 0 & 0 & \cdots & e_2 & e_{n-1} & \cdots & 0 & \cdots & \sqrt{2}e_k & \sqrt{2}e_{k+1} \\
0 & \cdots & e_n & e_1 & \cdots & 0 & 0 & \cdots & e_{n-1} & e_2 & \cdots & 0 & \cdots & \sqrt{2}e_{k+1} & \sqrt{2}e_k \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & \vdots \\
0 & \cdots & 0 & 0 & \cdots & 0 & \sqrt{2}e_{n-1} & \cdots & e_{k+1} & e_k & \cdots & \sqrt{2}e_2 & \cdots & 0 & 0 \\
\sqrt{2}e_n & \cdots & e_{k+1} & e_k & \cdots & \sqrt{2}e_1 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & 0
\end{pmatrix}$$

$$B_{2k+1}=\frac{1}{2}\begin{pmatrix}
\sqrt{2}e_1 & e_2 & \cdots & e_{k+1} & \cdots & e_{2k} & \sqrt{2}e_n & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
0 & e_1 & \cdots & 0 & \cdots & e_n & 0 & \sqrt{2}e_2 & \cdots & e_{k+1} & \cdots & \sqrt{2}e_{2k} & \cdots & 0 \\
\vdots & \vdots & & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\
0 & 0 & \cdots & e_1+e_n & \cdots & 0 & 0 & 0 & \cdots & e_2+e_{n-1} & \cdots & 0 & \cdots & 2e_{k+1} \\
\vdots & \vdots & & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\
0 & e_n & \cdots & 0 & \cdots & e_1 & 0 & \sqrt{2}e_{n-1} & \cdots & e_{k+1} & \cdots & \sqrt{2}e_2 & \cdots & 0 \\
\sqrt{2}e_n & e_{n-1} & \cdots & e_{k+1} & \cdots & e_2 & \sqrt{2}e_1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0
\end{pmatrix}.$$

Lemma 2. For $A \in R^{n\times n}$, $vec_A(A)$ is in the form of (Eq. 6) and (Eq. 7). Then

$$A \in SBR^{n\times n} \Leftrightarrow vec(A) = W_n vec_A(A), \tag{9}$$

where

$$W_{2k} = \frac{1}{2}\begin{pmatrix} e_2 & \cdots & e_k & e_{k+1} & \cdots & e_{n-1} & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ -e_1 & \cdots & 0 & 0 & \cdots & -e_n & e_3 & \cdots & e_k & e_{k+1} & \cdots & e_{n-2} & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & \cdots & -e_1 & -e_n & \cdots & 0 & 0 & \cdots & -e_2 & -e_{n-1} & \cdots & 0 & \cdots & -e_{k+1} & -e_{k+2} \\ 0 & \cdots & -e_n & -e_1 & \cdots & 0 & 0 & \cdots & -e_{n-1} & -e_2 & \cdots & 0 & \cdots & -e_{k+2} & -e_{k+1} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ -e_n & \cdots & 0 & 0 & & -e_1 & e_{n-2} & \cdots & e_{k+1} & e_k & \cdots & e_3 & \cdots & 0 & 0 \\ e_{n-1} & \cdots & e_{k+1} & e_k & \cdots & e_2 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \end{pmatrix},$$

$$W_{2k+1} = \frac{1}{2}\begin{pmatrix} e_2 & \cdots & e_{k+1} & \cdots & e_{2k} & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ -e_1 & \cdots & 0 & \cdots & -e_n & e_3 & \cdots & e_{k+1} & \cdots & e_{2k-1} & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & -e_1 - e_n & \cdots & 0 & 0 & \cdots & -e_2 - e_{n-1} & \cdots & 0 & \cdots & -e_k - e_{k+2} \\ \vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ -e_n & \cdots & 0 & \cdots & -e_1 & e_{n-2} & \cdots & e_{k+1} & \cdots & e_3 & \cdots & 0 \\ e_{n-1} & \cdots & e_{k+1} & \cdots & e_2 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}.$$

For $A = A_1 + A_2 j \in Q^{n\times n}$, we denote an identification by the symbol $\cong$, that is, $A \cong \widetilde{A} = (A_1, A_2)$,

$$vec(A_1) + vec(A_2)j = vec(A) \cong vec(\tilde{A}) = \begin{pmatrix} vec(A_1) \\ vec(A_2) \end{pmatrix}, \quad \| vec(A) = \| vec(\widetilde{A}) \| = \left\| \begin{pmatrix} vec(A_1) \\ vec(A_2) \end{pmatrix} \right\|.$$

We denote $\vec{A} = (\mathrm{Re}(A_1), \mathrm{Im}(A_1), \mathrm{Re}(A_2), \mathrm{Im}(A_2))$,

$$vec(\vec{A}) = \begin{pmatrix} vec(\mathrm{Re}(A_1)) \\ vec(\mathrm{Im}(A_1)) \\ vec(\mathrm{Re}(A_2)) \\ vec(\mathrm{Im}(A_2)) \end{pmatrix}, \quad vec_S(\vec{A}) = \begin{pmatrix} vec_A(\mathrm{Re}(A_1)) \\ vec_B(\mathrm{Im}(A_1)) \\ vec_B(\mathrm{Re}(A_2)) \\ vec_B(\mathrm{Im}(A_2)) \end{pmatrix}.$$

Notice that $\| vec(\tilde{A}) = \| vec(\vec{A}) \|$.

Lemma 3 [7] . If $A = A_1 + A_2 j \in Q^{m\times n}, B = B_1 + B_2 j \in Q^{n\times s}$, and $C = C_1 + C_2 j \in Q^{s\times t}$, then

$$vec(\widetilde{ABC}) = (f(C)^T \otimes A_1, f(Cj)^H \otimes A_2)\begin{pmatrix} vec(\widetilde{B}) \\ vec(-\widetilde{jBj}) \end{pmatrix}. \tag{10}$$

Lemma 4. If $A = A_1 + A_2 j \in Q^{m\times n}, X = X_1 + X_2 \in SBQ^{n\times n}, B = B_1 + B_2 \in Q^{n\times s}$, B_n, W_n are defined in the form (Eq. 8) and (Eq. 9), respectively. The vectors $vec_B(A)$ and $vec_A(A)$ are defined by Definition 1 and Definition 2, respectively. Then

$$vec(\widetilde{AXB}) = (f(B)^T \otimes A_1, f(Bj)^H \otimes A_2)\begin{pmatrix} W_n & iB_n & 0 & 0 \\ 0 & 0 & B_n & iB_n \\ W_n & -iB_n & 0 & 0 \\ 0 & 0 & B_n & -iB_n \end{pmatrix}\begin{pmatrix} vec_A(\mathrm{Re}(X_1)) \\ vec_B(\mathrm{Im}(X_1)) \\ vec_B(\mathrm{Re}(X_2)) \\ vec_B(\mathrm{Im}(X_2)) \end{pmatrix}. \tag{11}$$

The solution of Problem I and Problem I

Based on our earlier discussions, we now turn our attention to Problem I. The following notations are necessary for deriving the solutions of Problem I.

For $A = A_1 + A_2 j \in Q^{m\times n}$, $B \in Q^{n\times s}$, $C \in Q^{m\times s}$, set $Q_1 = (f(B)^T \otimes A_1, f(Bj)^H \otimes A_2)$,

$$Q = Q_1 \begin{pmatrix} W_n & iB_n & 0 & 0 \\ 0 & 0 & B_n & iB_n \\ W_n & -iB_n & 0 & 0 \\ 0 & 0 & B_n & -iB_n \end{pmatrix}, \quad P_1 = \mathrm{Re}(Q), P_2 = \mathrm{Im}(Q), \quad e = \begin{pmatrix} vec(\mathrm{Re}(\widetilde{C})) \\ vec(\mathrm{Im}(\widetilde{C})) \end{pmatrix},$$

and

$$R = (I - P_1^+ P_1)P_2^T, \quad H = R^+ + (I - R^+R)ZP_2P_1^+(P_1^+)^T(I - P_2^T R^+),$$

$$Z = [I + (I - R^+R)P_2P_1^+(P_1^+)^T P_2^T(I - R^+R)]^{-1}, \quad S_{11} = I - P_1^+P_1 + (P_1^+)^T P_2^T Z(I - R^+R)P_2P_1^+,$$

$$S_{12} = -(P_1^+)^T P_2^T(I - R^+R)Z, \quad S_{22} = (I - R^+R)Z.$$

From the results in [10], we have

$$(P_1, P_2)^+ = \begin{pmatrix} P_1^+ - P_1^+P_2H \\ H \end{pmatrix}, \qquad I - (P_1, P_2)^+(P_1, P_2) = \begin{pmatrix} S_{11} & S_{12} \\ S_{12}^T & S_{22} \end{pmatrix}.$$

Theorem 1. Given $A \in Q^{m\times n}, B \in Q^{n\times s}, C \in Q^{m\times s}$, B_n, W_n are defined in the form (Eq. 8) and (Eq. 9), $W = diag(W_n, B_n, B_n, B_n)$, respectively. $A = A_1 + A_2 j, A_1, A_2 \in C^{m\times n}$. Then the set S_L of Problem I can be expressed as

$$S_L = \left\{ X \mid vec(\vec{X}) = W(P_1^+ - H^T P_2 P_1^+, H^T)e + W(I - P_1^+P_1 - RR^+)y \right\}, \tag{12}$$

where y is an arbitrary vector of appropriate order.

Proof. By Lemma 4,

$$\| AXB - C \| = \| vec(\widetilde{AXB}) - vec(\widetilde{C}) \|$$

$$= \left\| \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} vec_S(\vec{X}) - e \right\|,$$

we have $vec_S(\vec{X}) = (P_1^+ - H^T P_2 P_1^+, H^T)e + (I - P_1^+P_1 - RR^+)y$. The proof is completed.

Theorem 2. Problem I has a unique solution $\widehat{X} \in S_L$, and $\widehat{X}$ can be expressed as

$$vec(\vec{\widehat{X}}) = W(P_1^+ - H^T P_2 P_1^+, H^T)e. \tag{13}$$

References

[1] Q.W. Wang: Bisymmetric and centrosymmetric solutions to system of real quaternion matrix equation, Computers Math. Applic. 49 (2005), pp. 641- 650.

[2] Y.H. Au-Yeung, C.M. Cheng: On the pure imaginary quaternionic solutions of the Hurwitz matrix equations, Linear Algebra and its Applications 419 (2006), pp.630-642.

[3] Q. W. Wang, S .W. Yu, C.Y. Lin: Extreme ranks of a linear quaternion matrix expression subject to triple quaternion matrix equations with applications, Applied Mathematics and Computation, 195 (2008), pp.733-744

[4] T. S. Jiang, M. S. Wei: Equality constrained least squares problem over quaternion field, Applied Mathematics Letters, 16 (2003) , pp. 883-888.

[5] T.S. Jiang, M.S. Wei: Real representations of quaternion matrices and quaternion matrix equations, Acta Math Sci (China) 26 A (2006), pp. 578-584.

[6] Yao tang Li and Wen jing Wu. Symmetric and skew-antisymmetric solutions to systems of real quaternion matrix equations. Computers Mathematics with Applications, 55 (2008), pp. 1142-1147.

[7] S.F. Yuan, A.P. liao, Y. Lei: Least squares Hermitian solution of the matrix equation $(AXB, CXD) = (E, F)$ with the least norm over the skew field of quaternions, Mathematical and Computer Modelling 48(2008), pp. 91-100.

[8] S. F. Yuan, A.P. Liao, Y. Lei: Minimization problem for Hermitian matrices over the quaternion field, Mathematica Acta Scientia, 29A (2009), pp. 1298-1306. (in Chinese)

[9] H. Dai and P. Lancaster: Linear matrix equations from an inverse problem of vibration theory, Linear Algebra and Its Applications, 246 (1996), pp. 31-47.

[10] J.R. Magnus: L-structured matrices and linear matrix equations, Linear and Multilinear Algebra 14 (1983), pp. 67-88.

doi:10.4028/www.scientific.net/AMM.50-51.195

Study on the Adaptability of Staining method for Acoustic Field Measuring in Ultrasonic Processing

K.H.Zhang[1, a], G.F.Wang[1, 2, b], S.Lu[1, c]

[1] Institute of Precision Mechanism & Instrument, Zhejiang Normal University, JinHua, 321004, China

[2] Lanzhou Jiaotong University, Lanzhou, 730070, China

[a]mature@126.com, [b]wgf_okgood@163.com, [c]lushuang@zjnu.cn,

Keywords: Ultrasonic Processing, Staining method, Cavitation

Abstract: In order to search for an ideal method to measure the acoustic field, the staining method was studied. First of all, the experiment device was designed to produce standing wave, and then the coated papers were inserted into the methylene blue solution. At last the experiment result was analyzed. It illustrated that in the power range of about *10w* to *250 w*, staining method was a good one which could described the distribution of acoustic pressure in container which used for acoustic processing or sonochemistry.

1 Introduction

In acoustic machining or sonochemistry, how to measure the distribution of the total acoustic field in the machining or the reaction container was very important. So far there have appeared many methods based on various kinds of principles, such as hydrophone method, heat-sensitive sensor method [1], aluminum foil corrosion method [2], the acoustic emission spectrometry and acoustical holography [3, 4]. However each acoustic field measuring technique has its advantages, disadvantages and application scope, for example: Hydrophone method can't adept to low frequency and large-power acoustic field. In aluminum foil corrosion method, the film on the aluminums foil surface is weak and could be broken and peeled easily. When we want to get the total acoustic pressure of the processing container using hydrophone, we have to put lots of test devices to every test points, as is inconvenient and costs a large of money. Since the influence of the test devices put into the acoustic field, the result may be inaccurate. So it is necessary to search for new solutions which could be easily in measuring the total ultrasonic field in container.

Many researchers have evaluated the distribution of acoustic field in liquid by means of staining method in container. In the ultrasonic study, the coated paper was put in methylene blue to dye, and by means of observing the quantity of dye spots and the dye depth to infer the acoustic field intensity or the cavitation intensity.

As to the dyeing mechanism, previous researchers tend to believed that the sound medium vibration, cavitation and a series of other complex factors led to the damage in the surface of the paper, adsorption and permeation made dye molecules attach to the fibers of the paper, as is a fatigue damage process.

As an easy-to-use testing method, the dyeing method is precise to a certain extent. In this paper, we wanted to give an illustration on the feasibility and effectiveness of the staining method through experiment.

2 Experimental results and discussion

Acoustic field with standing wave was regular and it has predictable acoustic intensity. For easy to study, it's necessary to get standing wave in our experiment. It is generally known that the standing wave could be obtained using waveguide.

The experimental device was designed by us to generate stable standing wave, as show in Fig. 1 and Fig. 2. The resonance frequency was 28.2 kHz, and the rated power of acoustic transducer was 55w. Chose the UGD ultrasonic source which could keep frequency automatic tracking and have an adjustable power. The container was a cylinder made of organic glass (PC) with an 8cm diameter.

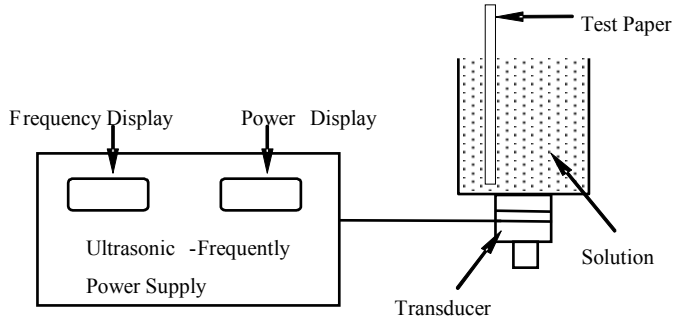


Fig.1 The schematic diagrammatic of experiment

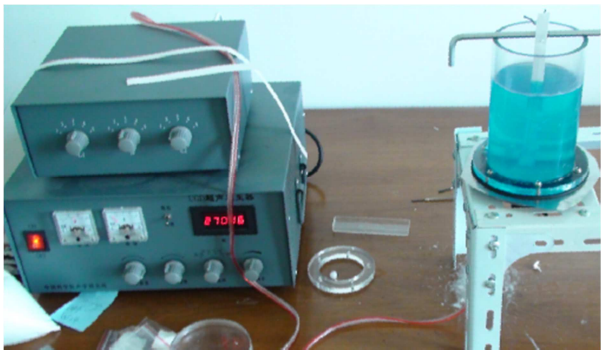


Fig.2 the photo of experimental device

According to cut-off frequency formula of cylindrical waveguide [4]:

$$f_c = f_{10} = 1.84\frac{C_0}{2\pi a} \tag{1}$$

Where c_0 is acoustic speed in the medium, and a is radius of cylindrical waveguide. Through calculation, we found that the acoustic was complex. There was a serious of high order harmonics expect for the main wave (0, 0). But because the fundamental wave was dominant and had little attenuation, we could consider that there was only plate wave propagating along the axis of the container.

Under ultrasonic condition, inserted the test paper in methylene blue solution， and take it out several minutes later. There would be blue-and-white strips on the test paper (Fig. 3).And the blue parts were composed of many discrete spots of blue color.

The distance of node to the surface of water is $H = (2n+1)*\lambda/4$. Through calculation, there are nodes when the value of H is *1.31, 3.94, 6.56, and 9.18* in the experiment with resonant frequency

28.2 kHz. The white parts and the blue parts corresponded to nodes and antinodes region respectively in figure 5. Customarily researchers indicated the sound field intensity by means of the dye deepness.

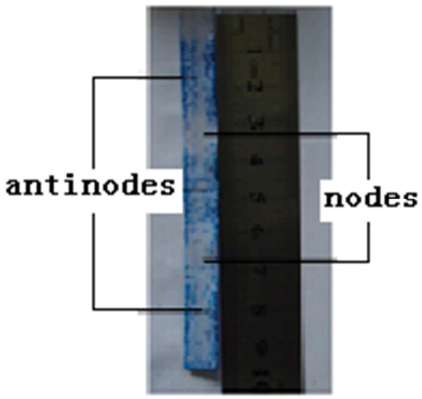


Fig.3 Photograph of a dyeing stripe on the test paper

The minimum value of acoustic pressure or acoustic intensity that could lead to cavitation was called cavitation threshold. Suppose the static pressure in the liquid is p_0, alternating acoustic pressure amplitude is P_m. Only if $P_m > P_0$, negative pressure arise. If negative pressure exceeds the structure intensity, cavitation will formed. Cavitation threshold P_c can be expressed according to this equation:

$$p_c = p_0 - p_v + \frac{2}{3\sqrt{3}}\sqrt{\left(\frac{2T}{R_0}\right)^3 \Big/ \left(p_0 - p_v + \frac{2T}{R_0}\right)} \tag{2}$$

where P_0 is Hydrostatic pressure; P_x is the vapour pressure; T is surface tension; R_0 is The initial radius of cavitation nuclear. The relationship between acoustic pressure and acoustic intensity could be formulated bellow:

$$p = \sqrt{\mathrm{I}\rho\mathrm{c}} \tag{3}$$

where I is acoustic intensity; ρ is the density of the media; c is the speed of ultrasound. With the input power of ultrasound increasing, the time to achieve the same dye depth reduced sharply. It just needed 1-2s to get dyeing spots when there is enough acoustic power. And at the same time there was dyeing strips as well.

Raising the input power to some extent, the fountain appeared on the water surface, which derived from acoustic stream phenomenon (Fig. 4). The axial speed of acoustic stream was gave by C.Eckart[5].

$$G(r) = G(r_1) = \frac{1}{2} K P_0^2 r_1^2 \tag{4}$$

where r_1 is the radius of radiant end of ultrasonic transducer.

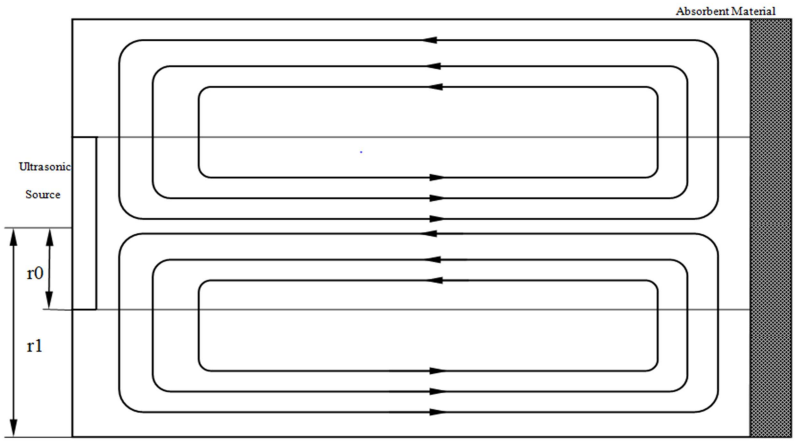
Under this condition the test paper dyed quickly and was damaged after a short time; however there were also stripes which indicated the acoustic intensity. The acoustic stream promoted the flow of acoustic medium and the movement of cavitation nucleus in fluid but wasn't able to effect the distribution of acoustic pressure. So it's feasible to use staining method for measuring the acoustic field under high power and no-linear condition.

Sometimes the blue stripes on test paper can't correspond to the acoustic antinodes, as can be explain bellow: in certain instances that the main wave couldn't reach the cavitation threshold but

the high order harmonic in acoustic field led to cavitation at the position of nodes.

According to previous study, cavitation included steady cavitation and transient one. There was mainly steady cavitation in the nodal regions, and in the antipodal regions was mainly transient cavitation. The test paper could be dyed quickly on the antinodes regions because the cavitation bubbles collapsed frequently here. Effected by acoustic stream, high order harmonic, steady and transient cavitation, it was difficult to infer an equation which could describes the relationship between acoustic intensity and the dyeing deepness or the quantity of dyeing spots.

(a) Photo of Acoustic stream [5]



(b) Schematic drawing of acoustic stream [6]

Fig. 4 Schematic drawing of acoustic field

Lowered the input power to *10w,* when the cavitation threshold was not reached, and the dyeing spots were also occurred, but it took quite a long time. According to the mechanism of staining method, the dyeing spots came from the fatigue damage in the surface of coated paper. So it was rational that the formation of dyeing spots need more time on the condition of low frequency.

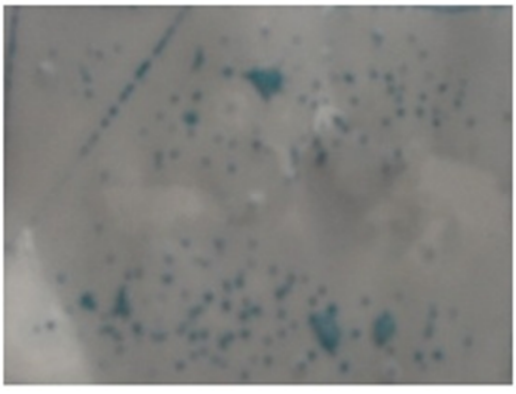


Fig.5 Wide test paper could indicate the total distribution of acoustic field in the container.

The coated paper was cut into the size and shape of vertical section of the PC container, and was inserted into the solution to dye. Several minute later, there appeared a picture on the test paper, and the dyeing depth indicated the acoustic intensity (Fig. 5). In contrast to the experiment result which came from hydrophone measurement, overall, it is exact. At this stage of the experiment we could come to the conclusion that the coated paper had little influence on the original acoustic field. The significance was that the experiment gave us an approach which could achieve the visualization of measurement result. We could use the wide coated paper to examine the uniformity of acoustic field, and predict the acoustic intensity.

4 Conclusions

Staining method could be used in a quite wide acoustic intensity scope, and it can be easily operated and could get a visual result. Learn from the experiment, the test paper could illustrate the whole distribution of the acoustic pressure by figures in container. So the staining method is an ideal method to test acoustic intensity in acoustic machining and acoustic chemistry, and could also provide references to improve ultrasonic cleaning process.

Staining method is suitable for the use of a wide range of 10w to 250w and even larger acoustic power. And we can get the conclusion that straining method has many barriers in pullulating to a utility technology to evaluate the acoustic field quantitatively. So we couldn't make use of the method to carry on accurate measurement, but it is simple and easy operating in special acoustic field which requires less accuracy.

Acknowledgement

The authors gratefully acknowledge the funding for this work from Natural Science Foundation of Zhejiang Province (Grant No. Y1091122).

References

[1] C.J.Martin, A.N.R.Law: The use of thermistor probes to measure energy distribution in ultrasound fields [J].Ultrasonics, 1980, 18(3): 127-133.

[2] E.A.Neppiras: Measurement of acoustic cavitation [J], Sonics Ultrason, 1986: 81-89.

[3] V S.Moholkar, S.P.Sable: Mapping the cavitation intensity in an ultrasonic bath using the acoustic emission, 2000, 46(4):684-194

[4] HU Shu-fang, WU Sheng-ju,Experiments on ultrasonic standing wave field in liqued, Journal of Shaanxi Normal University(Natural Science Edition), Vol.38 No.1 Jan.2010

[5] LI Hua-mao, Sound-streaming experiment, Technical Acoustics, 2004, 23(z2)

[6] Qian Zhuwen, Nonlinear Acoustics [M], 1992.

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.200

Simulations of Natural Vibration Characteristics of Large-size Radial Gate with Three Arms Considering Fluid-solid Coupling

Jiliang Liu [1, a], Xing Fang [2, b], Zhengzhong Wang [1, c], and Danxia Sun[3]

[1]Research Center of Water Engineering Safety and Disaster Prevention, Northwest A & F University, Yangling, Shaanxi, China,712100

[2]Department of Civil Engineering, Auburn University, Auburn, Alabama, USA

[3]Northwest Investigation Design and Research Institute of CHECC, Xi'an, Shaanxi, China, 710065

[a]liujiliang430@163.com, [b]xing.fang@auburn.edu, [C]wangzz0910@163.com

Keywords: Radial gate, Natural vibration characteristics, Fluid-solid coupling, Finite element method, ADINA software

Abstract. When orifice opening of flow under gate through spillways becomes large, large-size radial gates with three radial struts (arms) are normally recommended and have been used in various hydraulic engineering structures/projects. However, radial gates with three arms are easier to have operational failure caused by severe vibrations during opening and closing gate process than those with two arms. In this paper, a finite element numerical model was developed using ADINA software to simulate and analyze natural vibration characteristics of a radial gate with three arms at the Shuhe Hydropower Station using fluid-solid coupling theory. Simulation results of natural vibration characteristics for the gate were also developed without considering fluid-solid coupling. The model results indicate that fluid-solid coupling has remarkable effect on simulated natural vibration characteristics and can not be ignored. Simulated natural frequencies of the same mode are increased with the increase of openings, but the change amplitudes are not significant. This study provides useful information for structural dynamic design of large-size radial gates in order to avoid the natural frequency of radial gate close to the main pulsation frequency of water flow that may cause resonance.

Introduction

Radial gates are the most widely used spillway gates in hydraulic engineering projects because they are the simplest, most reliable, and usually the least expensive type of crest gates. They require considerably smaller forces during maneuvering compared with other types of gates, have more reliable seating on the sill, require less space, and have better conditions for the passage of ice and wood. Safe operation of radial gate is very important for any hydraulic structures and dams. However, for many radial gates installed and operated in many hydraulic structure projects, vibrations of the radial gates usually occur during opening and closing gate process. Under certain conditions, the vibration is very severe and can cause an operational failure accident, e.g., more than 20 failure accidents of radial gates have happened in China since 1960s [1]; the radial gate of Hezhi Dam in Japan also had failure accident. Many investigations show that failure accidents of radial gates are due to special hydrodynamic loads and dynamic characteristics of the gates, i.e., natural vibration characteristics. In fact, the dynamic response and dynamic stability of a radial gate under dynamic loads depend on its natural vibration characteristics. Therefore, it is necessary and important to analyze natural vibration characteristics of a radial gate in order to provide the reliable and useful information for structural dynamic design of the radial gate. As pointed out by several studies [2-7], the vibrations of radial gates normally happen due to water flow acting on the gates, and the vibrations are called flow-induced vibrations. Therefore, it is necessary and essential to consider fluid-solid coupling when analyzing natural vibration characteristics of radial gates.

In recent years, there are many hydraulic engineering projects having large gate or orifice opening of flow through spillways, and the radial gates with two arms (struts) are not strong enough

to adapt to this situation especially when height-span ratio is more than 1. Large-size radial gates with three arms are normally recommended to use when gate-orifice opening is large. However, Zhang and Liu [1] pointed out that radial gates with three arms were easier to have operational failure caused by severe vibrations than those with two arms. Therefore, the natural vibration characteristics of radial gates with three arms should be analyzed in order to provide useful information for structural dynamic design of this type of radial gates.

Based on fluid-solid coupling theory, a radial gate with three arms installed and operated in the Shuhe Hydropower Station was used as prototype to develop a finite element numerical model for this study. The software ADINA was used to develop the model and analyze natural vibration characteristics of the gate under different openings. Natural vibration characteristics were also analyzed by developing another numerical model without considering fluid-solid coupling. This study provides useful information for dynamic design of large-size radial gates in order to avoid the natural frequency of radial gate close to the main pulsation frequency of water flow that may cause resonance.

Materials and Methods

Fluid-solid coupling theory used to obtain natural vibration characteristics can be found in published papers and textbooks [3, 8]. The radial gate studied here is one of the three-arm radial gates installed and operated at the Shuhe Hydropower Station, which is located on the Hanjiang River with the total hydropower capacity of 270 MW. There are five spillways with the orifice size of 13 m×24.3 m (the span × the height). Because the orifice size is very large and the height-span ratio is 1.87 (>1), radial gates with three arms of which the cross sections are box shape were installed. The elevations of bottom sill and support pivot are 193.5 m and 217.6 m above the mean sea level, respectively. The radius of the curved panel is 32 m and the designed head is 23.8 m above the sill. The radial gates are operated by hydraulic hoist with the capacity of 2×4000 KN. The dimensions of the radial gate are shown in Fig. 1.

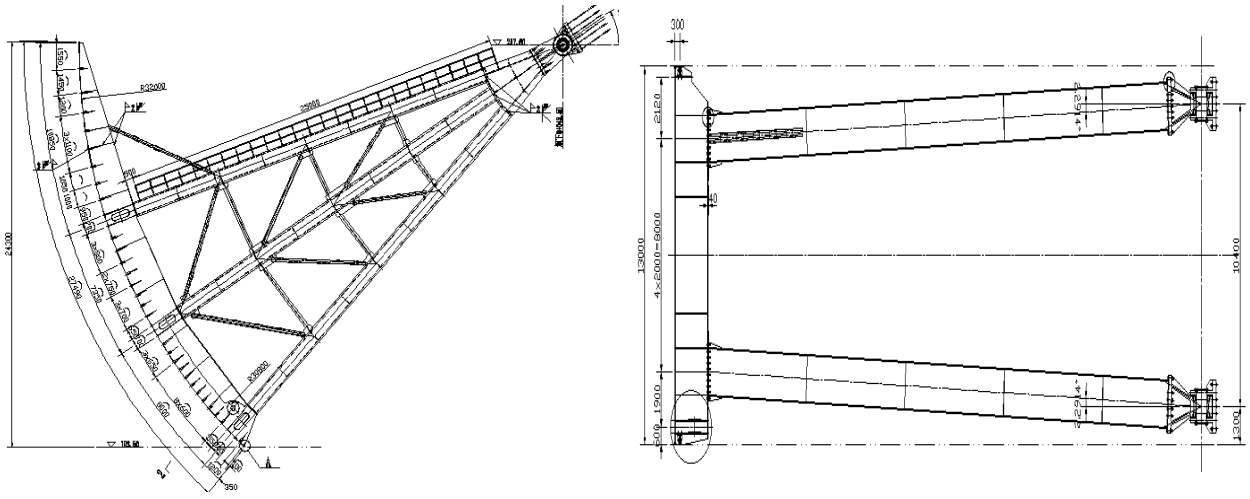


Fig. 1 Dimensions of Radial Gate with Three Arms Used in Finite Element Modeling Simulation

These five radial gates usually have operational vibrations during opening and closing gate process, therefore, it is necessary and essential to analyze their natural vibration characteristics in order to make sure that the radial gates are operated safely. We used ADINA software to develop the finite element numerical model, which is shown in Fig. 2, to simulate and analyze natural vibration characteristics of the radial gate under different openings. The numerical model consists of shell element, beam element, 3-D solid element, and truss element, and fluid region is divided by 3-D Fluid element. The 3-D fluid elements in the front of the gate are not shown in Fig. 2. Summary of number of elements for the finite element model is given in Table 1.

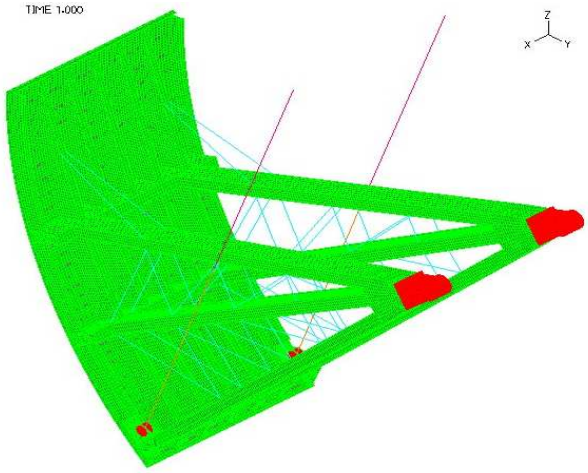


Fig. 2 Finite Element Model of the Radial Gate with Three Arms

Table 1 Element Types and Number of Elements Used in the Model

Element types	Shell	3-D Solid	Beam	Truss	3-D Fluid	Nodes
Number of elements	51445	39328	1810	54	4000	119350

Results and Discussions

Using the finite element numerical model, we obtained natural frequencies of the radial gate at the gate openings of 0.0 m, 0.5m, 1.0 m, 2.0 m, and 5.0 m. According to modal analysis theory, the low-order modes affect the safety of the radial gate much more than higher-order modes do, therefore, we only listed simulated natural frequencies of the first five modes in Table 2. Natural vibration characteristics were also analyzed by developing another numerical model without considering fluid-solid coupling (just like Fig. 2 without fluid elements) in order to compare with the results obtained by considering fluid-solid coupling, and the natural frequencies of the first five modes are shown in Table 3.

After analyzing simulation results from the finite element models, we have obtained the following information:

(1) Tables 2 and 3, and Figs. 3 and 4 show that fluid-solid coupling has remarkable effect on simulated natural vibration characteristics, and simulated natural frequencies with the consideration of fluid-solid coupling are smaller than those obtained without considering fluid-solid coupling. Therefore, fluid-solid coupling should be considered in the analysis of natural vibration characteristics of radial gates.

(2) The natural frequencies of the same mode are increased with the increase of openings, but the change amplitudes are not significant.

(3) With considering fluid-solid coupling, the curved skin plate or panel of the radial gate was found to have severe vibration, but the gate arms do not. It is because that the gate arms have box shape cross section instead of typical I-section and are designed with sufficient stiffness.

Simulated maximum natural frequency of radial gate in Shuhe Hydropower Station by considering fluid-solid coupling is 4.634 Hz. According to the investigation on pulsation frequency of water flow [9], for about 90 percent of radial gates, the main pulsation frequencies of water flow are between 1 Hz and 20 Hz; for about 48 percent of radial gates, the main pulsation frequencies of water flow are between 1 Hz and 10 Hz. Therefore, the radial gates of Shuhe Hydropower Station may have operational vibrations caused by resonance when the natural frequencies are equal to main pulsation frequencies of water flow under certain gate-opening conditions.

Because natural vibration characteristics of radial gates are affected by many factors, such as flow conditions, stiffness distribution, boundary conditions, and fluid-solid coupling, moreover, the

vibration mechanism of radial gates is very complex, up to now, there are no commonly accepted methods of structural dynamic design of radial gates especially for large-size radial gates provided for designers to use. In other words, researchers still need conduct lots of investigations (theoretical, experimental, and numerical modeling studies) in order to fully understand the vibration mechanism of radial gates.

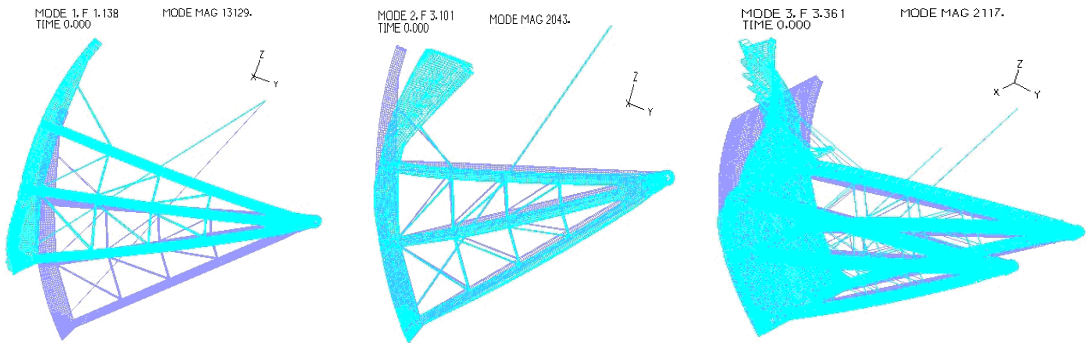
Table 2 Simulated Natural frequencies of the first 5 modes with the consideration of fluid-solid coupling

Mode	Gate openings				
	0.0 m	0.5 m	1.0 m	2.0m	5.0m
	Frequency（Hz）	Frequency（Hz）	Frequency（Hz）	Frequency（Hz）	Frequency（Hz）
1	1.032	1.087	1.138	1.246	1.617
2	2.540	2.775	3.101	3.760	4.403
3	2.735	2.999	3.361	4.026	4.468
4	3.879	3.926	3.957	4.095	4.603
5	3.945	3.991	4.022	4.102	4.634

Table 3 Simulated Natural frequencies of the first 5 modes without the consideration of fluid-solid coupling

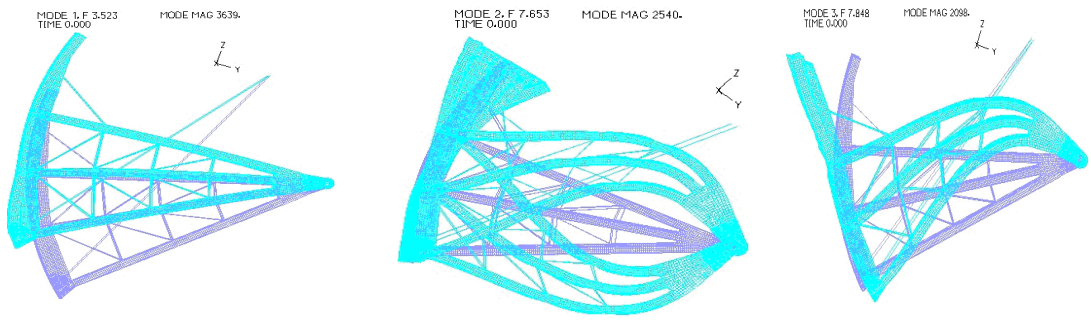
Mode	Gate openings				
	0.0 m	0.5 m	1.0 m	2.0m	5.0m
	Frequency（Hz）	Frequency（Hz）	Frequency (Hz)	Frequency（Hz）	Frequency（Hz）
1	3.434	3.48	3.523	3.605	3.819
2	7.652	7.652	7.653	7.654	7.657
3	7.847	7.848	7.848	7.85	7.855
4	9.348	9.348	9.348	9.348	9.349
5	9.434	9.435	9.435	9.435	9.436

We also obtained mode shapes and selected part of them at the opening of 1.0 m and showed them in Figs. 3 and 4.



(a) mode 1 (b) mode 2 (c) mode 3

Fig. 3 Mode Shapes at the Opening of 1.0 m with the Consideration of Fluid-solid Coupling



(a) mode 1 (b) mode 2 (c) mode 3

Fig. 4 Mode Shapes at the Opening of 1.0 m without the Consideration of Fluid-solid Coupling

Conclusions

This paper developed a finite element numerical model of a large-size radial gate with three arms using ADINA to analyze its natural vibration characteristics. The model results indicate that fluid-solid coupling has remarkable effect on natural vibration characteristic and can not be ignored, and the natural frequencies of the same mode are increased with the increase of openings, but the change amplitudes are not significant. Large-size radial gates with three arms should adopt arms with box shape cross section instead of I-section because box shape has sufficient stiffness. The model simulations show that curved panel of the radial gate has severe vibration, therefore, new designs of gate panel are needed, e.g., additional horizontal and/or vertical girders. This study provides useful and valuable information for dynamic design of large-size radial gates.

Acknowledgements

This study was supported by a state 863 project (Grant No. 2002AA62Z3191) of the P.R. China.

References

[1] J.G. Zhang, and G.R. Liu: Journal of Hydroelectric Engineering Vol. 38 (1991), p. 49-57.

[2] Z.X. Xie, and J.F. Zhou: Journal of Hohai University Changzhou Vol. 20 (2006), p. 25-28.

[3] Y.H. Wu, and X.Z. Xie: Journal of Hydraulic Engineering Vo1. 1 (1995), p. 27-33.

[4] J.J. Lian, X.M. Peng, G.T. Cui, and J.Y. Lin: Journal of Tianjin University Vol. 32 (1999), p. 171-176.

[5] N.D. Thang, and E. Naudascher: J. Hydr. Res. Vol. 24 (1986), p. 391-404.

[6] D.X. Qiu, and Z.Q. Zhu: Guangdong Water Resources and Hydropower Vol.10 (2010), p.10-12.

[7] S.W. Xian, G.H. Yan, P.Z. Ma, and S.Z. Shao: Journal of Hydroelectric Engineering Vol. 31 (1990), p.32-46.

[8] Z.Y. Ye, W.W. Z, and A.M S: *Fundamentals of Fluid-Structure Coupling and Its Application* (Harbin Institute of Technology Press, China 2010).

[9] Y.L. He: *Structural Stability Theory* (China Waterpower Press, China 1995).

doi:10.4028/www.scientific.net/AMM.50-51.205

The Influence of the Voltage Slope to the Homogeneous Multiple-pulse Barrier Discharge at Atmospheric Pressure★

Peng Xie[1,a], Yanling Liu[2, b] and Shaoyan Cui[1, c]

[1] School of Mathematics and Information, Ludong Universtiy, Yantai 264025, China

[2] Yantai Government East Service Center, Yantai 264003, China

[a]ytxiepeng@gmail.com, [b]neieplyl@163.com, [c]shycui@ldu.edu.cn

Corresponding author: ytxiepeng@gmail.com

Keywords: Homogeneous, Atmospheric-pressure, Barrier discharge, Numerical simulation, Plasma

Abstract. The homogeneous barrier discharge with multiple current pulses at atmospheric pressure in helium is numerically studied on the basis of one-dimensional fluid model .The influence of the applied periodic polygonal function voltage slope to the intervals of the corresponding current pulses and the power are discussed and analyzed. The simulation results show that ,in the case of invariable voltage amplitude and variable voltage slope, the number of current pulses , intervals of the corresponding current pulses and the operating factor of the power are changing.

Introduction

The homogeneous plasma at atmospheric pressure has been the research focus in the field of plasma due to the broad prospect of application [1]. A multi-pulse dielectric barrier discharge at atmospheric pressure has been reported in recent years, but the mechanism of this discharge is still not entirely clear [3-6]. In order to further research, on the basis of one-dimensional fluid model ,this paper siplified the applied voltage, studied the homogeneous barrier discharge with multiple current pulses, analyzed in detail the discharges parameters on multi-pulse discharge effects.

Theoretical model

The discharge is between two parallel plates electrodes, the specific discharge structure is from Reference [2-3]. Two plates were covered by dielectric layer,the voltage applied between two electrodes is periodic line function.the Numerical simulation is based on one-dimensional fluid model.the Charged particles and the transport of the Metastability atoms are discribed by continuity equation and momentum equation:

$$\frac{\partial n_i}{\partial t}+\frac{\partial j_i}{\partial x}=s_i, \tag{1}$$

$$j_i=\omega_i n_i-D_i\frac{\partial n_i}{\partial x}, \tag{2}$$

Where n_i, j_i, ω_i and D_i are the Particle density, Particle flow density, drift velocity and diffusion coefficient, s_i is the Net produce items.

Electric field determined by the current balance equation:

$$\varepsilon_0\varepsilon(x)\frac{\partial E(x,t)}{\partial t}+j(x,t)=j_0(t), \tag{3}$$

Where j is Conduction current density,is given by:

$$j(x,t) = e(j_+(x,t) - j_e(x,t)), \tag{4}$$

Where $j_+(x,t)$, $j_e(x,t)$ are ion and electron density,is given by （2）. $\varepsilon(x)$ is Relative permittivity,which depends on x.In the Media, $\varepsilon(x) = \varepsilon_1 = 7.5$, In the gas, $\varepsilon(x) = \varepsilon_0 = 1$, the Electric field Meets the following condition :

$$-\int_0^{d_s} E(x,t)dx = V(t), \tag{5}$$

Where d_s is the distance between the two electrodes , $V(t)$ is the periodic line function on the plate:

$$V(t) = \begin{cases} k \bmod(\frac{t}{100tal}, \frac{4V_0}{k}), 0 \le \bmod(\frac{t}{100tal}, \frac{4V_0}{k}) < \frac{V_0}{k} \\ -k \bmod(\frac{t}{100tal}, \frac{4V_0}{k}) + 2V_0, \frac{V_0}{k} \le \bmod(\frac{t}{100tal}, \frac{4V_0}{k}) < \frac{3V_0}{k} \\ k \bmod(\frac{t}{100tal}, \frac{4V_0}{k}) - 4V_0, \frac{3V_0}{k} \le \bmod(\frac{t}{100tal}, \frac{4V_0}{k}) < \frac{4V_0}{k} \end{cases}, \tag{6}$$

Where j_0 is discharge current density, which is integrated by (3):

$$j_0(t) = (\frac{2d_B}{\varepsilon_0\varepsilon_B} + \frac{d_g}{\varepsilon_0})^{-1} \left[\int_0^d \frac{j(x,t)}{\varepsilon(x)} dx - \frac{\partial V(t)}{\partial t} \right], \tag{7}$$

Where d_B is medium thickness, d_g is gas gap thickness, ε_B is medium relative permittivity.

The charge density accumulated on the dielectric panel is given by:

$$\sigma(t) = \int_0^t j(x,t')dt' \tag{8}$$

The nonlinear equations were solved by Finite difference method which raised by Scharferter and Gummel .the charged particles in the model are electron e, Atomic ion He^+,Molecular ion He_2^+. the Metastability atoms $He(2^3s)$ and $He(2^1s)$ can be considered into a kind of effective types He^* because of the similar characteristics . The particles mobility and the diffusion coefficients are from [7], neglect the effect of discharge photoionization, Secondary electron emission coefficient is constant, Both of two ions are 0.01.

Simulation Results and Analysis

The discharge parameters in this paper are as follows:the dielectric constant $\varepsilon_1 = 7.5$, the thickness of the Medium layer $d_B = 0.1cm$, The spark gap $d_g = 0.2cm$, Gas pressure is a standard atmospheric pressure(P=101325 Pa）.

The Influence of the Voltage Slope to the Interval of the Current Pulse Maintaining Constant Voltage Amplitude

When the voltage function slope kept getting bigger, achieve the same voltage amplitude of the time needed for shorter, therefore the number of current pulse. Decreasing when constant voltage amplitude and large slope. Figure 1 and Figure 2 shows the discharge current with the time evolution. It can be seen from the two chart, when applied voltage amplitude remain unchanged,

slope continuously variable, in the half cycle of the applied voltage, the number of the current pulse gradually decreased. When the slope is 0.96, there are three current pulses, When the slope increases to 2.8, appeared a current pulse. We have studied many sets of data, with the increase of applied voltage gradient, although the applied voltage amplitude keep constant, but the intervals of appearing corresponding pulse are not identical in the half cycle of the applied voltage. With the increasing of the applied voltage slope, The number of the current pulse decreased continuously, the time interval between the corresponding current pulse presents generally larger at first, then decrease, as shown in figure 3 and figure 4.

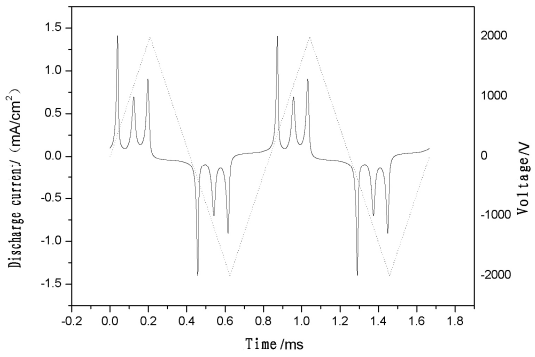


Fig. 1. Amplitude V0=2000V, slope K=0.96

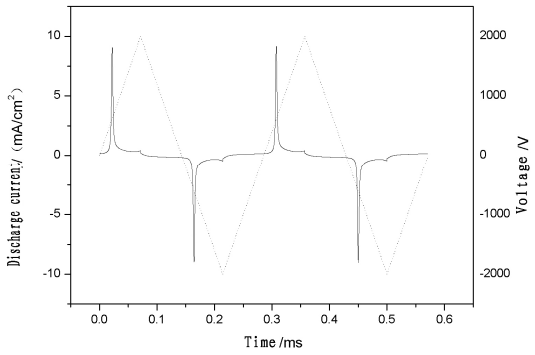


Fig. 2. Amplitude V0=2000V, slope K=2.8

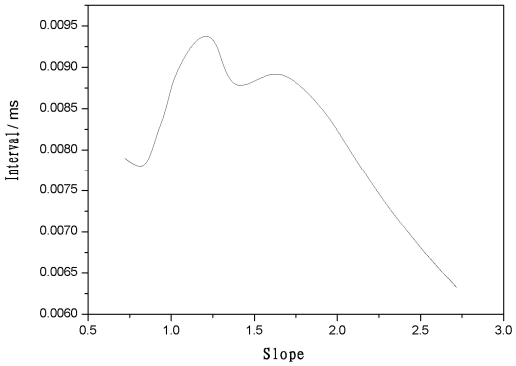


Fig. 3. The interval of the previous two pulse in the first half cycle

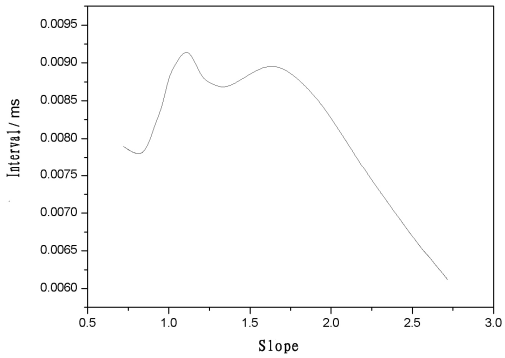


Fig. 4. The interval of the previous two pulse in the first half cycle

The Influence of the Voltage Slope to the Effective Utilization of Power Device Maintaining Constant Voltage Amplitude.

Figure 5 and figure 6 give the comparison of the power while the voltage amplitude unchanged, maintaining V0 = 4000, the slope becomes larger and larger .the ratio of the average power of the gas and the total power shows General trend, the first decline after easing, But the highest ratio is in the lower slope started,so to achieve high efficiency, should try to choose low voltage slope, the ratio of the maximum power of the gas and the total power shows General trend, first increased, then decreased, and then easing. the highest ratio is in the slope K=4.8, Therefore, we can choose the appropriate voltage according to need to ensure the maximum utilization of the discharge.

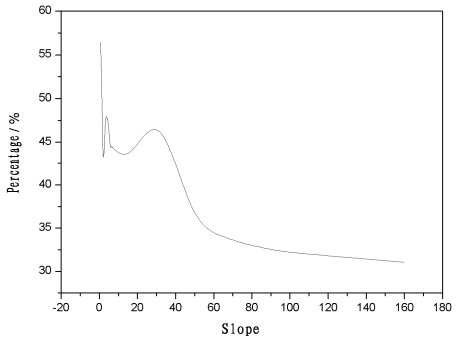


Fig. 5. The ratio of the average power

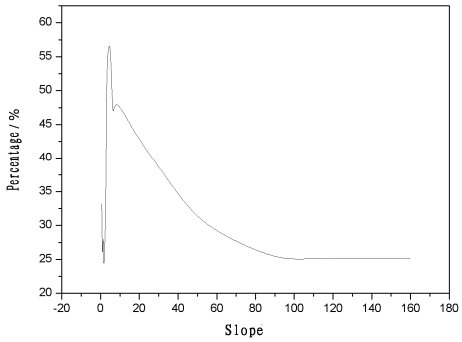


Fig. 6. The ratio of the maximum power

Conclusions

In this paper ,the homogeneous barrier discharge with multiple current pulses at atmospheric pressure in helium is numerically studied on the basis of one-dimensional fluid model . Get the following conclusions:

(1) the number of current pulses Decreased in the case of invariable voltage amplitude and the Voltage gradient kept getting bigger, the time interval corresponding between the current pulse is not the same. Generally First increases, and finally the trend is easing.

(2) the ratio of the average power of the gas and the total power shows General trend, the first decline after easing,.the ratio of the maximum power of the gas and the total power shows General trend, first increased, then decreased, and then easing. the highest ratio is in the slope K=4.8.

References

[1] F.Z. zhuang,Y.X. cai: Experimental Study on the Average Discharge Current in Dielectric Barrier Discharge Under Atmospheric Pressure. *High Voltage Engineering.* Vol. 34 (10), p. 2140-2144 (2008)

[2] Y.H. wang,D.Z. wang: Study on Homogeneous Multiple-pulse Barrier Discharge at Atmospheric Pressure. *Acta Physica Sinica.* Vol. 54 (3), p. 1295-1300 (2005)

[3] Y.H. wang,D.Z. wang: Numerical Simulation of Dielectric-barrier-controlled Glow Discharge at Atmospheric Pressure. *Acta Physica Sinica.* Vol. 52 (7), p. 1694-1699 (2003)

[4] X.X. wang: Dielectric Barrier Discharge and Its Applications. *High Voltage Engineering*. Vol. 35 (1), p. 1-11 (2009)

[5] Mangolini L: Radial structure of a low-frequency atmospheric-pressure glow discharge in helium. *Appl.Phys.Lett.* Vol. 80 (10), p. 1722-1726 (1722)

[6] Akishev Y S: Pulsed regime of the diffusive mode of a barrier discharge in helium. *Plasma Phys.Rep*. Vol. 27 (2), p. 164-168 (2001)

[7] Scharferter D L,Gummel H K: Analysis of a Silicon Read Diode Oscillator. *IEEE Trans. Electron. Devices*. Vol. 16(64), p. 64-67 (1969)

★ Project supported by Science Foundation of Ludong University(No.LY20082704, No.LZ20082704) and Promotive research fund for excellent young and middle-aged scientisits of Shandong Province(No. BS2010NJ003)

©(2014) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.209

A Study on the Acceptance and Use of Handheld E-Book Readers

MA Jian

School of Management, Zhejiang University of Media and Communications, Hangzhou, P. R. China

janmar@126.com

Keywords: e-book readers; self-efficacy; product feature; usefulness; ease of use

Abstract. This paper proposes a new model to test the technology acceptance model which proposed by Davis in order to study the influence factors on acceptance and use of handheld e-book readers. The main conclusions are as follows: 1.There are significant correlation between self-efficacy and usefulness, without significant correlation between self-efficacy and other variables; 2.There are significant correlation between product feature and intention to use, without significant correlation between product feature and other variables; 3.There are significant correlation between usefulness and ease of use as well as intention to use, without significant correlation between usefulness and other variables; 4.There are significant correlation between ease of use and usefulness, without significant correlation between ease of use and other variables; 5. There are no significant correlation among intention to use, self-efficacy, product feature, usefulness and ease of use.

Introduction

As early as 1998, the first generation e-book reader has appeared on the market, but the market prospect of e-book reader loomed in China until 10 years later. As far as the new type of e-book reader is concerned, what factors are the important affecting factors for acceptance and use of handheld e-book readers, and what's the extent of influence, are of great practical significance obviously.

In theory, a intensive study of influence factors on acceptance and use of handheld e-book readers can develop the applied scope of technology acceptance model, also can deepen understanding of problems in the use of handheld e-book readers, and have the possibility of completing original research results through prove and emend the technology acceptance model.

In practice, handheld e-book readers, as a reading way by digital media, actually represent the current reading trend of new digital media. It is highly probable that the magnetism of handheld e-book readers to the reading public is augmenting daily with the increase of e-book resources. And the exploratory research on this subject is of great practical significance obviously.

Literature Review

The theory of reasoned action, which Fishbein and Ajzen put forward in 1975, stated that the actual behavior of one person is determined by his behavior intention, and behavior intention is influenced by the attitude towards behavior and subjective norm simultaneously. This theory acknowledges the existence of external variable, and states that the external variable can affect the cognitive attitudes of people, and then affect the actual behavior of people. In other words, the external variable may affect the belief of people, and affect the attitude towards behavior, and then affect the behavior intention, and finally affect the actual behavior. This theory is widely used to predict and explain the relationship between attitude and behavior by social psychologists[1]. Although numerous individual, organizational, and technological variables have been investigated [2], research has been constrained by the shortage of high-quality measures. Past research indicates that many measures do not correlate highly with system use [3].

In 1985, Davis proposed a technology acceptance model(see Fig. 1) which based on the theory of reasoned action proposed by Fishbein and Ajzen to examine the relationship among cognition, affective factors, and the use of technology, and this model intend to explain the reasons why users accept or reject the new information system. The innovation of technology acceptance model lies in

that this model proposes two new concepts: perceived usefulness and perceived ease of use. Perceived usefulness is defined as "the degree to which a person believes that using a particular system would enhance his or her job performance." Perceived ease of use, in contrast, refers to "the degree to which a person believes that using a particular system would be free of effort."[4]

We propose a new model to test on this condition as follows (see Fig. 2).

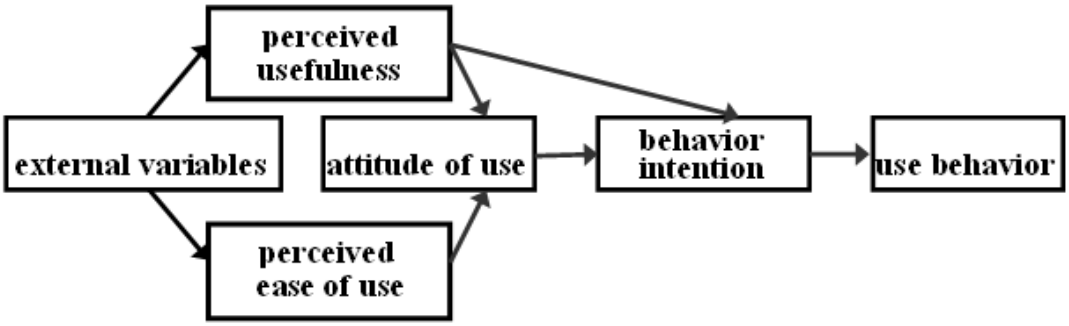


Fig. 1 Technology Acceptance Model

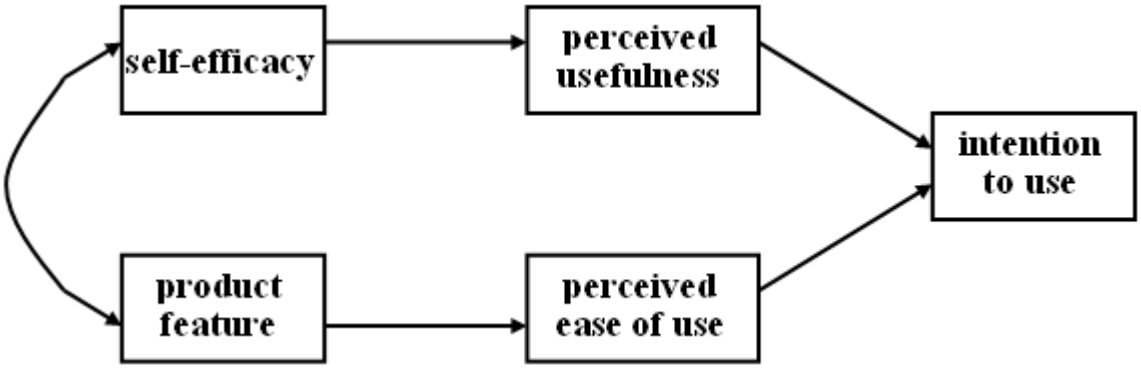


Fig. 2 A Model to Test

Data Analysis and Main Conclusions

We invited 160 students in Shanghai Jiaotong University to use EDO E600 handheld e-book reader on trial, and directly interfaced with them. And then we tackled and analyzed the questionnaire by SPSS.

Differences Test on Gender of Variables.

Table 1 Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Self-Efficacy	Male	82	3.9177	0.64437	0.07116
	Female	78	3.8397	0.71733	0.08122
Product Feature	Male	82	3.9360	0.58550	0.06466
	Female	78	3.8301	0.41011	0.04644
Usefulness	Male	82	3.2061	0.70507	0.07786
	Female	78	3.0269	0.71144	0.08055
Ease of Use	Male	82	3.4768	0.47513	0.05247
	Female	77	3.4039	0.51362	0.05853
Intention to Use	Male	82	3.4228	0.76989	0.08502
	Female	78	3.2436	0.88220	0.09989

The results of t-test show that there are no significant differences in variables, so the effection of gender on every variables can be omitted.

Differences Test on Educational Level of Variables.

Table2 Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Self-Efficacy	Bachelor	61	3.7623	0.72948	0.09340	3.5755	3.9491	1.00	5.00
	Master	74	3.9088	0.65110	0.07569	3.7579	4.0596	1.00	5.00
	Ph.D.	25	4.0800	0.60260	0.12052	3.8313	4.3287	2.50	5.00
	Total	160	3.8797	0.67988	0.05375	3.7735	3.9858	1.00	5.00
Product Feature	Bachelor	61	3.8197	0.42224	0.05406	3.7115	3.9278	2.67	4.83
	Master	74	3.8457	0.41830	0.04863	3.7488	3.9426	2.92	5.00
	Ph.D.	25	4.1567	0.80303	0.16061	3.8252	4.4881	3.33	7.50
	Total	160	3.8844	0.50883	0.04023	3.8049	3.9638	2.67	7.50
Usefulness	Bachelor	61	3.1311	0.66572	0.08524	2.9606	3.3016	1.40	4.90
	Master	74	3.1608	0.62457	0.07260	3.0161	3.3055	1.00	4.50
	Ph.D.	25	2.9640	1.01484	0.20297	2.5451	3.3829	1.10	4.30
	Total	160	3.1188	0.71165	0.05626	3.0076	3.2299	1.00	4.90
Ease of Use	Bachelor	61	3.5115	0.41637	0.05331	3.4048	3.6181	2.40	4.50
	Master	74	3.4703	0.52231	0.06072	3.3493	3.5913	2.30	4.60
	Ph.D.	24	3.1750	0.51773	0.10568	2.9564	3.3936	2.30	4.30
	Total	159	3.4415	0.49393	0.03917	3.3641	3.5189	2.30	4.60
Intention to Use	Bachelor	61	3.2295	0.80632	0.10324	3.0230	3.4360	1.00	5.00
	Master	74	3.4640	0.76215	0.08860	3.2874	3.6405	1.00	5.00
	Ph.D.	25	3.2133	1.03136	0.20627	2.7876	3.6391	1.00	5.00
	Total	160	3.3354	0.82881	0.06552	3.2060	3.4648	1.00	5.00

Table3 ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Self-Efficacy	Between Groups	1.906	2	0.953	2.090	0.127
	Within Groups	71.590	157	0.456		
	Total	73.496	159			
Product Feature	Between Groups	2.220	2	1.110	4.474	0.013
	Within Groups	38.947	157	0.248		
	Total	41.166	159			
Usefulness	Between Groups	0.739	2	0.369	0.727	0.485
	Within Groups	79.785	157	0.508		
	Total	80.524	159			
Ease of Use	Between Groups	2.064	2	1.032	4.414	0.014
	Within Groups	36.482	156	0.234		
	Total	38.546	158			
Intention to Use	Between Groups	2.280	2	1.140	1.673	0.191
	Within Groups	106.942	157	0.681		
	Total	109.222	159			

The results of variance analysis show that there are differences in educational level of product feature and ease of use, and what shows that educational level made significant effects on these two variables.

Differences Test on Subject of Variables.

Table4 ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Self-Efficacy	Between Groups	0.007	1	0.007	0.016	0.899
	Within Groups	73.489	158	0.465		
	Total	73.496	159			
Product Feature	Between Groups	0.032	1	0.032	0.123	0.726
	Within Groups	41.135	158	0.260		
	Total	41.166	159			
Usefulness	Between Groups	0.372	1	0.372	0.733	0.393
	Within Groups	80.152	158	0.507		
	Total	80.524	159			
Ease of Use	Between Groups	0.674	1	0.674	2.795	0.097
	Within Groups	37.872	157	0.241		
	Total	38.546	158			
Intention to Use	Between Groups	0.251	1	0.251	0.364	0.547
	Within Groups	108.971	158	0.690		
	Total	109.222	159			

There are no differences in variables except in different subjects of self-efficacy.

Differences Test on Daily Consumption of Variables.

Table5 ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Self-Efficacy	Between Groups	0.974	4	0.243	0.520	0.721
	Within Groups	72.523	155	0.468		
	Total	73.496	159			
Product Feature	Between Groups	0.133	4	0.033	0.125	0.973
	Within Groups	41.034	155	0.265		
	Total	41.166	159			
Usefulness	Between Groups	1.625	4	0.406	0.798	0.528
	Within Groups	78.899	155	0.509		
	Total	80.524	159			
Ease of Use	Between Groups	2.085	4	0.521	2.201	0.071
	Within Groups	36.461	154	0.237		
	Total	38.546	158			
Intention to Use	Between Groups	1.576	4	0.394	0.567	0.687
	Within Groups	107.645	155	0.694		
	Total	109.222	159			

There are no significant differences in five variables.

Relative Analysis among Five Variables.

Table6 Descriptive Statistics

	Mean	Std. Deviation	N
Product Feature	3.8635	0.42042	160
Self-Efficacy	3.8797	0.67988	160
Usefulness	3.1188	0.71165	160
Ease of Use	3.4415	0.49393	159
Intention to Use	3.3354	0.82881	160

Table7 Correlations

		Product Feature	Self-Efficacy	Usefulness	Ease of Use	Intention to Use
Product Feature	Pearson Correlation	1				
	Sig. (2-tailed)					
	N	160				
Self-Efficacy	Pearson Correlation	0.096	1			
	Sig. (2-tailed)	0.226				
	N	160	160			
Usefulness	Pearson Correlation	0.120	0.153*	1		
	Sig. (2-tailed)	0.131	0.053			
	N	160	160	160		
Ease of Use	Pearson Correlation	-0.037	0.113	0.254***	1	
	Sig. (2-tailed)	0.648	0.156	0.001		
	N	159	159	159	159	
Intention to Use	Pearson Correlation	0.171*	0.186**	0.696***	0.365***	1
	Sig. (2-tailed)	0.030	0.018	0.000	0.000	
	N	160	160	160	159	160

*. Correlation is significant at the 0.06 level (2-tailed).

**. Correlation is significant at the 0.05 level (2-tailed).

***. Correlation is significant at the 0.01 level (2-tailed).

The conclusions have been drawn from the above data analysis: 1.There are significant correlation between self-efficacy and usefulness, without significant correlation between self-efficacy and other variables; 2.There are significant correlation between product feature and intention to use, without significant correlation between product feature and other variables; 3.There are significant correlation between usefulness and ease of use as well as intention to use, without significant correlation between usefulness and other variables; 4.There are significant correlation between ease of use and usefulness, without significant correlation between ease of use and other variables; 5. There are no significant correlation among intention to use, self-efficacy, product feature, usefulness and ease of use.

References

[1] M. Fishbein, & I. Ajzen, Belief: Attitude, Intention, and Behavior: An Introduction to Theory and Research. Reading, Addison-Wesley, MA(1975).

[2] C. Franz, & D. Robey: Organizational Context, User Involvement, and the Usefulness of Information Systems. Decision Sciences, Summer(1986), p.329-356.

[3] C. Schewe: The Management Information System User: An Exploratory Behavioral Analysis. Academy of Management Journal, December(1976), p.577-590.

[4] F. Davis: Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, September(1989), p.319-340.

doi:10.4028/www.scientific.net/AMM.50-51.214

Fuzzy Vibration Control of Flexible Solar Panel Using Piezoelectric Stack

Rui Xu [a], Dongxu Li [b] and Jianping Jiang [c]

College of Aerospace and Material Engineering, National University of Defense Technology, Changsha, 410073, P.R.China

[a]alpc_xurui@139.com, [b]dongxuli@263.net, [c]jianping202@163.com

Keywords: Vibration control; Fuzzy control; Piezoelectric smart structure; PZT stack; Solar panel

Abstract. To effectively suppress vibrations of the flexible solar panel, the fuzzy logic control with piezoelectric smart structure is studied. The bending moment induced in the solar panel by the PZT stack actuator is formulated. The dynamical equations of the solar panel are derived. A fuzzy logic controller which uses universal fuzzy sets is designed. Considered the characteristic of the PZT stack, only positive control voltages were loaded to it. The finite element method simulation results demonstrate that the fuzzy logic controller can suppress the vibrations of the flexible spacecraft solar panel effectively.

Introduction

In order to meet limited launch weight and costs, and more power requirement, space craft solar panels are becoming lager and more flexible, and they usually have very low damping ratios, high dimensional order, low modal frequencies and parameter uncertainties in dynamics[1, 2]. Any disturbance, such as attitude or orbit maneuver, orbital debris collision, will result in low frequency vibration of solar panel, the unwanted vibrations of such flexible appendages will be caused unavoidably. Hence, the issue on modeling and active vibration control for such cantilever plate system is rather difficult and challenging[3, 4].

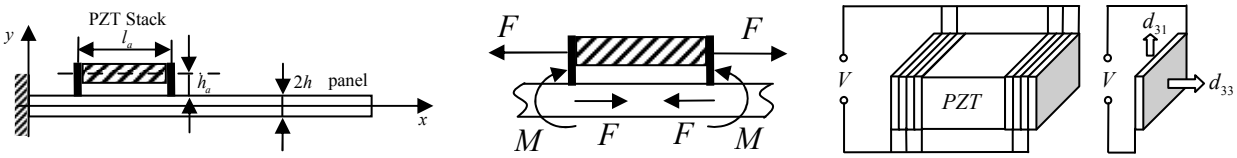
During the past few decades, the active vibration control has widely been studied. Their works focus on the development of actuators and control methodologies. Smart materials, such as piezoelectric lead zirconate titanate (PZT) patches and polyvinylidene fluoride (PVDF), have been used extensively as distributed sensors and actuators[1, 5-7], but they can not suppress the vibrations rapidly because of their limited force outputs. In this paper, a PZT stack actuators is utilized. Unlike conventional d_{31} patch actuators, the d_{33} geometry and stacked configuration of the actuators provide large control force in response to applied electric fields[8].

Control methodologies are also very important for vibration control. Since the flexible solar panel is a distributed parameter system of infinite order, model inaccuracies or uncertainties will have strong adverse effect on control system, when it is approximated by a lower-order model and controlled by a finite-order controller. In the absence of sufficiently precise process mathematic model and in the presence of non-linearity, fuzzy logic based control usually have an advantage over classical control and modern control[9]. In recent years, it has been widely used for vibration control[10-13]. In this paper, a fuzzy logic controller which use universal fuzzy sets is designed, it will receives the solar panel tip displacement and the velocity as inputs, while gives the control voltage as output.

Modeling and Design of Fuzzy Logic Controller

Modeling. A typical single-panel solar panel structure made of composite material is considered, it can be simplified as cantilever beam when we only care about its bending vibration, as shown in Fig. 1(a). The PZT stack actuator is fixed on the solar panel by rigid bracket. It is consists of many thin layers of PZT ceramic laminated together and electrically connected in parallel, as shown in Fig. 1(c).

The PZT stack is displacement generating device. It expands proportionally to the applied electric field or actuation voltage. Maximum motion or expansion of the PZT occurs at the maximum actuation voltage. When an external load resists the motion of the PZT expansion, the PZT stack applies a force that is a function of the stiffness of the external load. The stack elements can withstand high pressure and show the highest stiffness of all PZT actuator designs, but they *cannot* withstand large pulling forces.



(a) (b) (c)

Fig. 1 Smart solar panel model

Piezoelectric stack actuator model:

When an electric voltage V loaded in the z-direction of polarization between the two surfaces of the PZT-crystal, it will deform them oppositely by a strain

$$S_3 = d_{33}E_3 = d_{33}V/t \tag{1}$$

where $E_3 = V/t$ is the electric field intensity in the patch, t is the thickness of piezoelectric patch, d_{33} is the direction piezoelectric constant of the PZT.

So the voltage induced deformations of the PZT stack actuator when unconstrained is

$$\delta_E = S_3 n_a t = V n_a d_{33} \tag{2}$$

The condition of coincident deformation is

$$\delta_E + \delta_F = \delta_F' \tag{3}$$

where δ_F is the elastic deformation of the PZT stack, δ_F' is the elastic deformation of the beam.

$$\begin{cases} \delta_F = -Fl_a / E_a A_a \\ \delta_F' = Fl_a / E_b A_b \end{cases} \tag{4}$$

where E_a, E_b are the longitudinal elastic module of the PZT stack and solar panel, respectively (both materials remaining normally elastic), A_a, A_b are the cross section area of the PZT stack and solar panel, l_a is the length of the PZT stack.

From equation (3), the modal force result from a combination of the elastic and the voltage induced deformations of the PZT stack actuator assemblies is

$$F = \frac{n_a d_{33} E_a A_a E_b A_b}{la\left(E_a A_a + E_b A_b\right)} V \tag{5}$$

The equivalent moment produce by PZT stack can be express as:

$$M = F\left(h_a + h\right) \tag{6}$$

Finite element model:

The finite element method is a powerful numerical technique that provides solutions to many complicated engineering problems. In this paper, the finite element formulation of the smart solar panel structure is derived by applying Hamilton's principle. The equations of equilibrium governing the linear dynamic response of a system of finite elements are shown in equation (7).

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{C}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{F}_{\mathbf{f}} - \mathbf{F}_{\mathbf{control}} \tag{7}$$

where $\mathbf{M}$, $\mathbf{C}$, $\mathbf{K}$ are the mass, damping and stiffness matrix respectively; $\mathbf{q}$, $\dot{\mathbf{q}}$, $\ddot{\mathbf{q}}$ are the displacement, velocity and acceleration vectors of the finite element assemblage; and $\mathbf{F}_{\mathbf{f}}$, $\mathbf{F}_{\mathbf{control}}$ are the disturb force and control force matrix.

Damping matrix is calculated by Rayleigh damping model as followed equation:

$$\mathbf{C} = a\mathbf{M} + b\mathbf{K} \tag{8}$$

Design of Fuzzy Logic Controller. In order to reduce the displacement field of the cantilever beam system, a nonlinear fuzzy controller was constructed by using the Fuzzy Toolbox of Matlab. More specifically, a Mamdani-type Fuzzy Inference System, consisted of two inputs and one output, was developed. The system receives the solar panel tip displacement (u) and the velocity ($\dot{u}$) as inputs, while gives the control voltage (V) as output. Seven fuzzy sets (NL, NM, NS, ZO, PS, PM, PL), two input power sets (U, DU), and one output power set (V) are used. The labels NL, NM, NS, ZO, PS, PM, PL refer to the linguistic values: negative large, negative medium, negative small, zero, and positive small, positive medium and positive large, respectively. The triangular membership function (Fig. 2) is applied to the fuzzy sets.

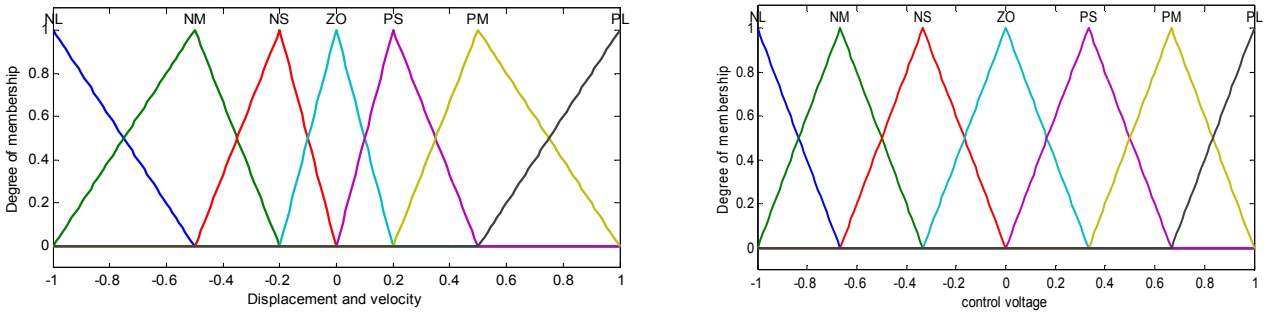


Fig. 2 Membership function plots

The rule base used in this paper is shown in Table 1. The first row and the first column of this table are the inputs from the system and the others are the output of the fuzzy logic controller. The knowledge about the vibrating beam is employed in constructing this rule base. For example, if the deflection of the beam is likely to be large in the positive y-direction, and the beam has small velocity of deflection in the negative y-direction, then the output will be a small force in the negative y-direction. The control surface is shown in Fig. 3.

Table 1 Fuzzy IF-THEN rule base

V / $\dot{u}$ / u	NL	NM	NS	ZO	PS	PM	PL
NL	NL	NL	NL	NM	ZO	ZO	ZO
NM	NL	NL	NL	NS	ZO	ZO	PS
NS	NL	NL	NM	ZO	ZO	PS	PM
ZO	NL	NM	NS	ZO	PS	PM	PL
PS	NM	NS	ZO	ZO	PM	PL	PL
PM	NS	ZO	ZO	PS	PL	PL	PL
PL	ZO	ZO	ZO	PM	PL	PL	PL

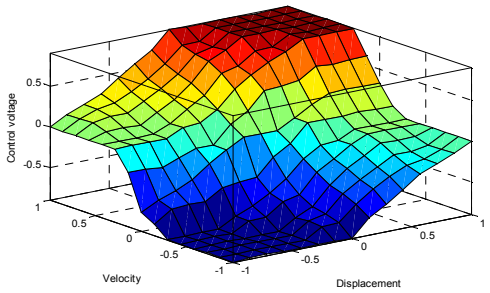


Fig. 3 Control surface

The scaling factors are important especially if the fuzzy sets are not customized for each case. In this paper, the universal fuzzy sets are used; therefore, the scaling factors are very important. By changing the scaling factors, the fuzzy sets can be customized for each system. To customize the fuzzy sets for the system, the input and out put power sets and the available output must be considered.

Since the PZT stack cannot withstand large pulling forces, negative voltage must not loaded to it. Hence, the fuzzy logic controller that designed above cannot be used directly. Its negative control signal should be inhibited. That means, *if the output of the fuzzy logic controller is negative, then let the control voltage be zero*.

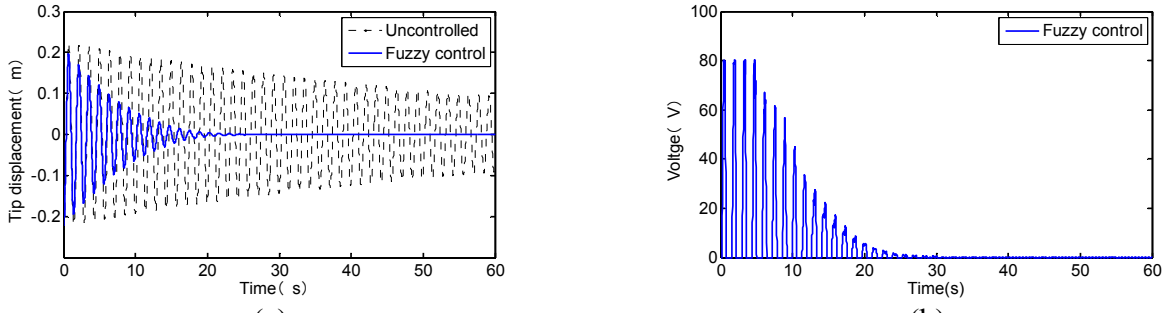
Simulation

Simulation Parameters. The solar panel structure is modeled by 60 beam elements. The patches number of the PZT stack is 80. The other parameters are shown in Table 2.

Table 2 The parameters of piezoelectric intelligent solar panel

Description	Notation	Value	Unit
Length\width\thicness of the solar panel	$l_s \backslash b_s \backslash t_s$	1200\50\4	$[mm]$
Length\width\thicness of the PZT stack	$l_a \backslash b_a \backslash t_a$	50\30\10	$[mm]$
Density of the solar panel	$d_b \backslash d_a$	2780	$\left[kg/m^3\right]$
Elastic module of the solar panel\PZT stack	$E_b \backslash E_a$	$3\times10^{10}\backslash7.6\times10^{10}$	$[Pa]$
Piezoelectric constant\maximue work voltage	$d_{33} \backslash U_{\max}$	$5\times10^{-10}\backslash200$	$[C/N]\backslash[V]$

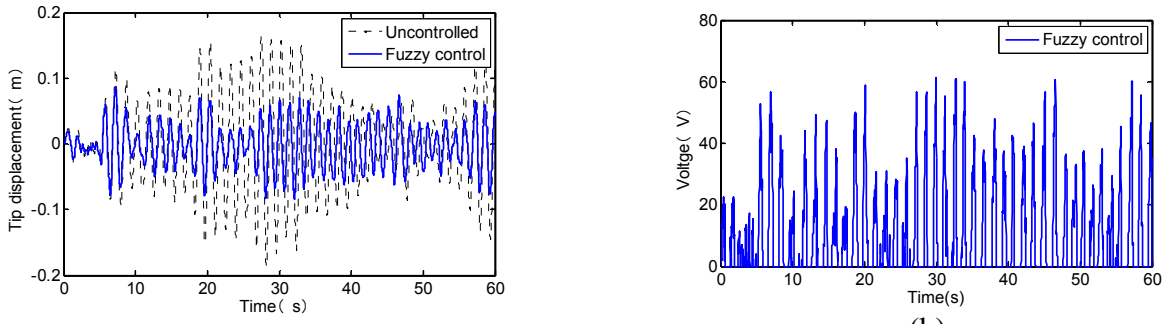
CASE 1: Initial Displacement Field. Considering an initial displacement field applied to the solar panel structure which is obtained by a mechanical force applied at the free tip that induces a tip displacement equal to 0.22m. The time-domain closed-loop response of the solar panel structure with the fuzzy logic controller is compared with the open loop response in Fig. 4(a). The decay time of the bending vibration with fuzzy logic control effect is about 20s, and the low amplitude vibration will last for a period of time. When compared with the open-loop decay time (more than 200s), it reveals a great improvement on the response attenuation. The control voltage is presented in Fig. 4(b). Note that the peak values of control voltage do not exceed the allowable range.



(a) (b)

Fig. 4 Tip displacements and control voltages for initial displacement disturbance

CASE 2: White Noise Excitation. Assume that a white noise excitation is applied on the tip of the solar panel. The time-domain closed-loop response of the solar panel structure with the fuzzy logic controller is compared with the open loop response in Fig. 4(a). It can be seen that fuzzy control system can suppress the vibration very efficiently. The control voltage is presented in Fig. 5(b). Note that the peak values of control voltage do not exceed the allowable range.



(a) (b)

Fig. 5 Tip displacements and control voltages for white noise excitation

Summary

An analytical review of the flexural actuation of a flexible solar panel has been presented in order to formulate the bending moment induced in the solar panel by the PZT stack actuator. A fuzzy logic controller is designed to control the vibrations of the smart solar panel. FEM-simulation researches show that the fuzzy logic controller can effectively suppress the vibrations of the smart solar panel in a short time with appropriate choice of input and out put scaling factors.

References

[1] Z.C. Qiu, H.X. Wu, C.D. Ye: Aerospace Science and Technology Vol.13 (2009) p.277.
[2] G.B. Maganti, S.N. Singh: Acta Astronautica Vol.61 (2007) p.575.
[3] J. Zheng, S.P. Banks, H. Alleyne: Acta Astronautica Vol.56 (2005) p.519.
[4] X.S. Ge, Y.Z. Liu: Chaos, Solitons & Fractals Vol.37 (2008) p.108.
[5] W. Liang, B. Rui-xiang, Y. Cheng: Composite Structures Vol.92 (2010) p.1410.
[6] Y.Y. Li, L. Cheng, P. Li: Composite Structures Vol.62 (2003) p.155.
[7] D. Sun, J.K. Mills, J. Shan, S.K. Tso: Mechatronics Vol.14 (2004) p.381.
[8] J. Redmond, G. Parker, P. Barney, R. Rodeman: ASME International Congress and Exposition (1996) p.1.
[9] V. KAMINSKAS, R. LIUTKEVICIUS: Institute of Mathematics and Informatics, Vilnius Vol.13 (2002) p.287.
[10] M. Marinaki, Y. Marinakis, G.E. Stavroulakis: Control Engineering Practice Vol.18 (2010) p.618.
[11] Q. Wenzhong, S. Jincai, Q. Yang: Journal of Sound and Vibration Vol.275 (2004) p.917.
[12] J. Lin, W.Z. Liu: Journal of Sound and Vibration Vol.296 (2006) p.567.
[13] D.Q. Sun: Yuhang Xuebao/Journal of Astronautics Vol.30 (2009) p.1537.

doi:10.4028/www.scientific.net/AMM.50-51.219

A Biomimetic Squid Funnel Actuated By Shape Memory Alloy Wires

Zhenlong Wang[1,a], Fei Gao[1, 2, b] , Yukui Wang[1, c] , Cheng Guo[1]

[1] School of Mechanical Engineering, Harbin Institute of Technology, Harbin, China

[2] State Key Laboratory of Robotics and System at Harbin Institute of technology, Harbin, China

[a]wangzl@hit.edu.cn , [b]mojiegaofei@163.com, [c]wangyukui@hit.edu.cn

Keywords: Biomimetic squid funnel, Shape Memory Alloy, Pulse Width Modulation

Abstract. Inspired by the squid's funnel which can bend flexibly, a kind of biomimetic squid funnel actuated by shape memory alloy (SMA) wires was developed for controlling the jet direction in this article. With the application of Pulse Width Modulation (PWM) technique, experimental investigation was conducted. And the results show that, the bending angle of the biomimetic squid funnel can be controlled by changing PWM duty cycle and the final stable resistance of SMA wire decreases while the duty cycle is increased. The relationship between the banding angle of the biomimetic squid funnel and the resistance of SMA wire is approximately inverse linear. The feasibility of jet direction control with the biomimetic squid funnel was verified by a simulated jetting experiment.

Introduction

In the ocean, the squid and jellyfish swim forward by jetting, which is unusual. Compared with the fin-propulsion or tai-propulsion that most fish swim through, this jet-propulsion not only can achieve a high speed (such as the escape jet behavior of the squid), but also can realize precise positioning in the slow swimming (such as the floating action of the jellyfish), and its propulsion efficiency is also high [1], therefore some researchers have tried to develop the biomimetic robots with jet-propulsion. Sung-Weon Yeom et al [2] designed and manufactured a biomimetic jellyfish robot based on ionic polymer metal composite actuators to mimic real locomotive behavior with pulse and recovery processes. Guo, shuxiang et al [3] have developed a new jellyfish type of micro robot which can imitate the jellyfish. This kind of micro robot has a movement process as the jellyfish does when it is floating and sinking. They have successfully imitated the action of jetting, but the control of jet direction has not been researched.

Jet-propulsion has been successfully used in the ships, which can precisely control the direction of jet by different forms of vector jet devices, such as box-type jet device and Spherical jet device [4]. However, those devices are all achieved by mechanical transmission structure, so it's difficult to apply in the small biomimetic underwater robots due to their complex structure. Though the structure of the squid funnel is so simple, it can control the jet direction by bending flexibly. So a kind of biomimetic squid funnel can be designed to fit some small biomimetic jet-propulsion robots.

In recent years, the smart materials [5] are widely used in different kinds of flexible bending structure, and shape memory alloys (SMAs) are one of them. Compared with other smart materials, SMAs have many advantages, such as large recoverable deformation, large output force [6], easy to actuate and self-sensing. So SMA wires have been chosen as the actuator for the biomimetic squid funnel. SMAs are a unique class of metal alloys that can recover apparent permanent strains when they are heated above a certain temperature [7]. It can be actuated easily by electrical current, and Pulse Width Modulation (PWM) technique is widely used for the control of SMA [8].

Inspired by the squid's funnel, a kind of biomimetic squid funnel based on SMA wires and its control system were developed in this article. The biomimetic squid funnel, with the characteristics of simple structure, small size, noiseless when acting and bending flexible, was easily integrated into the small biomimetic jet-propulsion robots and it offered a new solution for the control of jet direction of the jet-propulsion robots.

Structure of the biomimetic squid funnel

As shown in Fig.1, the cross section of the squid's funnel is approximate round, and the diameter of the circle is about 10 mm, which is a reference for design. The squid's funnel consists of the muscular hydrostat, and the muscular hydrostat consists mainly of muscles with no skeletal support [9]. This biological structure is capable of substantial movement and deformation, so the squid funnel can bend flexibly.

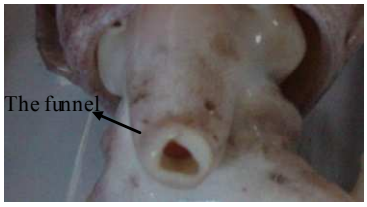


Fig.1 The shape of the squid's funnel

As shown in Fig.2, the biomimetic squid funnel consists of the corrugated pipe, the silica gel rod, the plexiglass sheet, SMA Wires, pipe electrode and the connector. The corrugated pipe is easy to bend compared with the common tube. There are three fan-shaped hollow areas and three pairs of small holes on the plexiglass sheet. The hollow areas are the water channel. The three pairs of holes are distributed in the same circle with interval of 120 degrees. The folded SMA wires were passed through the small holes on the plexiglass sheet in turn, and the ends of them were fixed with pipe electrode. The silica gel rod is made by mold, and the straightened SMA wires were imbedded in it.

When one of the three SMA wires was heated by electrical current and was contracted, the silica gel rod could bend toward the side of the contracted wire. At the same time, the curved silica gel rod and the other two stretched SMA wires were storing a certain amount of elastic energy. When the SMA wire was cooling down by cutting off power supply, the storage of elastic energy was released and the curved silica gel rod was recovered. When two of the three SMA wires were actuated simultaneously, the biomimetic squid funnel could bend toward a certain direction between them.

The photograph of the biomimetic squid funnel is shown in Fig.3. The length of the biomimetic funnel was 40 mm, and the diameter of it was 16mm. The NiTi SMA was chosen as the actuator, and the diameter of the SMA wire is 0.2 mm. A biomimetic funnel mechanical model was established, and the theoretical maximum of bending angle that is 33° was calculated.

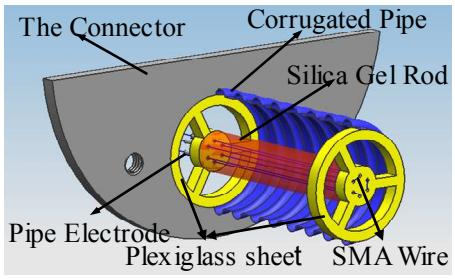


Fig.2 Structure of the biomimetic squid funnel

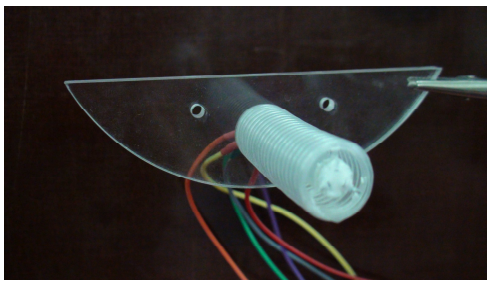


Fig.3 The biomimetic squid funnel

Control system of the biomimetic squid funnel

The block diagram of SMA actuators control system is shown in Fig.4, the LM3S8962 microcontroller was chosen as the control chip, and its PWM module can generate center-aligned PWM signals while the timer in PWM generator runs in Count-Up/Down mode. The advantage of this mode is that a load pulse in the center of pulse width can be used to generate an ADC trigger, which can simplify the program. The outputted PWM signals were amplified by the TPS2812P driver to drive the MOSFET and controlled the main circuit on and off.

The main circuit connects the drive power, MOSFET, SMA wire and a resistor R in series, which is shown in Fig.4. The voltage across the resistor R could be detected to calculate the resistance of SMA wire. The advantage of this method is to avoid the using of electric current sensor, which could

reduce circuit complexity and improve the device integrated level. The resistance of SMA wire can be calculated in following formula.

$$R_{SMA} = (V_{cc} - U) \times R / U \tag{1}$$

Where V_{cc} is the voltage of the drive power, U is the voltage across the resistor, R is the resistance of the resistor.

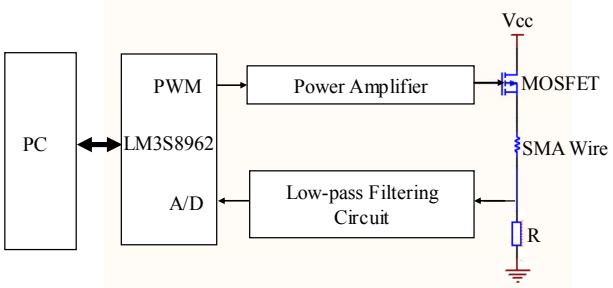


Fig.4 The block diagram of SMA actuators control system

The control system program consists of the PC program and the microcontroller program. The microcontroller program consists of the main program and the AD interrupt program. In the AD interrupt program, the voltage collected across the resistor is converted to the resistance of SMA wire according to Eq.1. All data types are converted to strings of ASCII characters then they are transmitted to the PC computer through a serial port.

Experiments and results

The PWM technique was applied to control the SMA actuators. When the PWM cycle was fixed, the bending angle of the biomimetic squid funnel and the change of the resistance of SMA wire were measured at different duty cycle. A simulated experiment of jetting was conducted at last.

The bending angle of the biomimetic squid funnel under different PWM duty cycle. The maximum bending angle that the biomimetic squid funnel could reach under water was measured under different PWM duty cycle. At beginning, the duty cycle is 10%, and then it is increased with intervals of 5%. The drive voltage is 8 V and the PWM cycle is 16 ms. The results showed that the biomimetic squid funnel could not bend when the duty cycle was less than 15%. After that, the bending angle increased linearly as the duty cycle increased, as shown in Fig.5.

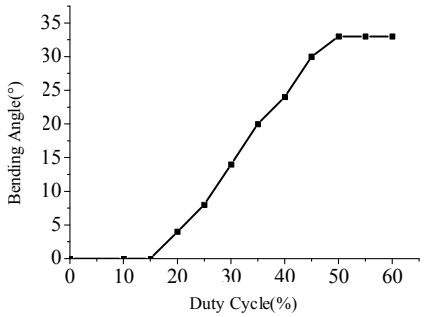


Fig.5 The bending angle under different duty cycle

The bending angle reached its maximum of 33° when the duty cycle was more than 50%, as shown in Fig.6.

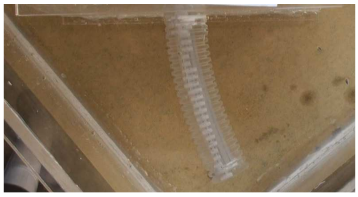


Fig.6 The maximum banding angle of the biomimetic squid funnel under water

Variation of the SMA wire resistance under the fixed duty cycle. A fixed duty cycle PWM signal was used to control the heating and cooling of SMA wire, and the voltage signal across the

resistor R was detected in the center of pulse width, then the SMA wire resistance signal was acquired by Eq.1. The experimental resistance curve which is detected under the condition of 40% duty cycle is shown in Fig.7. The curve1 is the original signal which exist some noise, and the curve2 is the signal after smoothing with the origin software.

Judging from the curve2, variation of the SMA wire resistance can be divided into three stages. The AB stage is the up stage. General speaking resistance of metals increases as the temperature increases, so the resistance of SMA wire increases slightly at the beginning. The BC stage is the down stage. As the temperature continued to raise, SMA wire started to transform from the martensitic phase to the austenitic phase, thus the resistance of SMA wire decreased as the austenite composition. The CD stage is the stable stage. When the heating and cooling of SMA wire reached dynamic equilibrium, the SMA wire temperature was kept in constant, and the ratio of austenite and martensite constituents remained unchanged, so did the SMA wire resistance.

The resistance of the SMA wire rose slightly in the AB stage, but since the reverse phase transformation (martensite-to-austenite) didn't start, the SMA wire generated no deformation; therefore the biomimetic squid funnel didn't bend at this stage. In the BC stage, the resistance of the SMA wire decreased linearly, and each value corresponds to a phase transition point. And different phase transition points correspond to different deformation of the SMA wire, therefore each resistance of SMA wire in the BC stage corresponds to a bending angle of the biomimetic squid funnel.

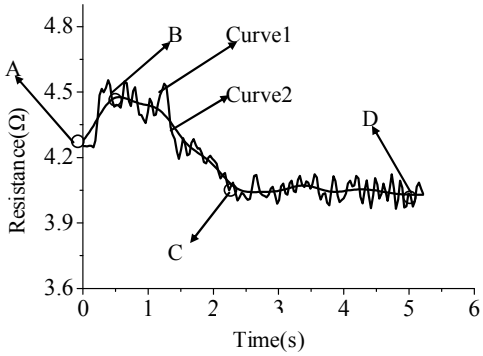


Fig.7 Variation of the SMA wire resistance under 40% duty cycle

Relationship between the resistance of SMA wire and the bending angle of the biomimetic squid funnel. Variations of the resistance of SMA wire under different duty cycle were measured by experiment. As shown in Fig.8, the curves were smoothed by origin software. As seen from the chart, the curves coincide together in the up stage and are distinguished by different slope in the down stage. The higher the duty cycle is, the faster the heating rate is and the faster the phase transition process is. So the slope of the curve becomes steeper as the duty cycle increases. In the stable stage, the resistances that reached stable under different duty cycle were different. The higher the duty cycle is, the smaller the resistance is.

The relationship between the banding angle of the biomimetic squid funnel and the resistance of SMA wire is approximately inverse linear, as shown in Fig.9. This provides a basis for the use of the electrical resistance of SMA wires as a feedback signal.

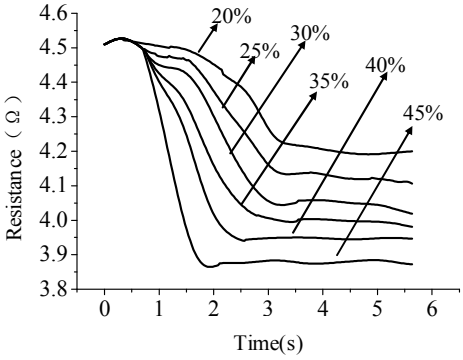


Fig.8 Variations of the resistance of SMA wire under different duty cycle

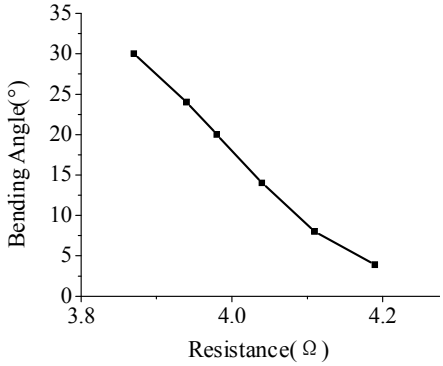


Fig.9 Relationship between the resistance of SMA wire and the bending angle of the biomimetic squid funnel

The simulated jetting experiment. As the Fig.10 shows, a square experimental device was designed for the simulated jetting experiment. The biomimetic squid funnel was fixed in the front of the device. The water flowed into the hole in the back of the device, and spurted out of the biomimetic squid funnel to simulate jetting. As shown in Fig.11, the biomimetic squid funnel verified its ability of control of the jet direction under the flow. Compared with static water, the processing of bending was slowed down and the maximum of bending angle was sharply decreased due to hydraulic impact. So the load capacity of the biomimetic squid funnel should be enhanced, like using SMA wires whose diameter is bigger.

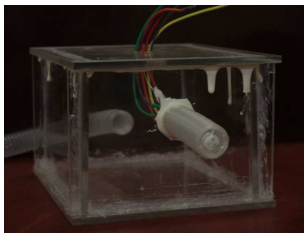


Fig.10 Experimental device

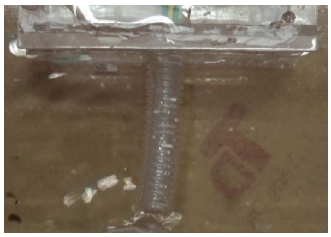


Fig.11 Control of jet direction under the flow

Conclusions

In conclusion, we proposed a kind of biomimetic squid funnel actuated by SMA wires that could bend flexibly to imitate the squid's funnel for controlling the jet direction. The PWM technique was applied to control SMA actuators. The bending angle of the biomimetic squid funnel can be controlled by changing duty cycle of PWM. The bending angle can reach its maximum of 33°, and the relationship between the banding angle of the biomimetic squid funnel and the electrical resistance of SMA wire is approximately inverse linear. The biomimetic squid funnel verified its ability of control of the jet direction through a simulated jetting experiment.

Acknowledgment

This work is supported by the National Natural Science Foundation of China under grant No. 50775049 and by Self-Planned Task (NO. SKLRS200805C) of State Key Laboratory of Robotics and System (HIT).

References

[1] Z.L. Wang, J. LI, G.R. Hang and Y.W. Wang: Robot, Vol. 31 (2009) No.3, pp.281-288.(in Chinese)

[2] Sung-Weon Yeom and Il-Kwon Oh: Smart Materials And Structures, Vol. 18 (2009) No.8, pp.1-10.

[3] S.X. Guo, L.W. Shi, X.F. Ye and L.F. *Li: Proceedings of the 2007 IEEE International Conference on Mechatronics and Automatio* (Harbin, China, August 5-8, 2007). pp.509-514

[4] S. Gao: *Study on Course/Speed Control of Ship with Waterjet Propulsion* (Ph.D., Harbin Engineering University, China, 2008) , pp.25-26. (in Chinese)

[5] J.H. Xie and W.G. Zhang: Chinese Journal of Sensors and Actuators, Vol. 17 (2004) No.1, pp.164-167. (in Chinese)

[6] Mingyen Ho and Jaydev P.Desai: *31st Annual International Conference of the IEEE EMBS* (Minneapolis, Minnesota, USA, September 2-6, 2009). pp.6856-6859.

[7] http://smart.tamu.edu/overview/smaintro/detailed/detailed.html

[8] Ma. N, Song. G: Smart Materials and Structures, Vol. 12 (2003) No.5, pp.712-715

[9] Information on http://www.absoluteastronomy.com/topics/Muscular_hydrostat

doi:10.4028/www.scientific.net/AMM.50-51.224

Convergence theorems for accretive operators in Banach spaces

Yan Guo[1,a], Jiwei He[2,b] and Yanmei Yang[3,c]

[1,2]Department of Mathematics and Physics, North China Electric Power University, Baoding, 071003, China;

[3]College of Science, Heibei Polytechnic University, Tangshan, 063009 , China

[a] guoyan12@126.com, [b] hejiwei313@163.com [c] yym04@163.com

Keywords: Inverse-strongly accretive; Nonexpansive mapping; Variational inequality; 2-uniformly smooth Banach Space.

Abstract. In this paper, a new iterative scheme for finding a common element of the set of fixed points of a nonexpansive mapping and the set of the variational inequality for an inverse-strongly accretive operator in a 2-uniformly smooth Banach space was introduced . It was verified that the sequence converged to a common element of two sets .

Introduction and Preliminaries

Let E be a smooth Banach space with norm $\|\cdot\|$, E^* denote the dual of E, and $\langle x, f\rangle$ denote the value of $f \in E^*$ at $x \in E$. Let C be a nonempty closed convex subset and A be an accretive operator of C into E. The variational inequality problem is to find a solution $u \in C$ such that $\langle Au, J(v-u)\rangle \geq 0, \quad \forall v \in C$, where J is the duality mapping of E into E^*.

Let $S_E = \{v \in E : \|v\| = 1\}$. A Banach space E is said to be uniformly convex if for each $\varepsilon \in (0,2]$, there exists $\delta > 0$ for each $x, y \in S_E$,

$$\| x - y \| \geq \varepsilon \quad \text{implies} \quad \| \tfrac{x+y}{2} \| \leq 1 - \delta \tag{1.1}$$

A Banach space E is said to be smooth if $\lim_{t\to 0} \frac{\|x+ty\|-\|x\|}{t}$ exists for all $x, y \in S_E$.. The form of E is said to be *Frećhet differentiable* if for each $x \in S_E$, $\lim_{t\to 0} \frac{\|x+ty\|-\|x\|}{t}$ is attained uniformly for $y \in S_E$. And we define a function $\rho : [0,\infty) \to [0,\infty)$ called the modulus of smoothness of E as follow:

$$\rho(t) = \sup\left\{\tfrac{1}{2}\left(\|x+y\| + \|x-y\|\right) - 1 : x, y \in E, \|x\| = 1, \|y\| = t\right\} \tag{1.2}$$

It is known that E is uniformly smooth if and only if $\lim_{t\to 0} \frac{\rho(t)}{t} = 0$. Let q be a fixed real number with $1 < q \leq 2$. Then a Banach space E is said to be *q-uniformly smooth* if there exists a

constant $c>0$ such that $\rho(t)\le ct^q$ for all $t>0$. For example, see[1] for more details.

Let C be a nonempty closed convex subset of a smooth Banach space E. For $\lambda>0$, a operator A of C into E is said to be α-*inverse strongly accretive* if for all $x,y\in C$

$$\langle Ax-Ay, j(y-x)\rangle \ge \alpha \| Ax-Ay\|^2 \tag{1.3}$$

where $j(y-x)$ is normalized duality mapping. It is obvious that $\|Ax-Ay\|\le \frac{1}{\alpha}\|x-y\|$. The variational inequalities can be characterized by using sunny nonexpansive retractions [2].

Lemma 2.1: Let C be a nonempty closed convex subset of a smooth Banach space E. Let Q_C be a sunny nonexpansive retraction from E onto C and let A be an accretive operator of C into E. Then for all $\lambda>0$, $\mathrm{VI}(C,A)=F(Q_C(I-\lambda A))$, where $\mathrm{VI}(C,A)=\{u\in C:\langle Au, J(v-u)\rangle\ge 0, \forall v\in C\}$.

Lemma2.2 Let C be a nonempty closed convex subset of a 2-uniformly smooth Banach space E. Let $\lambda>0$ and A be an α-inverse strongly accretive operator of C into E. If $0<\lambda\le \alpha/K^2$, then $I-\lambda A$ is a nonexpansive mapping of C into E, where K is the 2-uniformly smoothness constant of E.

Proof: From [3] we can get that for all $x,y\in C$,

$$\begin{aligned}\|(I-\lambda A)x-(I-\lambda A)y\|^2 &\le \|x-y\|^2-2\lambda\langle Ax-Ay, J(x-y)\rangle+2K^2\lambda^2\|Ax-Ay\|^2\\ &\le \|x-y\|^2-2\lambda\alpha\|Ax-Ay\|^2+2K^2\lambda^2\|Ax-Ay\|^2\\ &\le \|x-y\|^2+2\lambda(K^2\lambda-\alpha)\|Ax-Ay\|^2.\end{aligned}\tag{1.4}$$

Note that $0<\lambda\le\alpha/K^2$, so $I-\lambda A$ is a nonexpansive mapping of C into E.

Main results

In this paper, we introduce a new iterative scheme for a nonexpansive mapping S and an inverse-strongly accretive operator A in a Banach space. Then we prove a strong convergence. We generalize the results of Iiduka [4] into 2-uniformly smooth Banach spaces, and our proof is more succinctness than [4].

Theorem3.1: Let E be a uniformly convex and 2-uniformly smooth Banach space with the best smooth constants K and let C be a nonempty closed convex subset of E. Let A be an α-inverse strongly accretive operator of C into E and let $S:C\to C$ be a nonexpansive mapping such that $F(S)\cap \mathrm{VI}(C,A)\ne\phi$, where $F(S)$ is the set of fixed points of S. For arbitrary $x_1=x\in C$, define a sequence $\{x_n\}$ by

$$x_{n+1}=\alpha_n x+(1-\alpha_n)\left[\tfrac{1}{2}Sx_n+\tfrac{1}{2}Q_C(x_n-\lambda_n Ax_n)\right] \tag{2.1}$$

where $Q_C : E \to C$ is a sunny nonexpansive retraction, $\{\alpha_n\}$ and $\{\lambda_n\}$ are real sequences in (0,1) satisfying:

(i) $\lambda_n \in [a,b] \subset (0, \frac{\alpha}{K^2}]$ and $\sum_{n=1}^{\infty} |\lambda_{n+1} - \lambda_n| < \infty$;

(ii) $\lim_{n\to\infty} \alpha_n = 0$, $\sum_{n=1}^{\infty} \alpha_n = \infty$ and $\sum_{n=1}^{\infty} |\alpha_{n+1} - \alpha_n| < \infty$

then $\{x_n\}$ converges strongly to $Q_{F(S)\cap \mathrm{VI}(C,A)}x$.

Proof: Put $y_n = \frac{1}{2}Sx_n + \frac{1}{2}Q_C(x_n - \lambda_n Ax_n)$. For $\forall u \in F(S) \cap \mathrm{VI}(C,A)$, it follows from Lemma 2.1 and Lemma 2.2 that

$$\begin{aligned} \|y_n - u\| &= \tfrac{1}{2}\|(Sx_n - u) + [Q_C(x_n - \lambda_n Ax_n) - Q_C(u - \lambda_n Au)]\| \\ &\le \tfrac{1}{2}\|x_n - u\| + \tfrac{1}{2}\|(x_n - \lambda_n Ax_n) - (u - \lambda_n Au)\| \le \|x_n - u\| \end{aligned} \tag{2.2}$$

It follows from (2.2) that

$$\begin{aligned} \|x_{n+1} - u\| &= \|\alpha_n (x - u) + (1 - \alpha_n)(y_n - u)\| \\ &\le \alpha_n \|x - u\| + (1 - \alpha_n)\|x_n - u\| \le \|x - u\| \end{aligned} \tag{2.3}$$

Therefore, $\{x_n\}$ is nonincreasing and hence $\lim_{n\to\infty}\|x_n - u\|$ exists. So $\{x_n\}$ is bounded. Thus $\{y_n\}$ and $\{Ax_n\}$ are bounded. Suppose $\{Ax_n\}$ is bounded by M_1.

Next we will show $x_{n+1} - x_n \to 0$. Note that

$$\begin{aligned} \|y_{n+1} - y_n\| &\le \tfrac{1}{2}\|Sx_{n+1} - Sx_n\| + \tfrac{1}{2}\|Q_C(x_{n+1} - \lambda_{n+1}Ax_{n+1}) - Q_C(x_n - \lambda_n Ax_n)\| \\ &\le \tfrac{1}{2}\|x_{n+1} - x_n\| + \tfrac{1}{2}\|(x_{n+1} - \lambda_{n+1}Ax_{n+1}) - (x_n - \lambda_n Ax_n)\| \\ &\quad + (x_n - \lambda_{n+1}Ax_{n+1}) - (x_n - \lambda_n Ax_n)\| \\ &\le \tfrac{1}{2}\|x_{n+1} - x_n\| + \tfrac{1}{2}\|(I - \lambda_{n+1}A)(x_{n+1} - x_n)\| + \tfrac{1}{2}\|(\lambda_{n+1} - \lambda_n)Ax_n\| \\ &\le \tfrac{1}{2}\|x_{n+1} - x_n\| + \tfrac{1}{2}\|x_{n+1} - x_n\| + \tfrac{1}{2}|\lambda_{n+1} - \lambda_n|\|Ax_n\| \\ &\le \|x_{n+1} - x_n\| + \tfrac{M_1}{2}|\lambda_{n+1} - \lambda_n|. \end{aligned} \tag{2.4}$$

And then from this we have

$$\begin{aligned} \|x_{n+1} - x_n\| &= \|\alpha_n x + (1 - \alpha_n)y_n - (1 - \alpha_n)y_{n-1} + (1 - \alpha_n)y_{n-1} - \alpha_{n-1}x - (1 - \alpha_{n-1})y_{n-1}\| \\ &\le |\alpha_n - \alpha_{n-1}|\|x\| + (1 - \alpha_n)\|y_n - y_{n-1}\| + |\alpha_n - \alpha_{n-1}|\|y_{n-1}\| \\ &\le |\alpha_n - \alpha_{n-1}|M_2 + (1 - \alpha_n)\left(\|x_n - x_{n-1}\| + \tfrac{M_1}{2}|\lambda_n - \lambda_{n-1}|\right) \\ &\le (1 - \alpha_n)\|x_n - x_{n-1}\| + M_2|\alpha_n - \alpha_{n-1}| + \tfrac{M_1}{2}|\lambda_n - \lambda_{n-1}| \end{aligned} \tag{2.5}$$

where $M_2 = \max\{\|x\| + \|y_{n-1}\|\}$. Put $a_n = \|x_n - x_{n-1}\|$, $b_n = M_2|\alpha_n - \alpha_{n-1}|$, $c_n = \frac{M_1}{2}|\lambda_n - \lambda_{n-1}|$, then the last inequality (2.5) reduces to $a_{n+1} \le (1 - \alpha_n)a_n + b_n + c_n$, where $b_n = o(\alpha_n)$ and $\sum_{n=1}^{\infty} c_n < \infty$. By [5], we

conclude that $a_n \to 0$.

Following this and the inequality (2.4) we have $\|y_n - y_{n-1}\| \to 0$. It is obvious that

$$\begin{aligned} \|x_n - y_n\| &= \|\alpha_{n-1}(x - y_{n-1}) + (y_{n-1} - y_n)\| \\ &\le \alpha_{n-1}\|x - y_{n-1}\| + \|y_{n-1} - y_n\| \end{aligned} \tag{2.6}$$

Combining the condition $\|y_n - y_{n-1}\| \to 0$ and $\alpha_n \to 0$ we have $\lim_{n\to\infty}\|x_n - y_n\| = 0.$

Since $\{\lambda_n\}$ is bounded, there exists a subsequence $\{\lambda_{n_i}\}$ convergent to $\lambda \in [a,b]$. Without loss of generality it may assume that $\lambda_n \to \lambda$. Define a mapping $T : C \to C$ by $Tx = \frac{1}{2}Sx + \frac{1}{2}Q_C(I - \lambda A)x$ for all $x \in C$. Then we can prove T is nonexpansive and $F(T) = F(S) \cap \mathrm{VI}(C, A)$. By a result of Reich [6], for $x \in C$ and each $t \in (0,1)$, there exists a unique $x_t \in C$ such that $x_t = tx + (1-t)Tx_t$ and x_t converges strongly to $z = Q_{F(T)}(x)$ as $t \to 0$. From T is nonexpansive we have

$$\begin{aligned} \|x_t - y_n\| &= t\langle x - x_t, j(x_t - y_n)\rangle + t\langle x_t - y_n, j(x_t - y_n)\rangle + (1-t)\langle Tx_t - Ty_n, j(x_t - y_n)\rangle \\ &\quad + (1-t)\langle Ty_n - y_n, j(x_t - y_n)\rangle \\ &\le t\langle x - x_t, j(x_t - y_n)\rangle + t\|x_t - y_n\|^2 + (1-t)\|x_t - y_n\|^2 + (1-t)\|Ty_n - y_n\|\|x_t - y_n\| \end{aligned} \tag{2.7}$$

Hence $0 \le t\langle x - x_t, j(x_t - y_n)\rangle + (1-t)\|Ty_n - y_n\|M_3$, where $M_3 = \max\{\|x_t - y_n\|\}$. This is equivalent to $\langle x - x_t, j(y_n - x_t)\rangle \le \frac{1-t}{t}M_3\|Ty_n - y_n\|$, where $\forall t \in (0,1)$, $\forall n \ge 1$. Put $T_n = \frac{1}{2}S + \frac{1}{2}Q_C(I - \lambda_n A)$, then $y_n = T_n x_n$. So

$$\|Ty_n - y_n\| = \|Ty_n - T_n y_n + T_n y_n - y_n\| \le \|Ty_n - T_n y_n\| + \|y_n - x_n\| \tag{2.8}$$

Therefore, for any given $t \in (0,1)$, we obtain $\limsup_{n\to\infty}\langle x - x_t, j(y_n - x_t)\rangle \le 0$. Let $t \to 0$, we have that $\limsup_{t\to 0}\limsup_{n\to\infty}\langle x - x_t, j(y_n - x_t)\rangle \le 0$. Since E is uniformly smooth, then the norm of E is uniformly Fréchet differentiable and uniformly Gâteaux differentiable. Thus

$$\begin{aligned} \limsup_{n\to\infty}\langle x - z, j(y_n - z)\rangle &= \limsup_{n\to\infty}\limsup_{t\to 0}\langle x - x_t, j(y_n - x_t)\rangle \\ &= \limsup_{t\to 0}\limsup_{n\to\infty}\langle x - x_t, j(y_n - x_t)\rangle \le 0 \end{aligned} \tag{2.9}$$

Following this we can get

$$\begin{aligned}\|x_{n+1}-z\|^2 &\le \|\alpha_n(x-z)+(1-\alpha_n)(y_n-z)\|^2 \\ &\le (1-\alpha_n)^2\|y_n-z\|^2+2\alpha_n(1-\alpha_n)\langle x-z, j(y_n-z)\rangle+2K^2\alpha_n{}^2\|x-z\|^2 \\ &\le (1-\alpha_n)\|x_n-z\|^2+2\alpha_n\sigma_n+2K^2\alpha_n{}^2\|x-z\|^2 \\ &= (1-\alpha_n)\|x_n-z\|^2+o(\alpha_n),\end{aligned} \tag{2.10}$$

where $\sigma_n = \max\{0, \langle x-z, j(y_n-z)\rangle\} \to 0$. Put $a_n = \| x_n - z \|$, then

$$a_{n+1} \le (1-\alpha_n)a_n + o(\alpha_n).$$

Weng's lemma [7] implies $a_n \to 0$, that is, $x_n \to z$. The proof is completed.

References

[1] K. Ball, E. A. Carlen and E. H. Lieb: *Sharp uniform convexity and smoothness inequalities for trace norm*, Invent Math Vol.115 (1994) p.463.

[2] S. Reich: *Asymptotic behavior of contractions in Banach spaces*, J Math Anal Appl Vol. 44 (1973) p.57.

[3] H. K. Xu: *Inequalities in Banach spaces with applications*, Nonl Anal Vol.16 (1991) p.1127.

[4] H. Iiduka and W. Takahashi: *Strong convergence theorems for nonexpansive mappings and inverse-strongly monotone mappings*, Nonl Anal Vol.61 (2005) p.341.

[5] L. S. Liu: *Ishikawa and Mann iterative process with errors for nonlinear strongly accretive mappings in Banach spaces*, J Math Ana1 App1 Vol.194 (1995) p.114.

[6] S. Reich: *Strong convergence theorems for resolvents of accretive operatos in Banach spaces*, J Math Anal Appl Vol.75 (1980) p.287.

[7] X. L. Weng: *Fixed point iteration for local strictly pseudo-constractive mapping*s, Proc Math Soc Vol.1l3(3) (1991) p.727.

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.229

The working principle of Multi-layer electromagnetic screens and the dynamics analysis

LIU Weisheng[1, a], ZHANG Jinli[2,b]

[1] Department of phyisics, Tangshan teacher's college, Tangshan, China

[2] College of mathematics and information science, Qinghai normal university

[a]wsliu482@sohu.com, [b]zhangjl_happy@126.com

Keywords: Electromagnetic screen. Vibrator. Electromagnetic vibration.

Abstract. Introducing the structures of Multi-layer electromagnetic screens and elaborating the working principle of Multi-layer electromagnetic screens, we have carried on the computation to the related parameter in the foundation of dynamics analysis. The designing adopts the vibration modes and principles of the transmission, takes electro-magnetic vibrator as vibration source, encourages to drive four vibration shafts through the vibration transmission system at the same time, and eventually the materials are passed then sifted out the screen mesh. The products have small space occupation, the high vibration intensity, large capacity, easy to operate, stable performance, high efficiency, low power consumption, small dynamic and self cleaning sieves, etc.

Introduction

Multi-layer electromagnetic screens are a new type of mechanical solid materials screen. They adopt the vibration modes and principles of the transmission, take electro-magnetic vibrator as vibration source, encourage to drive four vibration shafts through the vibration transmission system at the same time, and eventually the materials are passed then sifted out the screen mesh.

The products have small space occupation, the high vibration intensity, large capacity, easy to operate, stable performance, high efficiency, low power consumption, small dynamic and self cleaning sieves, etc. er.

Working principles and Structure outline

Working principles. Multi-layer screens take electromagnetic vibrator as powered sources, and driving gears drive vibration mechanism to vibrate screen mesh but the screen boxes basically not intact. At the same time, the vibration system is in a State of near-resonance, so smaller driving forces will be able to meet the needs of amplitude. Vibration structures see Figure 1, Figure 2.

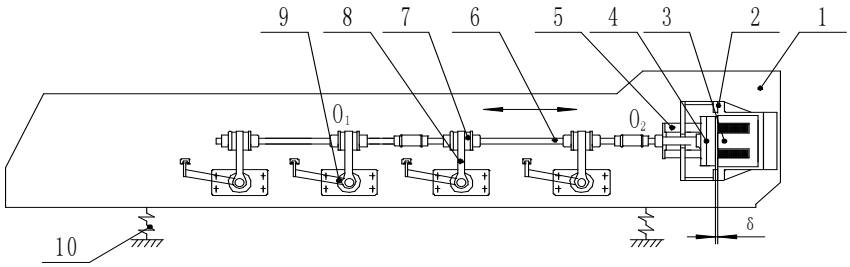


Fig. 1 drawing of Vibration structures

In the figure: 1-screen box, 2-shell of vibration generator,3-electromagnet , 4-bolt components of gag bit dynamical conditions, 5-The rubber host shakes spring,6-the force transmission,7-rubber spring 8-Vibrating rod, 9-Rubber torsion spring, 10-Vibration isolation spring.

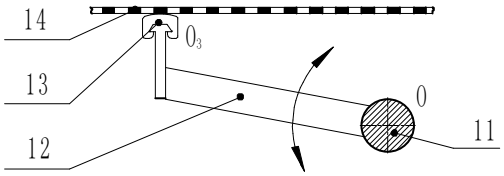


Fig. 2 drawing of Vibration

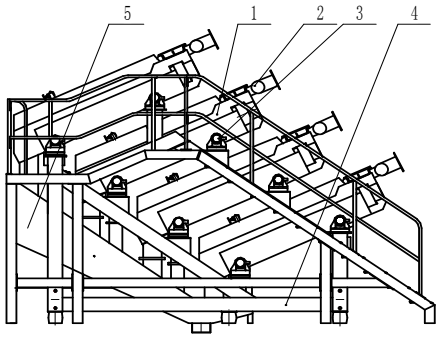


Fig.3 The structure electromagnetic screen

In the figure: 11-Vibration axis, 12-Vibration arm, 13-the cap of mount vibration apparatus,14−screen mesh

The electro-magnetic vibrators are composed by 2 shell of vibration generator, 3 electromagnet, 4 bolt components of gag bit dynamical conditions, 5 the rubber host shakes spring. Shell of vibration generator 2 is fixed the outside of screen boxes'1 side panel, electromagnet combined with shell of vibration generator, bolt components of gag bit dynamical conditions separately located at two rubber host to inspire the spring of shell's inside and outside, installed in the shell, among the electromagnets and bolt components of gag bit dynamical conditions remain far airgap δ.

Bolt components of gag bit dynamical conditions are fixed on the end of the O_2 point of the force transmission 6.The end of 8 vibrating rod articulates with 7 the rubber host shakes spring at O_1 point,and the other end of 8 vibrating rod is fixed on 9 vibration axis.9 vibration axis is across the screen box, and it puts below the screen mesh in a level which is fixed on several vibration arms 12 on a certain angle which loads 13 the cap of mount vibration apparatus .

Electricity flows through silicon-controlled rectifier half-wave rectification input coil, the closure magnetic which is composed by electromagnet and gag bit generates alternation magnetic flux. Electric magnet generates the electromagnetic force which makes the armature of frequencies by 50hz[2]. The gag bits' vibration drive O_1 the end of 8 vibrating rod through the force transmission 6, also the 8 vibrating rod drives 11 vibrating rod making the torsional vibration take O spot as the center of rotation, thus vibrating arm 12 and the cap of mount vibration apparatus 13 as in the figure a in the direction of reciprocating vibration (a to perpendicular to the vibration arm).

Structure outlines:

The gross structure of electromagnetic screen is as shown in Figure 3.

In the figure: 1-screen box, 2-feeding box, 3-Tilting fillet antivibration combination, 4-rack, 5-Receiving vat

Laminated screen box is divided into single and double with two forms. A single screen box only has a channel which only has installed a vibrator, duplex has two channels on both sides of which is equipped with a vibrator. The screen mesh is composed by three nets, the following contacts with the vibration hat is holds the net, generally for steel wire core polyurethane net or big aperture steel wool.

What above with material contact is the working net, is being composed together by two different aperture's strainer cementations. This kind of strainer has light quality, and the percentage

open area is high. The feeding box is transported the materials, which causes the materials along the screening surface uniform distribution, full screening. Because the feeding box has different basis structures, so it join with sieves by the different way. Tilting fillet antivibration combination and rack is connected with the main parts, installed in the tube beam supported by the screen box. The tilting fillet antivibration combination only plays screen box and rack's elastic joint role, simultaneously may have the adjustment block adjust the angle of the screen box.

Vibrating system's dynamics analysis and parameter computation

Dynamics analyze [3]. In the vibration partial quality system the vibrating arm, inspires the moving axis, to vibrate the arm, the rapid installation to vibrate hat's movement is take inspires moving axis's axle center O spot to make the twisting motion as the center.

According to the translation principle ,it is equal to translate into an equivalent quality O_1 points to the left, right, translational motion, that is the O point of inertia J is equal to the O_1 point of an equivalent quality m_{o1} product with the square of the O_1O distance: $J = m_{o1} \times (\overline{O_1O})^2$, $m_{o1} = \frac{J}{\overline{O_1O}^2}$; Supposes the first quality system's quality is M_1, the second quality system's quality is M_2.

According to the above principle, reverses spring's torsional rigidity is also equivalent to transform the O_1, which is compression spring's rigidity. In elastic system, main vibration rubber spring 5 and rubber torsion spring 9 is parallel. Therefore between M_1 and M_2 has two springs parallel, its total rigidity inspires the spring 5 rigidities and the rubber reverse spring for the rubber host 9 sum in the equivalent rigidity.

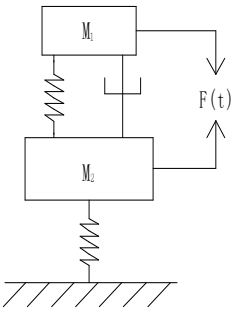


Fig. 4 Vibrating system's mathematical model

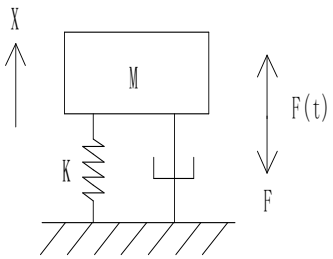


Fig.5 Forced vibration system of one freedom degree

According to the above analysis, vibrating system's mathematical model as shown in Figure 4, this is a double quality and two degree-of-freedom forced oscillation systems. Because the quality M_2 is bigger than the quality M_1, therefore the quality M_2 oscillation amplitude must be smaller than M_1 far the oscillation amplitude, M_2 approximate motionless, the small oscillation amplitude passes to the ground through the isolation spring, to move the load to be small, which is approximate in overall system's static load.

Appropriate identification M_1, M_2 main vibration between stiffness of the spring, you can make the system of vibration which is near its vibration, allowing smaller force to obtain necessary amplitude, power consumption can reduce.

Kinetic Parameters Calculation

Double mass forced vibration mathematical model of two freedom degrees are shown in fig.5. It can be converted into one freedom degree forced vibration system, which mass is *M* is shown in fig.6,and the kinematics differential equation is

$$M\frac{dx^2}{dt^2}+f\frac{dx}{dt}+kx=F(t)^{[4]} \qquad (1)$$

in the formula: *M*—Induction quality $M=\frac{M_1\times M_2}{M_1+M_2}$

f—Relative damping coefficient

K—stiffness of spring

F(t)—Vibration force

the particular solution is: $x=\lambda\sin(\omega t-\alpha)$ (2)

in the formula: *λ*—Relative Amplitudes between *M1* and *M2*

ω—rotary angular velocity vibration axis

t—time

α—phase difference between relative movement and exciting force

In this screen machine, electromagnetic force obtains by silicon half wave rectification controlled, which is periodic, discontinuity, and inharmonic, of which cycle is 2π. And the electromagnetic force is consisting of average electromagnetic force, first harmonic and secondary harmonic.

$$F(t)=\frac{3}{2}F_a+2F_a\sin\omega t_1+\frac{1}{2}F_a\sin 2\omega t_2 \qquad (3)$$

In the formula: F_a—Basic electromagnetic

ω—Supply angular frequency

t_1—the time of the harmonic excitation force

t_2—the time of the second harmonic excitation

putting (2) and (3) formula into(1)formula, the λ is obtained:

$$\lambda=\frac{F_a}{K}\frac{1}{\sqrt{(1-Z_0^2)^2+4b^2Z_0^2}} \qquad (4)$$

In the formula: λ—relative Amplitudes between *M1* and *M2*

K—spring staff

F_a—fundamental electromagnetic force

Z_0—frequency ratio, Z_0=0.9

b—relative damping ratio, b=0.05

From the (4) formula, we can obtain:

Fundamental electromagnetic force: $F_a=\lambda K\sqrt{(1-Z_0^2)^2+4b^2Z_0^2}$

Phase difference: $\alpha=\operatorname{arctg}\frac{2bZ_0}{1-Z_0^2}$

Innovations and application value

The structural design of multi-layer electromagnetic screens can enhance the single performances and screening efficiency. In addition, it is an effective way to reduce the equipment land area, and increases the work capacity in unit area, improving the performance and price.

The electro-magnetic vibrator drive the groups vibration structure to vibrate the screen mesh with drive devices, which forms a high performance electro-magnetic vibrator. So this device has the virtue of stringently screening, and high efficiency.

Dynamic characteristic of vibration structures are easy to converge, and the equipment adjustment decreases. There's been a measureable improvement in new electromagnetic screen, small quantity, low cost and low energy consumption.

The screen products of every screen box flow above and under of screen, then assemble together to come out.

The test results and screening operation indicates that this device has the virtue of smooth runs, high processing ability and high efficiency. So it has broad application prospect in in many fields, for example, concentrates in particular in ore dressing industry, the dry wet screening of metal and nonmetallic ore powder, drives technology innovation, and gains great economic benefits and social benefits.

References

[1] Gao Song,Yan Zhiquan,Zhu Manping. 3DZS2418 Electromagnetic High Frequency Vibrating Screen [J].coal mine machine,2001,(1), p,125-126.

[2] Jiang Shuxing,Huang Shengsheng.The sieving performances of the standard test sive shanker of the electromagnetic Vibrative pattern[J].GuiLin college of technology journal, 2001, 1(1),p, 88-90.

[3] Lu Zhingming, Liang Jingan, An Lianhong. Research on compound electro-magnetic vibrator high-frequency coal slime screen [J].ore dressing techonolgy, 2008,1(6), p,18-19.

[4] Ti Anshen, Du Shanying. mechanics[M]. BeiJing: high education press, 2005.

© (2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.234

Stability Analysis of Finite Element Method for Stokes Problem on Overlapping Non-matching Grids*

DENG Kang [1], WANG Qisheng [2,a]

[1]Mathematics and Computing Science College, Hunan University of Science and Technology, China, Xiangtan 411201

[2]School of Mathematics and Computational Science, Wuyi University, China, Jiangmen, 529020

[a]Correspondence author: Wang Qisheng, email: wqs9988@yahoo.cn

Keywords: Stokes problem, finite element method, overlapping non-matching grids, partition of unity, stability analysis..

Abstract. In this paper, the steady-state Stokes problem is discussed on a plane polygonal domain $\Omega \in R^2$. We propose a finite element method on overlapping non-matching grids for the Stokes problem based on the partition of unity method. The construction and stability analysis of a global conforming finite element method are studied, and our results are valid for any spatial dimensions.

Introduction

We consider the steady-state Stokes Problem on a plane polygonal domain $\Omega \in R^2$

$$\begin{cases} -\Delta u - \nabla p = f, \quad in\ \Omega, \\ \nabla u = 0, \quad in\ \Omega, \\ u = 0, \quad on\ \partial\Omega, \\ \int_\Omega p(x,y)dxdy = 0. \end{cases} \tag{1}$$

Where $u = (u_1, u_2)^T$ denotes the velocity vector function, p represents the pressure scalar function. Let Ω be a bounded domain with boundary $\partial\Omega$ and Γ be a closed subset of $\partial\Omega$. By $H_0^1(\Omega;\Gamma)$ we denote the closure in H^1-topology of $C^\infty(\overline{\Omega})$ functions that vanish in a neighborhood of Γ, simply $H_0^1(\Omega) = H_0^1(\Omega;\partial\Omega)$. Define space of velocity and pressure function $V = (H_0^1(\Omega))^2 = \{u = (u_1,u_2) \mid u_1, u_2 \in H_0^1(\Omega)\}$, $P = \{p \mid p \in L^2(\Omega), \int_\Omega p(x,y)dxdy = 0\}$. The variational formulation of the problem (1) is, find the pairs $(u,p) \in (V,P)$, such that

$$\begin{cases} a(u,v) + b(v,p) = (f,v), \forall v \in V, \\ b(u,q) = 0, \forall q \in P. \end{cases} \tag{2}$$

Where a bilinear form $a(\cdot,\cdot)$ is the Dirichlet form onΩ been definded by $a(u,v) = \int_\Omega \nabla u \cdot \nabla v dxdy$, and $b(v,q) = (q, \nabla v) = -(\nabla q, v)$. We assume that the $Babuska - Brezzi$ condition

$$\sup_{v \in V} \frac{(p, \nabla v)}{\| v \|_{1,\Omega}} \geq C_0 \| p \|, \forall p \in P, \tag{3}$$

holds for a positive constant C_0.Consequently, there is a unique solution $(u,p) \in (V,P)$ of the problem (2).

In recent years, there have been considerable interest in the use of overlapping grids.Using a partition of unity technique on overlapping nonmatching grids (see [1-3],[5-6]), the conforming finite element method has roots from Huang and Xu in [1]. The new finite element discretization for elliptic boundary value problems was introduced by Huang and Xu in [1]. A significant amount of literature was dedicated to numerical solution of the Stokes Problem (cf. [2, 4] and their references).

In this paper, following the ideas of Huang and Xu in [1] and [2],the stability analysis of a global conforming finite element for the Stokes Problem shall be given, and the results of literature [2] have been improved in some kind of significance.

Construction of the finite element spaces

Let Ω_1, Ω_2 be overlapping sub-domains of Ω satisfying $\Omega = \Omega_1 \cup \Omega_2$ and $\Omega_0 = \Omega_1 \cap \Omega_2$ .We further assume that Ω_1, Ω_2 are partitioned by quasi-uniform finite element triangulation (or quadrilateral) T_{h_1} and T_{h_2} of maximal mesh sizes h_1 and h_2 (with might not match on Ω_0). Again, just for the sake of simplicity, we assume that Ω_0 is a strip-type domain of width $d = O(h_1)$ and $h_1 \geq h_2, h_1 = O(h_2)$.

To obtain a conforming discretization of the variational problem (2), we define the local finite element subspaces as

$$\begin{cases} V^{h_i}(\Omega_i) = \{v \in (H_0^1(\Omega_i; \partial\Omega \cap \partial\Omega_i))^2, v|_E \in (P_1)^2, E \in T^{h_i}\}, \\ P^{h_i}(\Omega_i) = \{p \in C^0(\Omega_i),\ p|_E \in P_1, E \in T^{h_i}\}, \\ \hat{P}^{h_i}(\Omega_i) = \{p \in P^{h_i}(\Omega_i),\ p|_{\partial\Omega_i \setminus \partial\Omega} = 0\}, \end{cases} \tag{4}$$

where P_1 denotes the set of polynomials in two variables of degree at most one. On subdomain Ω_i , for the finite element pairs $(V^{h_i}(\Omega_i), P^{h_i}(\Omega_i))$ and $(V^{h_i}(\Omega_i), \hat{P}^{h_i}(\Omega_i))$, the $Babuska - Brezzi$ condition are satisfied, such that

$$\sup_{v \in V^{h_i}} \frac{(p, \nabla v)}{\| v \|_{1,\Omega_i}} \geq C_0 \| p \|, \forall p \in P^{h_i} \ (or \ \forall p \in \hat{P}^{h_i}), i = 1, 2. \tag{5}$$

The definition of Bubble Function Space $B^2_{T^h}$ associated with the partition $T^h = T^{h_1} \cup T^{h_2}$ shall be given (see section 3 and section 4). We can employ the partition of unity functions $\{\varphi_i\}$ to glue all the local subspaces together. Therefore, the first type finite element pairs is

$$\begin{cases} V^h = \sum_{i=1}^{2} \varphi_i V^{h_i}(\Omega_i) + B^2_{T^h}, \\ P^h = \sum_{i=1}^{2} \hat{P}^{h_i}(\Omega_i) \cap P(\Omega). \end{cases} \tag{6}$$

The second type finite element pairs is

$$\begin{cases} V^h = \sum_{i=1}^{2} \varphi_i V^{h_i}(\Omega_i) + B^2_{T^h}, \\ P^h = \sum_{i=1}^{2} P^{h_i}(\Omega_i) \cap P(\Omega). \end{cases} \tag{7}$$

In the next two sections, we shall obtain that the pairs of conforming finite element space $(V^h(\Omega), P^h(\Omega))$ as described in (6), (7), and the stability analysis of a global conforming finite element for the Stokes Problem shall be given.

The stability for the first type finite element pairs

For a triangle element $E \in T^h$, we define the first type bubble function $B_E(x, y)$ supported on E as the product of the nodal functions associated with the vertices of E , namely,

$$B_E(x, y) = \begin{cases} \lambda_1(x, y)\lambda_2(x, y)\lambda_3(x, y), & if \ (x, y) \in E, \\ 0, & if \ (x, y) \in \Omega \setminus E, \end{cases} \tag{8}$$

where $\lambda_i(x, y)$ is the nodal basis function with the ith-vertex of E . The first type bubble function space be called

$$B_{T^h}=\{B_E \mid E\in T^h\}. \tag{9}$$

Lemma 1[2]. For the pairs is defined as (6), (7), if exists two operators $\Pi_i : V\to V^h (i=1,2)$, with the
following properties

$$|v-\Pi_1 v|_{1,\Omega}\le C|v|_{1,\Omega},\quad forall\quad v\in V, \tag{10}$$

$$|\Pi_2(I-\Pi_1)v|_{1,\Omega}\le C|v|_{1,\Omega},\quad forall\quad v\in V, \tag{11}$$

$$b(v-\Pi_2 v,q)=0,\quad forall\quad v\in V, q\in P^h. \tag{12}$$

Then the $Babuska-Brezzi$ condition is satisfied in Ω.

Theorem 1. For the pairs is defined as (6), if overlapping sub-domain Ω_0 is a strip-type domain, and $\forall E\in T^h$, such that $|E|=O(h^2), |\Omega_0|=O(h)$. Then the first type finite element pairs (V^h, P^h) are a stable pairs.

Proof. For $V_i=(H_0^1(\Omega_i;\partial\Omega\cap\partial\Omega_i))^2$, thanks to standard theory of finite element approximations, $\forall v\in V_i, \exists\ \Pi_1^i : V_i\to V^{h_i}$, such that

$$\|v-\Pi_1^i v\|_{0,\Omega_i}\le Ch_i|v|_{1,\Omega_i}, \tag{13}$$

$$|v-\Pi_1^i v|_{1,\Omega_i}\le C|v|_{1,\Omega_i}. \tag{14}$$

Let $\Pi_1 v=\sum_{i=1}^{2}\varphi_i\Pi_1^i v$, such that $\Pi_1 : V\to V^h$, and

$$\|v-\Pi_1 v\|_{0,\Omega}\le\sum_{i=1}^{2}\|\varphi_i(v-\Pi_1^i v)\|_{0,\Omega_i}\le\sum_{i=1}^{2}\|(v-\Pi_1^i v)\|_{0,\Omega_i}\le C\sum_{i=1}^{2}h_i|v|_{1,\Omega_i}\le Ch|v|_{1,\Omega}. \tag{15}$$

Furthermore, $|v-\Pi_1 v|_{1,\Omega}\le\sum_{i=1}^{2}|\varphi_i(v-\Pi_1^i v)|_{1,\Omega_i}\le\sum_{i=1}^{2}|\nabla\varphi_i(v-\Pi_1^i v)|_{0,\Omega_i}+\sum_{i=1}^{2}|v-\Pi_1^i v|_{1,\Omega_i}$

$$\le C\sum_{i=1}^{2}(d^{-1}|v-\Pi_1^i v|_{0,\Omega_i}+|v|_{1,\Omega_i})\le C(d^{-1}h+1)\sum_{i=1}^{2}|v|_{1,\Omega_i}\le C|v|_{1,\Omega}.$$

Then the formula (10) is satisfied. Next, we define operator $\Pi_2, \Pi_2 v|_E=\alpha B_E(x,y), \forall v\in V, E\in T^h$, where $\alpha=(\alpha_1,\alpha_2)^T$ is determined such that $\int_E(v-\Pi_2 v)dxdy=0$, i.e., $\alpha=\int_E vdxdy/\int_E B_E dxdy$. Therefore, $forall\quad v\in V$, $q\in P^h$, we have

$|b(v-\Pi_2 v,q)|=|(\nabla q, v-\Pi_2 v)|\le\sum_{i=1}^{2}\sum_{E_i\in T_{h_i}}|\nabla q|\cdot|\int_{E_i}(v-\Pi_2 v)|=0$. Then the formula (12) is satisfied.

Thanks to (8), $\forall E\in T^h$, we have

$$|\Pi_2 v|_{1,E}^2=|\alpha|^2\cdot\int_E|\nabla B_E|^2\,dxdy\le\frac{1}{O(|E|^2)}\int_E v^2dxdy\cdot\int_E dxdy=\frac{|v|_{0,E}^2}{O(|E|)},\quad \text{and}$$

$$|\Pi_2 v|_{1,E}\le Ch^{-1}|v|_{0,E}. \tag{16}$$

Because of (15) and (16), we obtain

$$|\Pi_2(I-\Pi_1)v|_{1,\Omega}\le\sum_{i=1}^{2}\sum_{E_i\in T_{h_i}}|\Pi_2(I-\Pi_1)v|_{1,E_i}\le C\sum_{i=1}^{2}\sum_{E_i\in T_{h_i}}h^{-1}|v-\Pi_1 v|_{0,E_i}$$

$$\le\sum_{i=1}^{2}\sum_{E_i\in T_{h_i}}|v|_{1,E_i}\le C\sum_{i=1}^{2}|v|_{1,\Omega_i}\le C|v|_{1,\Omega}.$$

Which proves that (11) holds and concludes the proof of the theorem.

The stability for the second type finite element pairs

For right-angled triangle element $E \in T^h$, assume R is inscribed square in E, with sides that are $2l$ in length, such that $R = \{(x,y) \mid \bar{x} - l \le x \le \bar{x} + l, \bar{y} - l \le y \le \bar{y} + l\}$, and $(\bar{x}, \bar{y})$ is the center coordinate of square R. We define the second type bubble function $B_{E,j}(x,y)(j=1,2)$ supported on R as follow

$$B_{E,1}(x,y) = \begin{cases} \dfrac{(x-\bar{x}+l)^2(\bar{y}+l-y)(y-\bar{y}+l)}{16l^4}, & if\ (x,y) \in R, \\ 0, \quad if\ (x,y) \in \Omega \setminus R; \end{cases}$$
$$B_{E,2}(x,y) = \begin{cases} \dfrac{(\bar{x}+l-x)^2(\bar{y}+l-y)(y-\bar{y}+l)}{16l^4}, & if\ (x,y) \in R, \\ 0, \quad if\ (x,y) \in \Omega \setminus R. \end{cases} \tag{17}$$

The second type bubble function space be called $B_E = span\{B_{E,1}, B_{E,2}\}$.

Theorem 2. For the pairs is defined as (7), if overlapping sub-domain Ω_0 is a strip-type domain, and $\forall E \in T^h$, such that $|E| = O(h^2)$, $|\Omega_0| = O(h)$ .Then the second type finite element pairs (V^h, P^h) are a stable pairs..

Proof. The operator Π_1 and proof of (15) are similar to Theorem 1. For $\forall v \in V$ and right-angled triangle element $E \in T^h$, assuming $\Pi_2 : V \to V^h$, we have $\Pi_2 v|_E = \alpha_E B_{E,1}(x,y) + \beta_E B_{E,2}(x,y)$, where $\alpha_E = (\alpha_{1E}, \alpha_{2E})^T$ and $\beta_E = (\beta_{1E}, \beta_{2E})^T$, such that $\int_E (v - \Pi_2 v) \nabla q dxdy = 0$.

Let $\varphi_i(x,y) = a_i x + c_i\ (i = 1,2)$, $\forall q \in P^h$ and $E \in T^h$, then

$$\nabla q|_E \in span\{\begin{pmatrix} x-\bar{x} \\ 0 \end{pmatrix}, \begin{pmatrix} y-\bar{y} \\ x-\bar{x} \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\},$$

where $(\bar{x}, \bar{y})$ is the center coordinate of square R. Because $\Pi_2 v|_E = (\alpha_{1E} B_{E,1} + \beta_{1E} B_{E,2}, \alpha_{2E} B_{E,1} + \beta_{2E} B_{E,2})^T$,

Therefore

$$\begin{cases} (B_{E,1}, x-\bar{x})\alpha_{1E} + (B_{E,2}, x-\bar{x})\beta_{1E} = (v_1, x-\bar{x}), \\ (B_{E,1}, 1)\alpha_{1E} + (B_{E,2}, 1)\beta_{1E} = (v_1, 1), \\ (B_{E,1}, 1)\alpha_{2E} + (B_{E,2}, 1)\beta_{2E} = (v_2, 1), \\ (B_{E,1}, y-\bar{y})\alpha_{1E} + (B_{E,2}, y-\bar{y})\beta_{1E} + \\ (B_{E,1}, x-\bar{x})\alpha_{2E} + (B_{E,2}, x-\bar{x})\beta_{2E} = (v_1, y-\bar{y}) + (v_2, x-\bar{x}), \end{cases}$$

and

$$|\alpha_{1E}| = \frac{9}{2l^2} \left| \int_E (1 - 2(x-\bar{x})) v_1(x,y) dxdy \right| \le C \frac{\sqrt{|E|}}{l^2} |v|_{0,E}.$$

In the same way

$$|\alpha_{2E}|, |\beta_{1E}|, |\beta_{2E}| \le C \frac{\sqrt{|E|}}{l^2} |v|_{0,E}.$$

Therefore

$$|\Pi_2 v|_{1,E} \le C \frac{\sqrt{|E|}}{l^2} |v|_{0,E},$$

and

$$|\Pi_2(I-\Pi_1)v|_{1,\Omega} \le \sum_{i=1}^{2}\sum_{E_i\in T^{h_i}} |\Pi_2(I-\Pi_1)v|_{1,E_i} \le C\sum_{i=1}^{2}\sum_{E_i\in T^{h_i}} \frac{h_i\sqrt{|E_i|}}{l_i^2}|v|_{1,E_i} \le C\sum_{i=1}^{2}|v|_{1,\Omega_i} \le C|v|_{1,\Omega},$$

then (11) is satisfied, thus completes the proof.

References

[1] Yunqing Huang and Jinchao Xu. A conforming finite element method for overlapping and non-matching grids [J]. Math. Comput. 2002,72:1057-1066.

[2] C. Bacuta, J. Chen, Y. Huang, J. Xu, and L. Zikatanov. Partition of unity method on nonmatching grids for the Stokes problem[J]. J. Numer. Math., Vol. 13, No. 3, pp.157-169(2005).

[3] I.Babuska and J.M. Melenk. The partition of unity method [J]. Internat.J.Numer. Methods Engrg., 1997,40:727-758.

[4] V.Girault and P.A.Raviart. Finite Element Methods for Navier-Stokes Equations[M]. Springer-Verlag, Berlin, 1986.

[5] Qisheng Wang, Kang Deng, Zhiguang Xiong and Yunqing Huang. The convergence analysis of finite element method for parabolic equation on overlapping nonmatching grids. Advances in Mathematics, 2007(36), No6:737-748(in Chinese).

[6] Qisheng Wang, Kang Deng, Zhiguang Xiong and Yunqing Huang.Partition of Unity for a Class of Nonlinear Parabolic Equation on Overlapping Nonmatching Grids. Numer. Math. J. Chinese Univ.English Series, 2007(16), No1: 1-13.

*This work is supported by Natural Science Fund of China (No.61070150) and Guangdong Natural Science Fund of China (No.10452902001005845).
Corresponding author. This work is supported by National Natural Science Foundation of China Nos 10531080.

doi:10.4028/www.scientific.net/AMM.50-51.239

Estimation of Origin-Destination Matrix with Tolling Data

Hua Wang, Xiaoning Zhang

Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, shanghai 200092, China

ceh_wang@yahoo.com.cn, cexzhang@tongji.edu.cn

Keywords: origin-destination matrix estimation, tolling data, volumes on pricing entry-exit, travel time, bi-level programming.

Abstract. Prior matrix and surveyed link volumes were, in most cases, employed to estimate origin-destination matrix. With the development of BOT and of congestion pricing, charged links become an important component of road network, due to the fact that the tolling data: volumes and travel time on pricing entry-exit are traffic information, both cost-free and accurate. In this paper, we put forward a bi-level programming model, taking account of data on charging entry-exit to estimate OD matrix based upon the traditional model. Meanwhile, a heuristic method -the simulated annealing approach - is utilized to solve the OD estimation problem. Results of examples indicate that the accuracy of estimation will be improved while adding the tolling data, and that it is feasible to calculate OD matrix by combining the volumes and travel time on entry-exit with partial common link flows. In this light, this way can be applied to enhance accuracy, and also to reduce the cost spent on surveying the link flows in common OD matrix estimation.

Introduction

Origin-destination (OD) information is commonly deemed the important fundamental traffic data for transportation planning and management. In direct survey method of OD table there are some intractable problems (e.g., high cost, low sample rate, low statistical accuracy and long updated period). Thus, the estimation approach, which is depended on observed link flow and prior OD table, is used to get high effective and accurate OD demand.

There are some classical mathematical methods to estimate static OD matrix based on traffic link flows, such as maximum likelihood method, maximum entropy model, generalized least square model (GLS), Bayesian model and bi-level programming model. For instance, in order to solve the problem of sampling errors of observed link flows, Zhang and Lam presented O-D matrix estimation model combining maximum likelihood and Bayesian estimators[1]; Wang and Sun made use of an improved entropy maximum model and path iteration algorithm to estimate OD matrix on large-sized transport network[2]. Meanwhile, some researchers adopted GLS method to calculate OD matrix. For example, Martin L. Hazelton[3] discussed a statistical model of route choice proportions and developed a GLS OD estimation model depended on this route choosing method; Javier Doblas et al. estimated and updated observed OD matrices based upon available link flow information by means of the non-linear programming approach [4].

To estimate OD table, Fisk in 1989[5] firstly proposed the bi-level program, whose upper model always was expressed as the aforementioned models. Some pursuers focused on this method and commonly utilized GLS model to develop upper objective function. For example, in Yang's research, a bi-level programming model, which integrated the conventional GLS model and the standard network equilibrium model, has been adopted for the purpose of OD matrix estimation in congested networks[6]. Given a target OD matrix and link flows, Lundgren and Peterson[7] presented a bi-level model using GLS method to estimate OD matrix. Maher and Zhang formulated as bi-level programming models [8] the problems of trip matrix estimation and traffic signal optimization on congested road networks. In fact, the bi-level program is not only a mathematical optimization problem, but also a computation problem. Many effective algorithms, such as spread Spiess's gradient algorithm[4], heuristic iterative algorithm and sensitivity analysis method [6] and descent heuristic algorithm[7], have been developed to solve such problems.

The studiers mainly paid attention to link flows and prior OD matrix in both static and dynamic OD matrix estimations. Though some researchers have referred to entry-exit volume[9], they simply applied these data to calculate OD pairs directly or to adjust errors of link flows other than integrated them into the process of back-calculation. Actually, because the vehicles in different OD zones may operate in one entry or exit, the direct statistic of OD and calculation of link flows may be incorrect. With BOT development and congestion pricing implementation, it is easier and more precisely to collect traffic data on tolling entry-exits. The entry-exit flows are not only the linear maps of link flows but also are reflective of the path choice probabilities. Furthermore, the travel time on entry-exit is the subset of path cost that exerts a great impact on equilibrium traffic assignment. For these two reasons, flows and travel time on entry-exits are added into the OD matrix estimation model using least square method in the section2. The improved model may have practical meaning to reduce the cost spent on link flows collection to satisfy the models and to improve the accuracy of estimation.

The OD estimation model with tolling data

In this section, four kinds of traffic data were integrated into upper objective function with GLS method and the lower traffic assignment problem was modeled by the fixed demand deterministic user equilibrium program. The four specific data are prior OD demand, common link flows, volumes and travel time on tolling entry-exits. They are dealt as four dimensionless items, and named as the sum of relative error.

Upper model:

$$\min_{D} F(D) = \lambda_1 \sum_{w\in W}\left(\frac{D_w - \widetilde{D_w}}{\widetilde{D_w}}\right)^2 + \lambda_2 \sum_{b\in A'}\left(\frac{V_b(\mathbf{D}) - \overline{V_b^*}}{\overline{V_b^*}}\right)^2 + \lambda_3 \sum_{a\in A}\left(\frac{V_a(\mathbf{D}) - \overline{V_a^*}}{\overline{V_a^*}}\right)^2 + \lambda_4 \sum_{w0\in W0}\left(\frac{\tau_{w0}(\mathbf{D}) - \overline{\tau_{w0}^*}}{\overline{\tau_{w0}^*}}\right)^2 \quad (1)$$

Subject to:

$$\pi_1^w \bullet \widetilde{D_w} \le D_w \le \pi_2^w \bullet \widetilde{D_w},\ w \in W \quad (2)$$

Where D_w and $\widetilde{D_w}$ respectively represents the estimated traffic demand and prior traffic demand on OD pair w; $\mathbf{D}$ is the vector of D_w. Let $\overline{V_a^*}$ denote the actual volume of common link; $\overline{V_b^*}$ and $\overline{\tau_{w0}^*}$ stand for actual flow on tolling entry-exits and real travel time on tolling entry-exits respectively. And, $V_a(\mathbf{D})$, $V_b(\mathbf{D})$, $\tau_{w0}(\mathbf{D})$ are the optimal solution of the lower programming model. π_1^w and π_2^w are the lower and upper percentage bounds of each OD pair demand. The formula in the upper model includes four sections, which are the sums of relative error of corresponding traffic data item. Let $\lambda_i (i = 1,2,3,4)$ indicate the weight of each section.

Lower model:

$$\min_{v,\tau} \sum_{a\in A}\int_0^{v_a} t_a(\omega)d\omega + \beta \sum_{i\in I}\sum_{j\in J} q_{ij}x_{ij} \quad (3)$$

Subject to:

$$v_a = \sum_{w\in W}\sum_{p\in P_w} f_p^w \delta_{ap}^w,\ a \in A \quad (4)$$

$$\sum_{p\in P_w} f_p^w = D_w,\ w \in W \quad (5)$$

$$q_{ij} = \sum_{w\in W}\sum_{p\in P_w} f_p^w \delta_{ijp}^w,\ i \in I, j \in J \quad (6)$$

$$f_p^w \ge 0, p \in P_w, w \in W \quad (7)$$

Where travel time of each link t_a is a positive, strictly monotonously increasing and continuously differentiable function with respect to its own traffic flow v_a. The notation f_p^w represents the traffic flow on route $p \in R_w, w \in W$. The symbol δ_{ap}^w equals 1 if route p between O-D pair $w \in W$ uses link $a \in A$, and 0 otherwise. Let I denote the set of entry nodes and J the set of exit nodes; and let q_{ij} be the traffic flow between entry-exit pair $\langle i, j \rangle, i \in I, j \in J$ on the tolling sub-network (equate to $V_b(\mathbf{D})$); and x_{ij} stand for the tolling level. Where $\delta_{ijp}^w = 1$ if path p connecting OD pair $\langle r, s \rangle$ uses the tolling sub-network through entry-exit pair $\langle i, j \rangle$, and 0 otherwise.

There are many existing effective methods to solve the bi-level program, such as sensitivity analysis, simulated annealing method and genetic algorithms. In this paper, we use the simulated annealing method to solve the upper level program and the Frank-Wolfe method to solve the traffic assignment problem in a transformed network[10]. Given the charged network, how to transform the network for standard Frank-Wolfe method and the detailed procedure of SA please refer to reference [11].

Numerical example

Figure 1 is a network with 14 nodes and 46 inks. Links 33-38 are charged and the corresponding toll entry-exits are connected by node 4, 5, 6 and 7. The link travel time function is

$$t_a(v_a) = t_a^0\left(1 + 0.15\left(\frac{v_a}{C_a^0}\right)^4\right)$$

The values of t_a^0 and C_a for each link are given in Table 1. There are 8 OD pairs in the network, and the actual OD demands are shown in Table 2. The pricing levels for toll roads are provided in Table 3.

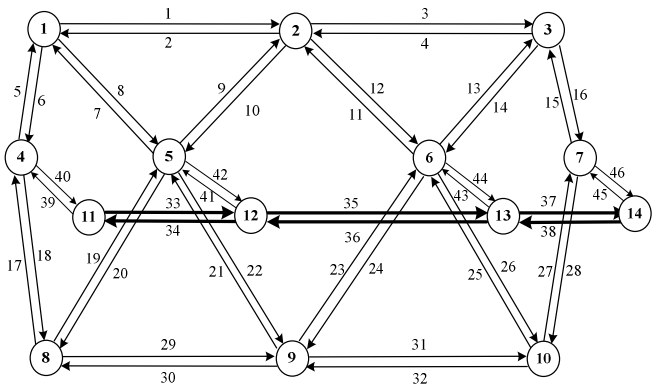


Fig1. Traffic network

Table 3. Charges of entry-exit on network

Entry node	Exit node			
	4	5	6	7
4	—	2	5	7
5	2	—	3	5
6	5	3	—	2
7	7	5	2	—

Table1. Input data of network

Link No.	t_a^0(min)	C_a(veh/h)	Link No.	t_a^0(min)	C_a(veh/h)	Link No.	t_a^0(min)	C_a(veh/h)
1	25	4000	17	2	6000	33	7	10000
2	25	4000	18	2	6000	34	7	10000
3	25	4000	19	9	4000	35	12	16000
4	25	4000	20	9	4000	36	12	16000
5	2	6000	21	10	4000	37	7	10000
6	2	6000	22	10	4000	38	7	10000
7	9	4000	23	9	4000	39	0.2	6000
8	9	4000	24	9	4000	40	0.2	6000
9	11	4000	25	10	4000	41	0.2	6000
10	11	4000	26	10	4000	42	0.2	6000
11	9	4000	27	2	6000	43	0.2	6000
12	9	4000	28	2	6000	44	0.2	6000
13	9	4000	29	24	3800	45	0.2	6000
14	9	4000	30	24	3800	46	0.2	6000
15	2	6000	31	26	4200			
16	2	6000	32	26	4200			

Table2. Actual traffic demand as prior OD matrix

O → D	1→3	1→10	3→1	3→8	8→3	8→10	10→1	10→8
Demand(veh/h)	10000	10000	9000	11000	10000	10000	12000	8000

The transformed network contains 10 nodes and a basic sub-network comprised of 32 initial links, and a pricing sub-network that involves 12 hypothetical links. And the new network is used to calculate the link flows and travel time by the lower UE model. Then, the link flows of original network can be obtained from the computation result of transformed network. Meanwhile, we take the UE assignment result of true OD demand as the actual statistical data ($\overline{V_b^*}, \overline{V_a^*}, \overline{\tau_{w0}^*}$) for the upper model. In an attempt to compare the impacts from tolling data and prior OD matrix, the true OD demands are treated as prior OD matrix. In such, there will be high precisions of estimation results. The discussion of different prior OD matrices and biased observed link flows can be seen in reference[6].

In this example, $\pi_1^w = 85\%, \pi_2^w = 115\%$, and weights of each data item for each group are given in Table 4. In order to compare with traditional model in terms of accuracy, the ratio of λ_1 and λ_3 is set as constant 1.0. The calculation under Group A is to analyze the effectiveness of volume on pricing entry-exit and the weights of Group B are employed to check the improvement of OD estimation while travel time data is added. In Group C, two data items are integrated simultaneously, and in Group D only partial uncharged link flows are used to estimate OD matrix. In Group E, uncharged link flows are not taken into consideration; i.e., common link volumes are completely replaced by entry-exit flows. The computation results of all groups are given in detail in Table 5 and 6.

Table 4. Weights for each group of upper model

	λ_1	λ_2	λ_3	λ_4		λ_1	λ_2	λ_3	λ_4
A1	0.5	0	0.5	0	B1	0.5	0	0.5	0
A2	0.4	0.2	0.4	0	B2	0.4	0	0.4	0.2
A3	0.35	0.3	0.35	0	B3	0.35	0	0.35	0.3
A4	0.3	0.4	0.3	0	B4	0.3	0	0.3	0.4
A5	0.25	0.5	0.25	0	B5	0.25	0	0.25	0.5
A6	0.2	0.6	0.2	0	B6	0.2	0	0.2	0.6
C1/D1	0.4	0.1	0.4	0.1	E1	0.3	0.4	0	0.3
C2/ D2	0.3	0.2	0.3	0.2	E2	0.2	0.5	0	0.3
C3/ D3	0.2	0.3	0.2	0.3	Note: The difference of group C and D is that all link volumes are considered in C, but partial link (1-6).flows are taken in D.				
C4 /D4	0.1	0.4	0.1	0.4					

Table5. OD matrix estimation results in Group A, B and E

Group	A1	A2	A3	A4	A5	A6	E1
O-D $\sum \Delta D_i$	4040	3042	3079	3440	4055	4509	3157
error(%)	5.05	3.80	3.85	4.30	5.07	5.64	3.95
Group	B1	B2	B3	B4	B5	B6	E2
O-D $\sum \Delta D_i$	4040	3345	3604	4047	4094	4188	2731
error(%)	5.05	4.18	4.51	5.06	5.12	5.24	3.41

Table6. OD matrix estimation results in Group C and D

Group	C1	C2	C3	C4	D1	D2	D3	D4
O-D $\sum \Delta D_i$	1897	2039	2735	3175	2229	3298	3175	3041
error(%)	2.37	2.55	3.42	3.97	2.79	4.12	3.97	3.80

From Table 5, we can see that the accuracy of OD matrix estimation, compared with that of the traditional model, will be improved firstly and then be worsened as the weight of entry-exit flows grows up. Meanwhile, the variation trend of the estimation accuracy influenced by travel time in Group B on entry-exit is similar to that operated by entry-exit flows. The error variation trends illustrate that the sensitiveness of flow and travel time, which exerts a positive effect upon estimation accuracy, decreases gradually. Finally, the increasing error will overweigh the improvement because the weights of prior matrix and link flow are reduced. Furthermore, the estimation results under A2-A4, B2 and B3 are superior to those estimated by traditional model. This means that the additional data items are probably beneficial to enhance estimation accuracy.

From Table 6, it can be found that the estimation results under all weight series in Group C are better than those calculated by traditional model and the error under C1 decreases about 2.7%. In Group C and D, it is clear that adding two tolling data to estimation is more effective to enhance accuracy than adding one of the tolling data. By analyzing the results in Group C and D, we can conclude that, as the weights of two additional data items increase, the estimation accuracy of all link flows is close to that with partial link volumes. The calculation result of D3 and D4 are approximately same, and it means that new data items have the same effect to improve estimation result as that yielded by all link flows in C4. In other words, compared with C3, the tolling data compensate the lack of link flows (link: 8-28). At the same time, the estimation results under Group D are superior to those generated by traditional model. This it to say, in this network, it is applicable to compute OD matrix by combining partial link flows with traffic data on tolling entry-exits.

However, we should realize that the result in E2 is better than in E1, it may imply that the flow on entry-exit is more sensitive than prior OD constraint to ensure estimation accuracy. Meanwhile, the

accuracies under both E1 and E2 are better than that under A1; i.e., tolling data are more useful in improving estimate precision than other common links' in this network because more vehicles operate on charged links. Here, the tolling data are almost cost-free traffic information, so that the high cost on surveyed link flows might be saved. It also should be stated that, in practice, the observed traffic data may be largely biased, except the tolling data. If there are high observation errors on surveying common link flows and prior OD matrix in practical scenario, the difference of estimation accuracies between traditional model and new proposed model may be more remarkable. Hence, the unbiased and cost-free tolling data may have the advantage to pursue high accuracy in OD estimation. Nevertheless, for different road structures, the new data items may play different roles in estimation.

Conclusion

Prior OD matrix is combined with link flows to estimate OD matrix in existing methods. In highway network and urban network, volume on charged roads and related travel time information can be used to enhance the OD estimation precision. The expenditure on investigating link flows could be saved and the problem of surveying location choice could be avoided. Therefore, we proposed a bi-level program, whose upper objective function takes account of tolling data. The numerical example indicates that the additional data items are effective on improving estimation accuracy and new model outperforms traditional model in terms of estimation accuracy. In conclusion, it is a feasible approach to combine partial common link flows with tolling data, and the improved model can be applied to save surveying cost of fundamental traffic data and to improve the estimation accuracy.

Acknowledgement

The research described in this paper was substantially supported by three research grants from the National Natural Science Foundation of China (Projects 70871092) and a project sponsored by Shanghai Rising-Star Program (09QA1406200).

Reference

[1]H.P.LO, N.ZHANG, W.H.K.LAM. Estimation of an Origin-Destination Matrix with Random Link Choice Proportions: A Statistical Approach. Transportation Research Part B. 30(1996):309-324.

[2]Wang Wei, Sun Jun. Research on Method of Estimating O-D Matrices for Large-Sized Transportation Network[J]. Journal of Southeast University. 1996,26(6A):49-54.

[3]Martin L. Hazelton. Estimation of origin-destination matrices from link flows on uncongested networks. Transportation Research Part B 34 (2000) 549-566.

[4]Javier Doblas, Francisco G. Benitez. An Approach to Estimating and Updating Origin-Destination Matrices Based upon Traffic Counts Preserving the Prior Structure of A Survey Matrix[J].Transportation Research Part B, 2004, 39(2005): 565-591.

[5]Fisk CS. Trip matrix estimation from link traffic counts: the congested network case[J].Transportation Research,1989,23(B): 331-336.

[6]Hai Yang. Heuristic algorithms for the bilevel origin-destination matrix estimation problem. Transportation research, 1995(29B): 231-242.

[7]Jan T. Lundgren, Anders Peterson. A heuristic for the bilevel origin–destination-matrix estimation problem. Transportation Research Part B 42 (2008) 339–354.

[8]Michael J. Maher, Xiaoyan Zhang, Dirck Van Vliet. A bi-level programming approach for trip matrix estimation and traffic control problems with stochastic user equilibrium link flows. Transportation Research Part B 35 (2001) 23-40.

[9]Chang Yun-tao, Peng Guo-xiong, Yang Xiao-guang. Freeway OD matrices estimation with on/off ramps traffic counts[J].Journal of Traffic and Transportation Engineering,2003, 3(4):89-94.

[10]Sheffi Y. Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. Prentice-Hall, Englewood Cliffs, NJ, 1985:116-122.

[11]Zhang Xiaoning. Theories of Congestion Pricing in Transportation Networks[M].Hefei: Heifei University of Technology Press, 2009.

doi:10.4028/www.scientific.net/AMM.50-51.245

A important property on planar graph

Nana Li[1,a], Shujun Wang[1,b] and Xianrui Meng [2,c]

[1] Tangshan College,Tangshan, P.R. China

[2] College of Science, Hebei united University, Tangshan, P.R. China

[a]lalazinana@126.com, [b]shujunwang88@hotmail.com, [c]xianruimeng@yahoo.com.cn

Keywords: planar graph; $(L,d)^*-colouring$; $(k,d)^*-choosable$.

Abstract. In 1994, C. Thomassen proved that every planar graph is $5-choosable$ (namely $(5,0)^*-choosable$). In 1993, M. Voigt shown that there are planar graphs which are not $(4,0)^*-choosable$. But no one know whether every planar graph is $(4,1)^*-choosable$. In this paper, we give a important property on planar graph that "Every planar graph is $(4,1)^*-choosable$" and "Every planar graph is free $(4,1)^*-choosable$" are equivalent.

Introduction

List colourings of graphs are generalizations of usual colourings that were introduced by V.G. Vizing [1] and independently by P. Erdos [2] .

A vertex-colouring, or just colouring, of a finite simple graph G is an assignment of a colour to each vertext of G. A colouring is proper if adjacent vertices always get different colours. A graph is k-colourable if it has a proper colouring using at most k different colours. A list-assignment L to (the vertices of) G is the assignment of a list (set) $L(v)$ of colours to every vertex v of G; and a k-list-assignment is a list-assignment such that |L(v)|=k for every vertex v of G. If L is a list-assignment of G, then an $L-colouring$ of G is a proper colouring in which each vertex receives a colour from its own list. The graph G is k -list-colourable or k-choosable if there exists an $L-colouring$ for every k-list-assignment L of G. The chromatic number $\chi(G)$ of G is the smallest number k such that G is k-colourable. The list chromatic number, or choice number, or choosability ch(G) of G is the smallest number k such that G is $k-choosable$. C. Thomassen [3] has proved that ``Every planar graph is $5-choosable$" in 1994. The first example of a $non-4-choosable$ planar graph was given by M. Voigt [4], and further examples were given by S. Gutner [5] and M. Mirzakhani [6].

In a vertex-coloured graph, the defect $def(v)$ of a vertex v is the number of vertices adjacent to v that have the same colour as v; so a colouring is proper if and only if every vertex has defect 0. A graph G is $(k,d)^*-colourable$ if its vertices can be coloured with k colours in such a way that no vertex has defect greater than d. If L is a list-assignment of G, then an $(L,d)^*-colouring$ is an L-colouring in which no vertex has defect greater than d, and L is $(L,d)^*-colourable$ if it has an $(L,d)^*-colouring$. Finally, G is $(k,d)^*-choosable$ if it is $(L,d)^*-colourable$

whenever L is a k-list-assignment. Obviously, $(k,0)^*-colourable$ means the same as (properly) $k-colourable$,and $(k,0)^*-choosable$ means the same as $k-choosable$. R.Skrekovski [7] proved that every planar graph is $(3,2)^*-choosable$, and N. Eaton and T. Hull [8] gave a simpler version of this proof.

The free list colouring was first raised by M. Voigt [9] in 1996. Let L be a list assignment of G, G is called free $(L,d)^*-colourable$ if for every vertex $v\in V(G)$ and for every colour $f\in L(v)$ there exists an $(L,d)^*-colouring$ $\varphi_{v,f}$ with $\varphi_{v,f}(v)=f$. And G is called $free\,(k,d)^*-choosable$, if it is $free\,(L,d)^*-colourable$ whenever L is a k-list assignment. When d=0, we call it $free\,k-choosable$. M. Voigt [9] has proved that the equivalence of "Every planar graph is $5-choosable$"with "Every planar graph is $free\,5-choosable$". In this paper, we shall prove that "Every planar graph is $(4,1)^*-choosable$ " and "Every planar graph is free $(4,1)^*-choosable$" are equivalent. In order to prove our main result, we first introduce several useful lemmas and theorems in the next section.

Some lemmas and theorems

Lemma 2.1: (1) Every free $(k,d)^*-choosable$ graph is $(k,d)^*-choosable$;

(2) There are $(k,d)^*-choosable$ graphs which are not free $(k,d)^*-choosable$.

Proof(1)$\Rightarrow$(2) It is trivial.

(2) $\Rightarrow$ (1) Consider the complete bipartite graph $K_{2,2}$ and $L(x_1)=\{1,2\}$, $L(x_2)=\{1,3\}$, $L(y_1)=\{2,3\}$, $L(y_2)=\{1,3\}$, where $\{x_1,x_2\}$ and $\{y_1,y_2\}$ are two partite classes of $K_{2,2}$. If we assign color 3 to x_2, then it can not be continued to a $(L,0)^*-$list colouring of the whole graph.

We have seen that the property "to be $free\,(k,d)^*-choosable$" is a stricter requirement for a graph than the property "to be $(k,d)^*-choosable$".

Now assume the graph G is $(k,d)^*-choosable$ but not $free\,(k,d)^*-choosable$, this means there exists a k-list assignment L of G, a vertex $v^*\in V(G)$ and a colour $f\in L(v^*)$ such that $\varphi(v^*)\in L(v^*)\backslash\{f\}$ for all L-list colouring of G

In the following, such a vertex v^* is called a bad vertex and such a colour f a bad colour.

Lemma 2.2: Let G be a graph which is $(k,d)^*-choosable$ but not $free\,(k,d)^*-choosable$, v^* a bad vertex of G and $F:=\{f_1,f_2,\cdots,f_{k-1}\}$ an arbitrary set of $k-1$ colours. There exists a k-list assignment $L_{v^*,F}$ of G, so that $\varphi(v^*)\in F$ is satisfied for every $(L_{v^*,F},d)^*-$list colouring φ.

Proof. Because G be a graph which is $(k,d)^*-choosable$ but not $free\,(k,d)^*-choosable$,we can find a k-list assignment L to G, such that G is $(L,d)^*-colourable$ but not $free\,(L,d)^*-colourable$. Let v^* be a bad vertex and $L(v^*)=\{g_1,g_2,\cdots,g_{k-1},g_k\}$ and g_k is a bad colour. We use the known list assignment L with $\varphi(v^*)\in L(v^*)\backslash\{g_k\}$ (for all $(L,d)^*-$ list

colouring φ of G) and rename the colours in a suitable way: Denote colour set $T_1=\bigcup_{v\in V(G)}L(v)=\{g_1,g_2,\cdots,g_{k-1},g_k,g_{k+1},\cdots,g_n\}$, and let colour set $T_2=\{f_1,f_2,\cdots,f_{k-1},f_k,\cdots,f_n\}$, where $\{f_k,\cdots,f_n\}\cap F=\Phi$. Define an injection $\psi:\psi(g_i)=f_i,i=1,2,\cdots,n$, and let the resulting list-assignment be L^*.Because for every $(L,d)^*-colouring\ \varphi$, $\varphi(v^*)\in L(v^*)\backslash\{g_k\}$, it is easy to see that for every $(L^*,d)^*-$colouring $\varphi^*,\varphi^*(v^*)\in L^*(v^*)\backslash\{f_k\}=F=\{f_1,f_2,\cdots,f_{k-1}\}$.

Lemma 2.3 ([10]): There are planar graphs which are not $(3,1)^*-choosable$.

Main result

Theorem 3.1: The following results are equivalent:

(1) Every planar graph is $(4,1)^*-choosable$.

(2) Every planar graph is free $(4,1)^*-choosable$.

Proof. (2) $\Rightarrow$ (1):Trivial.

(1) $\Rightarrow$(2): Assume a planar graph G' is $(4,1)^*-choosable$, but G' is not free $(4,1)^*-choosable$. In the following, using G' we will construct a planar graph G^* which is not $(4,1)^*-choosable$. This contradicts to (1).

Let $G^{(3,1)}$ be a planar graph which is not $(3,1)^*-choosable$ with an m-element vertex set $V(G^{(3,1)})=\{v_1,v_2,\cdots,v_m\}$ for some m by Lemma 2.3. Let $L^{(3,1)}$ be a 3-list assignment of $G^{(3,1)}$ and $G^{(3,1)}$ is not $(L^{(3,1)},1)^*-$ list colourable.

Choose a bad vertex v^* of G' and let G' be embedded in the plane in such a way that v^* belongs to the boundary of the exterior face.

Take m copies $\{G_1',G_2',\cdots,G_m'\}$ of this graph G' with the bad vertices $v_1^*,v_2^*,\cdots,v_m^*$ respectively.

Define $G^*(V^*,E^*)$, where $V^*:=\bigcup_{i=1}^m V(G_i')$, $E^*:=\bigcup_{i=1}^m E(G_i')\cup\{(v_i^*,v_j^*)|(v_i,v_j)\in E(G^{(3,1)})\}$

Construct a 4-list assignment L^* of G^* in the following way: for $v_j^*(j=1,2,\cdots,m)$,take $L^*(v_j^*)=L^{(3,1)}(v_j)\cup\{f_k\}$, where f_k is bad colour of bad vertex v^* in G'. And for each one of the other vertices in G^*, take the same colour list as its in G'.

Since G^* is planar, then G^* is $(4,1)^*-choosable$ by (1). And hence there exists an $(L^*,1)^*-$ list colouring φ^* of G^* with $\varphi^*(v_j^*)\in L^{(3,1)}(v_j)$ for all $j=1,2,\cdots,m$, by Lemma 2.2.

Therefore $G^{(3,1)}$) is $(L^{(3,1)},1)^*-$ list colourable, which contradicts to the assumption..

Remark 3.1 Note that there are planar graphs that are not $(2,2)^*-choosable$ [2].Using the same method, we can prove the equivalence every planar graph is $(3,2)^*-choosable$ [11]" with “every planar graph is free $(3,2)^*-choosable$ „.

References

[1] V.G. Vizing: Coloring the vertices of a graph in prescribed colors(in Russian). Diskret. Anal. Vol. 29(1976).

[2] P. Erdos, A.L. Rubin and H. Taylor:Choosability in graphs, Congr. Numer. Vol.26(1979).

[3] C. Thomassen:Every planar graph is 5-choosable, J.Combin.Theory Ser. B Vol.62(1994).

[4] M. Voigt:List colourings of planar graphs, Discrete Math. Vol.120(1993).

[5] S. Gutner: The complexity of planar graph choosability, Discrete Math. Vol.159(1996).

[6] M. Mirzakhani: A small non-4-choosable planar graph, Bull Inst. Combin. Appl. Vol.17(1996).

[7] R. · Skrekovski:List improper colourings of planar graphs, Combin. Prob.Comput. Vol.8(1999).

[8] N. Eaton and T. Hull:Defective list colorings of planar graphs, Bull Inst. Combin.Appl. Vol.25(1999).

[9] M. Voigt:Choosability of planar graphs, Discrete M ath. Vol.150(1996).

[10] D.R. Woodall:Defective choosability results for outerplanar and related graphs,Discrete Math. Vol.258(2002).

[11] D.R. Woodall:List colourings of graphs, Survey in combinatorics, 2001, London Math. Soc. Lecture Note Series 288, Cambridge University Press(2001).

Project supported by Natural Science Foundation of Hebei Province of China (A2010000910) and Science Research Foundation of Hebei Polytechnic University (No. z200916).

doi:10.4028/www.scientific.net/AMM.50-51.249

Vacuum Casting Automatic Control System Based on Configuration Software

Jie Huang[1, a], Yuanyuan Liu[1,b], Haiguang Zhang[1], Zhaofu Chi[1], Qingxi Hu[1]

[1]Rapid Manufacturing Engineering Centre, Shanghai University, Shanghai 200444, China
[a]jerrry521@shu.edu.cn, [b]yuanyuan_liu@shu.edu.cn

Keywords: Vacuum casting; Configuration software; DDE; MATLAB

Abstract. Bubbles and craft parameters in vacuum casting are the key factors to the quality of products. How to eliminate bubbles and set optimal parameters are the key technical problems to the automation development of vacuum casting. This paper advances a new vacuum casting automatic control system based on configuration software. Dynamic data exchange（DDE） among control module, detection module and execution module is realized based on the strong capability of communication and human-machine interface in configuration software. The realization of this control system can effectively eliminate bubbles and provide corresponding optimal craft parameters automatically and intelligently. In a word, the efficiency and quality of the products can be improved and the dependency of experience parameters can be reduced in vacuum casting.

Introduction

Vacuum casting, a kind of rapid tooling manufacturing technology, is characterized by the use of vacuum technique during the processes of mould fabrication and the casting of parts [1]. Because of its capability to produce products in a short time and low price, especially its unique advantage in the thin-wall and transparent products, vacuum casting technology is widely used and studied in product design verification and small batch production.

Bubbles in reaction and craft parameters such as pressure, velocity and casting time in vacuum casting are the main factors that affect the quality of products. Manual control, electrical control and PLC control are the control methods in current vacuum casting equipments. At present, most vacuum casting equipments employ the artificial method to eliminate bubbles, which is both time-consuming and labor-consuming with a low efficiency. Besides, the optimal craft parameters are not available in the manufacturing of the machines. With the development of the CIMS, advanced and complex control system has been applied to the RP (rapid prototyping) equipments, which effectively improves the productivity and quality of the products [6,7]. But the advanced control system has not been applied in vacuum casting equipments yet. In this paper, a vacuum casting automatic control system based on configuration software is proposed. Vacuum casting bubble automatic elimination system based on on-line machine vision detection [2] is included in this system. MATLAB is used to design fuzzy algorithm to recommend craft parameters.

The structure of the automatic control system

As is show in Fig.1, two-level distributed structure based on PLC and PC is adopted in the control system, which comprises three main modules: the control module, the detection module and the execution module. The control module contains main-controller PC and sub-controller PLC. The detection module is used to inquire the signals of pressure, bubble, position and velocity of the material cups. And the execution module includes vacuum pump, motors, and electromagnetic

valves. They are all controlled by PLC, which mainly complete the sequential control in the whole process of vacuum casting, such as degassing, stirring, air-in and pouring.

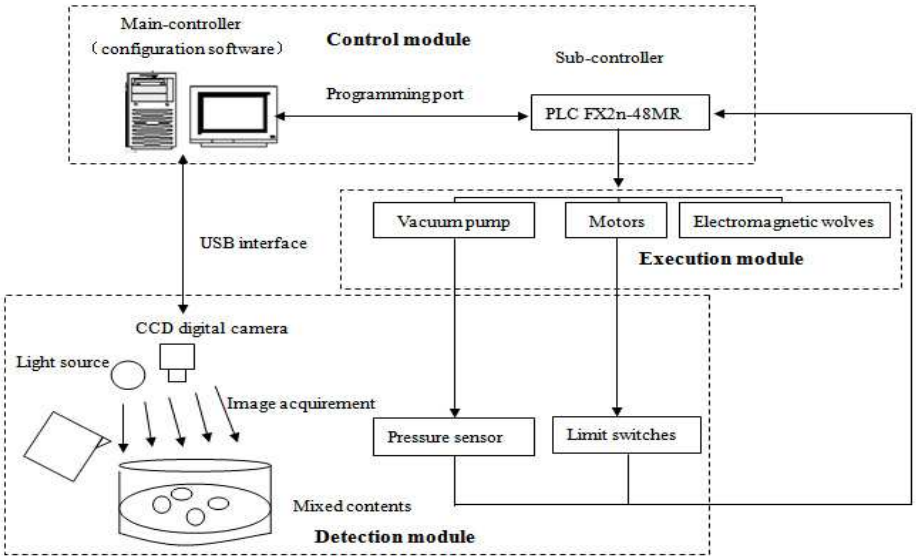


Fig.1 The structure of the control system

The core of the whole system is control module. In the industrial control area, there are two major ways in PC control system: One is to use programming languages such as VB, VC++ to design procedures to reach the requirement of the system and the other one is to use industrial configuration software for secondary development [3]. The advantages of configuration software include its short design period, excellent I/O interface and easy maintenance. In order to realize full automatic control in the whole vacuum casting process, data exchange and feedback between modules in real time must be implemented. So configuration software is chosen in this system for the integrated communication protocols, short design period and good human-machine interface.

In the system, Pressure sensor, CCD visual sensor, MATLAB, ACCESS database and PLC are connected to configuration software and data exchange is done via communication protocols. ForceControl configuration software is integrated with Mitsubishi programming port protocol, MODBUS protocol, DDE protocol and SQL function that is based on ODBC database. In vacuum casting control system, force control 6.1 SP is selected, the communication style is shown in Fig.2.

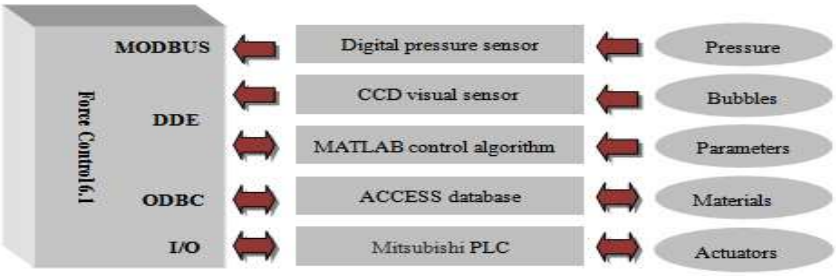


Fig.2 The data exchange based on force control 6.

DDE based on ForceControl configuration software

The access to material database. The following functions are achieved by ACCESS database through configuration software: Build vacuum casting material database by ACCESS and save the information of the materials including physical property, chemical property and craft information. The database is convenient for users to search and choose the appropriate casting materials according to corresponding products. Besides, the information can be added, deleted and modified by the defined users.

ForceControl6.1 provides SQL functions in scripts to operate database based on ODBC. By SQL functions, the database can be inserted, added, deleted and modified [4].

The specific method to achieve access to the ACCESS database by SQL functions is as follows:

(1) To create a SQL model table on configuration interface and define each field name of the table.

(2) To create a binding table corresponding to ACCESS database table and link the field name and middle variable. SQLCreateTable (ConnectID, “vacuum casting”, “model”);

(3) To create windows table “table” by windows control and link ODBC source data; ACCESS database and windows table can be linked through the source data.

(4) The corresponding operation scripts:

Link the table: SQLConnect (ConnectID, DataSourceDesc) ;

Select, display table: SQLSelect (ConnectID, “vacuum casting”, “”); SQLDisplayToGrid (“table”);

Delete the data: SQLDelete (ConnectID, “vacuum casting”, “condition”);

Insert data: SQLInsert (ConnectID, “ vacuum casting”, “model”);

Modify data: SQLUpdate (ConnectID, “vacuum casting”, “field name”, “condition");

Search data: SQLSelect (ConnectID,”vacuum casting “, “field name =’“+desc+”’”);

Communication between configuration and MATLAB. At peasant, industrial control configuration software is with excellent encapsulation, I/O communication and human-machine interface, but it is hard to apply complex control algorithms. On the contrary, MATLAB possesses strong analysis computing capability and complex control algorithm modules. Besides, MATLAB can provide interface with other software, which makes it easy to exchange data. Therefore, it is better to use configuration software as the main controller to acquire and display data, while using MATLAB as a background program to do complex algorithms, such as fuzzy control, parameter tuning, signal processing [5].

In the automatic control system of vacuum casting, the optimal craft parameters should be provided to the craft process according to the material information and product properties searched by the operator. We can get the information and properties from ACCESS database. These information variables can be viewed as the inputs for the fuzzy algorithm and sent to MATLAB via DDE protocol. In the program of MATLAB, fuzzy algorithm is designed with these variables. Once the program is executed, the output variables which are optimized by the algorithm are shown on the configuration. At the same time, the optimized variables are linked with the parameters of PLC and execution module, which are sent to execution module to operate the vacuum casting process.

The input variables are curing time, temperature, style and property of materials from ACCESS database. The output variables are casting and stirring velocity as well as pressure.

The steps of communication between configuration and MATLAB include:

(1) To create a DDE project in I/O module and set the service, topic name in configuration software.

(2) To set the input and output variables such as a1.pv, a2.pv.

(3) To create an M file in the MATLAB software using fuzzy algorithm.

(4) To translate the M file to executable program and set a button to link the program.

The main program of M files in MATLAB is as follows:

```
Global channel      (declare global variable channel)
Channel=ddeinit ('DB','DB');     (initialize communication, the two DB are service and topic name )
Ddeadv (channel,'ctime.pv','disp(x)','x');      (create input and output channel connection)
A0=ddereq(channel,'ctime.pv');      (connect the inputs and variables in matalb)
ddepoke(channel,'T0.pv',B0);         (B0 is the result of Matlab,T0.pv is the output variable)
```

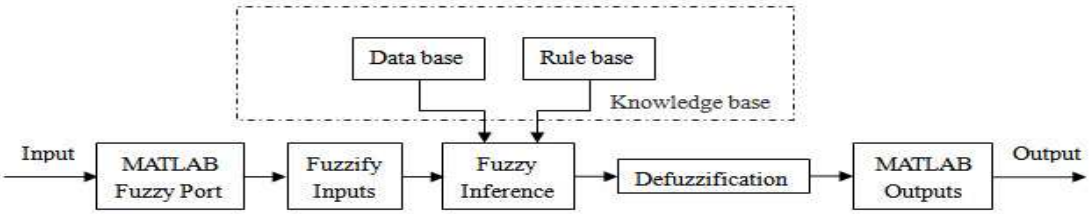


Fig.3 The structure of MATLAB fuzzy controller

Communication between configuration and PLC. Mitsubishi FX serial programming port protocol is integrated in the ForceControl 6.1, the communication can be achieved through the programming port without any extension card. The sub-controller easily gets data from main-controller through the digital variables and analog variables.

Communication between configuration and bubble automatic elimination system. In vacuum casting bubble automatic elimination system based on on-line machine vision detection, Image processing procedure is designed via VC++ and connected to configuration by DDE protocol supported by both VC++ and configuration software. The whole process of image processing is as follows: Image acquisitions-> Image processing-> Analysis and judgment-> Sending signal to PLC [2]. The final result is sent to PLC, which decides PLC whether to execute bubbles elimination.

Communication between configuration and pressure sensor. In vacuum casting process, two digital pressure sensors are needed to detect the pressure of the two chambers. ACS-P digital pressure sensor of ANCN Company is chosen to get the pressure value in real time, while MODBUS communication protocol is applied to display pressure value on configuration interface. MODBUS protocol defines a message structure that can be identified by controller and establishes a public format of message feedback. Through this protocol, the pressure values are easy to collect and display, which are sent to PLC to control the execution module. Through this protocol, the pressure values are easy to collect and display, which are sent to PLC to control the execution module.

Conclusions

The process to use the automatic control system is shown on Fig.4 and Fig.5 is the simulation interface.

Bubbles in reaction and craft parameters are the main factors that affect the final quality of vacuum casting products. To improve the degree of automation and apply automatic and intelligent control system is an effective way to improve the quality and efficiency and reduce the dependency of experience parameters in vacuum casting. An automatic control system of vacuum casting based on configuration software is designed in this paper. Dynamic data exchange among control module, detection module and execution module is realized based on the strong capability of communication and human-machine interface in configuration software. Once the feedback is realized among these modules, the vacuum casting equipment can be controlled in an automatic and intelligent way.

DDE protocol is adopted to exchange data between configuration interface and bubble automatic elimination system based on on-line machine vision detection. Bubbles can be detected by CCD visual sensor and eliminated by the bubble automatic elimination system, while the feedback is sent to PLC to control the whole craft. Besides, material information can be obtained from ACCESS database and sent to MATLAB, and the feedback parameters optimized from MATLAB by the fuzzy algorithm are sent to control the chamber pressure and motor speed.

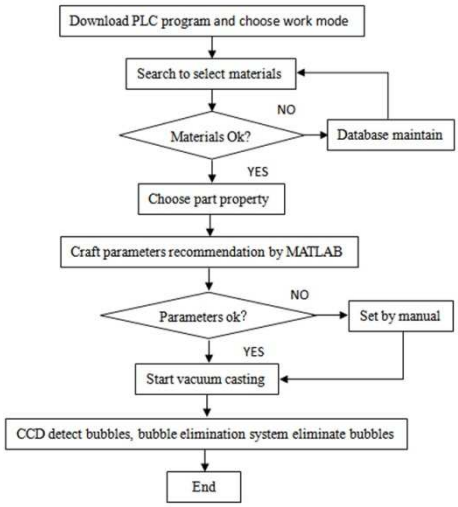


Fig.4 The process to use the system

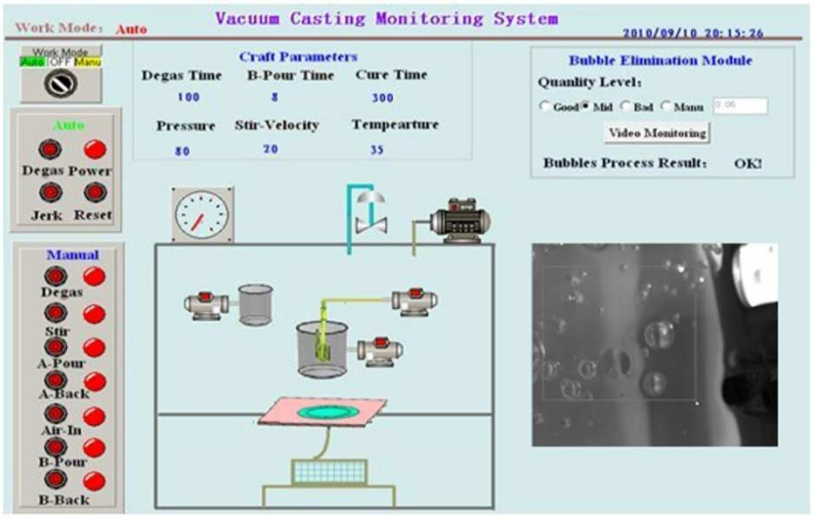


Fig.5 The simulation of the system

Acknowledgements:

The authors would like to thank the supports by Innovation Foundation of SHU (A.10-0109-08-010), Shanghai Science Foundation for the Excellent Youth Scholars (B.37-0109-08-004), and the fund of Shanghai Key Laboratory of Mechanical Automation and Robotics (D.10-0109-09-006).

References

[1] Y.Tang, W.K.Tan, *et al.*: J.Mater.Process. Technol. Vol.192-193 (2007), p. 334.
[2] Y.Y.Liu, Z.F.Chi, *et al.*: Key Engineering Materials Vol.426-427(2010), p260.
[3] H.O.Zhang, H.Ai, G.L.Wang: Machinery & Electrolics.Vol.28 (2004), p59.
[4] In: *ForceControl 6.1 Reference manual,* edited by Limited company of Beijing ForceCotrol BJ (2007), in press.
[5] Y.shi,Y.Hong and X.L.Qian:Chinese Journal of Scientific Instrument,Vol.24(2003),p337.
[6] A. Kirschbaum and M. Glesner: In 8th IEEE Int. Workshop on Rapid System Prototyping, June 24-26(1997) p. 136.
[7] C.R. Luo *et al.*:IEEE International Conference on Robotics and Automation, 1999, p. 2203
[8] J.Huang, Q.X.Hu *et al.*: IEEE International conference on MVHI 2010, 2010, p.646

doi:10.4028/www.scientific.net/AMM.50-51.254

A hyperchaotic system and its synchronization via scalar controller

Wuchun Dai [1,a], Zhengfu Cheng[1]

[1] Physics Department of Chongqing University of Arts and Sciences, Chongqing, 402160, China

[a]z3889@163.com

Keywords: hyperchaotic system, Lyapunov exponents, Poincar´e mapping, synchronization.

Abstract. In this paper, a 4D hyperchaotic system is proposed. Some basic dynamical behaviors are explored by calculating its Lyapunov exponents, Poincar´e mapping, etc.. Finally, synchronization for this new hyperchaotic system is achieved via scalar control. The nonlinear terms in the response system are not dropped. The proposed synchronization scheme is simple and theoretically rigorous. The mathematical proof of this method is provided. Some numerical simulations are obtained. The numerical simulations coincide with the theoretical analysis.

1. Introduction

There has been increasing interest in exploiting chaotic dynamics in engineering applications, such as electrical engineering, telecommunications, information processing, material engineering, etc.. This is due to the desirable features of chaos in some engineering and technological applications, where chaos can provide certain unique features for special needs. In the past decades, many novels and simple nonlinear circuits exhibiting chaotic dynamics have proposed [1,2,3,4,5,6,7,8,9,10,11,12].

However, Perez and Cerderia have proved that the messages masked by a simple chaotic system with only a single positive Lyapunov exponent, can be easily extracted sometimes when they were intercepted [13]. It is suggested that this problem can be overcome by using the higher-dimensional hyperchaotic systems. Because of hyperchaos systems has more than one positive Lyapunov exponent, and has more complex dynamic behavior. Therefore, it is very important to construct hyperchaotic system.

This paper reports a new 4D hyperchaos system, which is contained only one nonlinear term. We yield the Lyapunov exponents, Poincar´e mapping, and phase diagram of the 4D hyperchaotic system. Furthermore, a synchronization scheme via scalar control is proposed. The mathematical proof of this method is provided. Some numerical simulations are obtained. The numerical simulations coincide with the theoretical analysis.

2. A 4D novel hyperchaotic system with only one nonlinear term

Consider the following new 4D dynamical system

$$\begin{cases} \dot{x} = 0.56x - y \\ \dot{y} = x - 0.1yz^2 \\ \dot{z} = 4y - z - 6w \\ \dot{w} = 0.5z + 0.8w \end{cases}, \tag{1}$$

where x, y, z and w are state variables. The new 4D system (1) is symmetric with respect to the origin, which is invariant for the coordinate transformation $(x, y, z, w) \rightarrow (-x,- y, -z, -w)$.

System (1) has three real equilibrium, which are denoted by $S_0=[0,0,0,0]^{\boldsymbol{T}}$, $S_+=[-5.5p/4.48,-11p/16, p, -5p/8]^{\boldsymbol{T}}$, $S_-=[5.5p/4.48,11p/16,- p, 5p/8]^{\boldsymbol{T}}$, respectively. Where, $p = 0.056^{-1/2}$.

The characteristic equation at equilibrium S_0 is

$$\lambda^4 - 0.36\lambda^3 + 3.088\lambda^2 - 1.032\lambda + 2.2 = 0. \quad (2)$$

The four eigenvalues are 0.28-0.96i, 0.28+0.96i, -0.2-1.4142i, -0.2+1.4142i, respectively. So, equilibrium S_0 is a two dimensional unstable saddle point.

We obtain the characteristic equation at equilibrium $S_{\pm}$ as following

$$\lambda^4 + (0.1p^2 - 0.36)\lambda^3 - (0.586p^2 - 3.088)\lambda^2 + (0.9568p^2 - 1.032)\lambda + 2.2 - 0.3696p^2 = 0. \quad (3)$$

The four eigenvalues at equilibrium $S_{\pm}$ are 1.1005+1.3901i, 1.1005-1.3901i, 0.3135, -4.1402, respectively. So, equilibrium $S_{\pm}$ is a three dimensional unstable saddle point.

The Lyapunov exponents of system (1) are are $\lambda_1 = 0.074726, \lambda_2 = 0.021509, \lambda_3 = 0, \lambda_4 = -0.43386$, respectively. So system (1) is a hyperchaotic system. Fig.1 shows the hyperchaotic phase portraits.

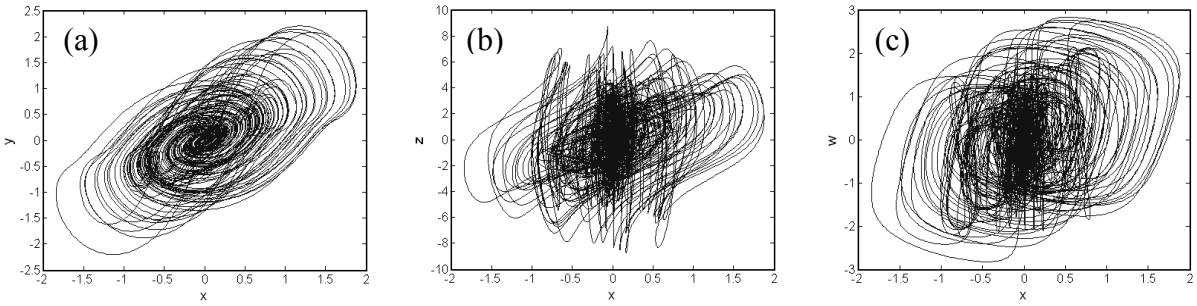


Fig. 1 Hyperchaotic phase portraits of the system (1): (a) x–y plane; (b) x– z plane; (c) x–w plane

The Poincar´e mapping of this hyperchaotic system is also analyzed. It can be seen that the Poinca´e mappings are these points in confusion as shown in Fig.2.

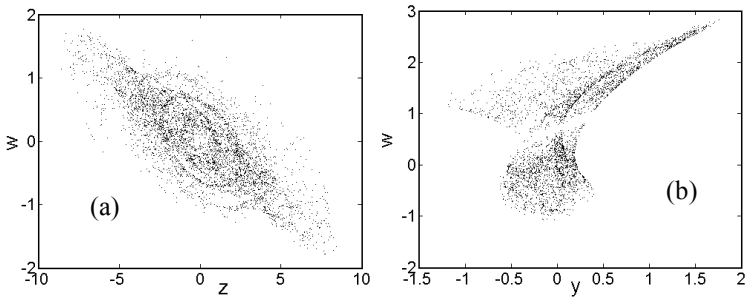


Fig.2 The Poincar´e mapping of system (1): (a) y=0, (b) z=-4.2233

3. Synchronization for the 4D hyperchaotic system (1)

In this section, we discuss the chaotic synchronization for hyperchaotic system (1). System (1) is regarded as the driving system. The scalar feedback controlled slave system is

$$\begin{cases} \dot{x}' = 0.56x' - y' \\ \dot{y}' = x' - 0.1y'z'^2 + 0.1z^2(y' - y) + 0.2yz(z' - z) + \beta(y' - y) \\ \dot{z}' = 4y' - z' - 6w' \\ \dot{w}' = 0.5z' + 0.8w' \end{cases}, \quad (4)$$

where β is a tuning parameter that controls the strength of the feedback into the response system, and to be designed. We note that when the system (4) and (1) are realized to synchronization, the driving system (1) and response system (4) essentially become uncoupled, and system (4) become uncontrolled.

Theorem : If $-\frac{1}{0.56} < \beta < -0.56$, then system (4) and (1) can be achieved to synchronization.

Proof:

Define the state error $e_1 = x' - x$, $e_2 = y' - y$, $e_3 = z' - z$, and $e_4 = w' - w$. Then, the corresponding error dynamics system can be obtained by subtracting Eq. (1) from Eq. (4)

$$\begin{pmatrix} de_1/dt \\ de_2/dt \\ de_3/dt \\ de_4/dt \end{pmatrix} = \begin{pmatrix} 0.56 & -1 & 0 & 0 \\ 1 & \beta & 0 & 0 \\ 0 & 4 & -1 & -6 \\ 0 & 0 & 0.5 & 0.8 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix} + \begin{pmatrix} 0 \\ -0.1e_2e_3^2 - 0.1ye_3^2 - 0.2e_2e_3z \\ 0 \\ 0 \end{pmatrix}, \tag{5}$$

Form error dynamics system (5), we can obtain that $e_i = 0 (i = 1,2,3,4)$ is an equilibrium point of system (5). If the zero solution $e_i = 0 (i = 1,2,3,4)$ of error dynamics system (5) is globally asymptotically stable under a suitable constant β, which implies that the two systems (1) and (4) are realized to synchronization.

Linearizing error dynamics system (5) at equilibrium $e_i = 0 (i = 1,2,3,4)$ yields the Jacobian matrix

$$J = \begin{bmatrix} 0.56 & -1 & 0 & 0 \\ 1 & \beta & 0 & 0 \\ 0 & 4 & -1 & -6 \\ 0 & 0 & 0.5 & 0.8 \end{bmatrix}. \tag{6}$$

The four eigenvalues of matrix J are $\lambda_{\pm} = 0.5(\beta + 0.56) \pm 0.5\sqrt{(\beta - 0.56)^2 - 4}$, $\lambda_{\pm}' = -0.1 \pm i\sqrt{2.19}$, respectively.

Because $-\frac{1}{0.56} < \beta < -0.56$, so the following two cases are satisfied:

Case 1 If $-\frac{1}{0.56} < \beta \le -2 + 0.56$, then $(\beta - 0.56)^2 - 4 \ge 0$ and $0.56\beta \times 4 > -4$. So, $0.56\beta \times 4 > -4 \Rightarrow 0.56\beta \times 2 > -4 - 0.56\beta \times 2 \Rightarrow \beta^2 + 0.56^2 + 0.56\beta \times 2 > -4 - 0.56\beta \times 2 + \beta^2 + 0.56^2 \Rightarrow (\beta + 0.56)^2 > (\beta - 0.56)^2 - 4$.

Therefore, $\beta + 0.56 < 0$, $(\beta - 0.56)^2 - 4 \ge 0$ and $(\beta + 0.56)^2 > (\beta - 0.56)^2 - 4$. So, the eigenvalues $\lambda_{\pm}$ of matrix J is negative real number, and the eigenvalues $\lambda_{\pm}'$ of matrix J have negative real part.

Case 2 If $-2 + 0.56 < \beta < -0.56$, then $\beta + 0.56 < 0$ and $(\beta - 0.56)^2 - 4 < 0$. So, the four eigenvalues of matrix J have negative real part.

Therefore, the equilibrium point $e_i = 0 (i = 1,2,3,4)$ of error dynamics system (5) is globally asymptotically stable, which indicates that the system (4) and (1) are realized to synchronization. This finishes the proof.

The above results manifest the system (4) and (1) which are synchronized under suitable constant β. Some simulation results are described below. Now, we choose $\beta = -0.76$ and $\beta = -1.56$ for numerical simulations, respectively. The initial states of the drive system (1) and the response system (5) are taken as $x(0)$=-0.1, $y(0)$=0, $z(0)$=0.2, $w(0)$=-0.4 and $x'(0)$=0.3, $y'(0)$=0.5, $z'(0)$=-0.05, $w'(0)$=0.7, respectively. The simulation results can be depicted in Fig. 3 and Fig. 4, respectively. Where $\varepsilon(t) = (\sum_{i=1}^{4} e_i^2)^{1/2}$. According to the simulation results, we obtain that numerical simulations coincide with the theoretical analysis.

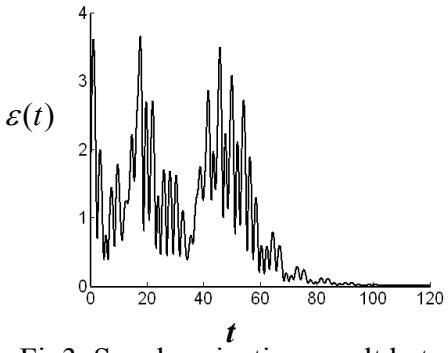


Fig3. Synchronization result between system (1) and (5)

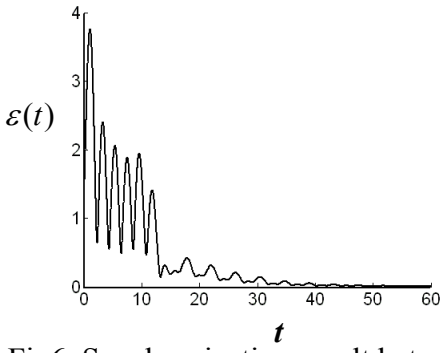


Fig6. Synchronization result between system (1) and (5)

4. Conclusion

A new 4D hyperchaos system, which is contained only one nonlinear term, is constructed. Its dynamical behaviors such as the Lyapunov exponents, Poincar´e mapping, and phase diagram of the 4D hyperchaotic system are obtained. Furthermore, a synchronization scheme via scalar control is proposed. The numerical simulations coincide with the theoretical analysis.

References

[1] D. Cafagna, G. Grassi: Int. J. Bifurcation Chaos Vol. 13(2003), p.2889

[2] K. Thamilmaran, M. Lakshmanan and A. Venkatesan: Int. J. Bifurcation Chaos Vol.14 (2004), p. 221

[3] Y. X. Li, W. K. S. Tang and G. R. Chen: Int. J. Circuits Theor. Appl Vol.33 (2005), p.235

[4] C. D. Li, X. F. Liao and K. W. Wong: Chaos, Solitons & Fractals Vol. 23 (2005), p.183

[5] P. Zhou, X.H. Luo, H. Y.A. Chen: Acta Physica Sinica Vol.54(2005), p.5048

[6] Y. X. Li, G. R. Chen and W. K. S. Tang: IEEE Trans. Circuits Systems II Vol.52 (2005), p.204

[7] Z. Q. Chen, Y. Yang, G. Y. Qi and Z. Z. Yuan: Physics Letters A Vol.360 (2007), p.696

[8] P. Zhou, X. F. Cheng, N.Y. Zhang: Chinese Physics Vol.17 (2008), p.3252

[9] P. Zhou, L. J. Wei, X.F. Cheng: Chinese Physics B Vol.18 (2009), p.2674

[10] P. Zhou, Y.X. Cao, X.F. Cheng: Chinese Physics B Vol.18 (2009), p.1394

[11] P. Zhou, L. J. Wei, X.F. Cheng: Acta Physica Sinica Vol.58 (2009), p.5201

[12] G. Y. Wang, X. Zhang, Y. Zheng and Y. X. Li: Physica A Vol.371 (2006), p.260

[13] G. Perez, H. A. Cerdeira: Phys. Rev. Lett. Vol. 74(1995), p.1970

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.258

Generalized projective synchronization for a class of continuous chaotic systems

Yafei Zhou [1,a], Dong Zhang[1]

[1] Physics Department of Chongqing University of Arts and Sciences, Chongqing, 402160, China

[a]z3889@163.com

Keywords: generalized projective synchronization, continuous chaotic systems, sate variable, time derivatives.

Abstract. In this paper, a generalized projective synchronization (GPS) scheme for a class of continuous chaotic systems is investigated by using only one sate variable and its time derivatives. The construction method of response system is proposed. The mathematical proof of the GPS scheme is provided. The synchronization technique is simple and theoretically rigorous. Finally, the corresponding numerical simulation results demonstrate the effectiveness of the proposed schemes.

1. Introduction

In 1990, Pecora and Carroll presented the concept of "chaotic synchronization" for the first time [1] and introduced a method to synchronize two identical chaotic systems with different initial conditions [2]. Chaos synchronization has become an active research subject in nonlinear science because of its potential applications in some engineering and technological applications [3-9]. In 1999, Mainieri et al. discovered a new method of synchronization, i.e., projective synchronization [10]. Chee et al. have applied the projective synchronization technique to chaotic secure communications [11]. In application to secure communications, projective synchronization (PS) can be used to extend binary digital to M-nary digital communication for achieving fast communication [12]. So the research on chaos synchronization has important theory significances and application value.

On the other hand, most of the methods for synchronize two chaotic systems is the state feedback controlling method. Several state variables or all the state variables of the driving system are transmitted to the response system. Therefore, it is necessary to measure several state variables or all the state variables state of driving system. However, it is apparent that it is not always possible to detect all state variables in a physical system, especially in a complicated system or a large system [13], and some state variables can not been measured in some cases. However, the variables that can be observed or measured are worthy to be investigated, and only the observable or measured variables of physical significance are usually interested in. For example, frequency in an oscillator, and temperature (pressure) in a thermodynamically system [13].

This work, inspired by the above discussion, addresses one observable state variable for a class of continuous chaotic systems. The main contributions of this paper are follows: first, by using only a specific sate variable and its time derivatives, all the state variables can be yielded; second, several well-known chaotic systems are used to demonstrate this view; finally, a generalized projective synchronization scheme is presented to show its application.

2. A class of continuous chaotic system

First, let us introduce some notions and results presented in Ref. [14]. Consider the following chaotic system with scalar output z:

$$\begin{aligned} d\mathbf{x}/dt &= \mathbf{f}(\mathbf{x}) \\ z &= h(\mathbf{x}) \end{aligned}, \tag{1}$$

where $\mathbf{x} \in R^n$, $\mathbf{f}: R^n \to R^n$. Let $k(\mathbf{x})$ be a scalar differentiable function and $\mathbf{V}(\mathbf{x}): R^n \to R^n$ be a vector field. The Lie derivative of $k(\mathbf{x})$ along $\mathbf{V}(\mathbf{x})$ is defined as $L_V k =< \mathbf{V}(\mathbf{x}), \text{grad}k(\mathbf{x}) >$, where $<.>$ is the Euclidean inner product. The *i*th Lie derivative of $k(\mathbf{x})$ along $\mathbf{V}(\mathbf{x})$ is defined as $L_V^{(i)} k =< \mathbf{V}(\mathbf{x}), \text{grad}L_V^{(i-1)} k(\mathbf{x}) >$.

Lemma 1 [14] If the map $\Phi = [h(\mathbf{x}), L_\mathbf{f} h(\mathbf{x}), L_\mathbf{f}^{(2)} h(\mathbf{x}), \cdots L_\mathbf{f}^{(n-1)} h(\mathbf{x})]^{\mathrm{T}}$ is a diffeomorphism. Then Eq.(1) is observable through scalar output z and its ith derivative for $i \le n-1$. In addition, $\mathbf{x} = \Phi^{-1}[z(t), dz(t)/dt, d^2 z(t)/dt^2, \cdots, d^{n-1} z(t)/dt^{n-1}]^{\mathrm{T}}$

According to **Lemma 1**, for a class of continuous chaotic systems, for simplicity, we can choose one of the variables, denoted by $x_i(t)$, which can be observed or measured, as output, then the other variables can be denoted by $x_j = g_j(x_i, dx_i/dt, d^2 x_i/dt^2, \cdots, d^{n-1} x_i/dt^{n-1})$, $j \ne i$ $(i, j = 1,2,...,n)$.

This indicates that all the state variables of such chaotic system will be yield via a specific sate variable (a scalar output) and its time derivatives, which can be verified by the following example.

Example Consider the unified chaotic system [15]:

$$
\begin{aligned}
dx_1/dt &= -(25\alpha + 10)(x_1 - x_2), \\
dx_2/dt &= (28 - 35\alpha)x_1 + (29\alpha - 1)x_2 - x_1 x_3, \\
dx_3/dt &= x_1 x_2 - (\alpha + 8)x_3/3,
\end{aligned}
$$

where $0 \le \alpha \le 1$. Now, choose variable $x_1(t)$ as the specific sate variable and as the scalar output (i.e., $h(x_1, x_2, x_3) = x_1$), then the other state variables can be denoted as:

$$
\begin{aligned}
x_2 &= (dx_1/dt)/(25\alpha + 10) + x_1 = g_2(x_1, dx_1/dt, d^2 x_1/dt) \\
x_3 &= [-dx_2/dt + (28 - 35\alpha)x_1 + (29\alpha - 1)x_2]/x_1 \\
&= [-d^2 x_1 dt/(25\alpha + 10) + dx_1/dt]/x_1 + \{(29\alpha - 1)[dx_1/dt/(25\alpha + 10) + x_1] + (28 - 35\alpha)x_1\}/x_1 = g_3(x_1, dx_1/dt, d^2 x_1/dt^2) \qquad (x_1 \ne 0).
\end{aligned}
$$

By a similar argument, one can see that many chaotic systems, such as Liu chaotic system [16], Rossler chaotic system, Sprott's chaotic system, a hyperchaotic system [6] also have the same property with example 1.

Without loss of generality, in the following, we rewrite chaotic system (1) as

$$
d\mathbf{x}/dt = \mathbf{A}\mathbf{x} + \mathbf{h}(x_1, dx_1/dt, ..., d^{n-1} x_1/dt^{n-1}) \tag{2}
$$

$\mathbf{A} \in R^{n \times n}$ is a constant matrix.

3. The new generalized projective synchronization (GPS) scheme

In this section, a generalized projective synchronization scheme is presented for chaotic drive system (2). Given a nonlinear map $\boldsymbol{\psi} = \boldsymbol{\psi}(\mathbf{y}): R^n \to R^n$, $\mathbf{y} \in R^n$, and its Jacobian matrix is,

$$
\mathbf{D}\boldsymbol{\psi}(\mathbf{y}) = \begin{pmatrix}
\partial \psi_1(\mathbf{y})/\partial y_1 & \partial \psi_1(\mathbf{y})/\partial y_2 & \cdots & \partial \psi_1(\mathbf{y})/\partial y_n \\
\partial \psi_2(\mathbf{y})/\partial y_1 & \partial \psi_2(\mathbf{y})/\partial y_2 & \cdots & \partial \psi_2(\mathbf{y})/\partial y_n \\
\cdots & \cdots & \cdots & \cdots \\
\partial \psi_n(\mathbf{y})/\partial y_1 & \partial \psi_n(\mathbf{y})/\partial y_2 & \cdots & \partial \psi_n(\mathbf{y})/\partial y_n
\end{pmatrix}.
$$

Assume that $\mathbf{D}\boldsymbol{\psi}(\mathbf{y})$ is an invertible matrix, and then construct the following response system (3),

$$
\dot{\mathbf{y}} = \mathbf{D}^{-1}\boldsymbol{\psi}(\mathbf{y})[\mathbf{A}\boldsymbol{\psi}(\mathbf{y}) + \mathbf{h}(x_1, dx_1/dt, ..., d^{n-1} x_1/dt^{n-1}) + \mathbf{K}\begin{pmatrix}
\psi_1(\mathbf{y}) - x_1 \\
\psi_2(\mathbf{y}) - g_2(x_1, dx_1/dt, ..., d^{n-1} x_1/dt^{n-1}) \\
\psi_3(\mathbf{y}) - g_3(x_1, dx_1/dt, ..., d^{n-1} x_1/dt^{n-1}) \\
\cdots \\
\psi_n(\mathbf{y}) - g_n(x_1, dx_1/dt, ..., d^{n-1} x_1/dt^{n-1})
\end{pmatrix}], \tag{3}
$$

where $\mathbf{k} \in R^{n \times n}$ is a real matrix, and to be decided later, $\mathbf{D}^{-1}\boldsymbol{\psi}(\mathbf{y})$ is the inverse matrix of $\mathbf{D}\boldsymbol{\psi}(\mathbf{y})$. Obviously, only sate variable x_1 of drive system (2) and its time derivatives are transmitted to the response system (3). The drive system (2) and response system (3) are said to be in generalized projective synchronization (GPS) if $\lim_{t \to \infty} \|\boldsymbol{\psi}(\mathbf{y}) - \mathbf{x}\| = 0$.

Theorem 1: Given drive system (2) and response system (3), there exists a feedback gain matrix $\mathbf{k}$ such that a generalized projective synchronization (GPS) between (2) and (3) can be achieved if all the eigenvalues of $\mathbf{A}+\mathbf{k}$ satisfy $\mathrm{Re}(\lambda_i)<0$ ($i=1,2,\cdots,n$).

Proof:

Denote $\mathbf{e}=\boldsymbol{\psi}(\mathbf{y})-\mathbf{x}$ as the GPS error between drive system (2) and response system (3), then the dynamic equation of GPS error can be obtained easily from Eq. (2) and (3), which is expressed as $\dot{\mathbf{e}}=\mathbf{D}\boldsymbol{\psi}(\mathbf{y})\dot{\mathbf{y}}-\dot{\mathbf{x}}$. Then, it follows that

$$\dot{\mathbf{e}}=\mathbf{A}(\boldsymbol{\psi}(\mathbf{y})-\mathbf{x})+\mathbf{K}\begin{pmatrix}\psi_1(\mathbf{y})-x_1\\ \psi_2(\mathbf{y})-g_2(x_1,dx_1/dt,...,d^{n-1}x_1/dt^{n-1})\\ \psi_3(\mathbf{y})-g_3(x_1,dx_1/dt,...,d^{n-1}x_1/dt^{n-1})\\ \cdots\\ \psi_n(\mathbf{y})-g_n(x_1,dx_1/dt,...,d^{n-1}x_1/dt^{n-1})\end{pmatrix}. \tag{4}$$

Since $x_j=g_j(x_1,dx_1/dt,...,d^{n-1}x_1/dt^{n-1})$, $(j=2,3,\cdots,n)$, then, Eq. (4) can be simplified as

$$\dot{\mathbf{e}}=(\mathbf{A}+\mathbf{k})\mathbf{e}, \tag{5}$$

Form error dynamics system (5), we can obtain that $e_i=0(i=1,2,\cdots,n)$ is a unique equilibrium point of system (11). Since all the eigenvalues of $\mathbf{A}+\mathbf{k}$ satisfy $\mathrm{Re}(\lambda_i)<0$ ($i=1,2,\cdots,n$), then the unique equilibrium point $e_i=0(i=1,2,\cdots,n)$ of error dynamics system (5) is asymptotically stable, i.e. $\lim_{t\to\infty}\|\boldsymbol{\psi}(\mathbf{y})-\mathbf{x}\|=0$, which implies that the generalized projective synchronization between (2) and (3) can be achieved. The proof is completed. ■

4. Simulation results

To illustrate the effectiveness of the GPS scheme, the GPS for unified chaotic system is considered and the numerical simulations are performed.

Let $\alpha=0.5$, the chaotic attractor of unified chaotic system is shown as Fig.1. Choose the nonlinear map $\boldsymbol{\psi}(\mathbf{y}):R^3\to R^3$ as follows,

$$\boldsymbol{\psi}(\mathbf{y})=\begin{pmatrix}\psi_1(\mathbf{y})\\ \psi_2(\mathbf{y})\\ \psi_3(\mathbf{y})\end{pmatrix}=\begin{pmatrix}y_1y_3\\ 2y_2+\sin y_2\\ y_1-y_2\end{pmatrix},$$

where $y_1\neq 0$. Its Jacobian matrix $\mathbf{D}\boldsymbol{\psi}(\mathbf{y})$ is $\mathbf{D}\boldsymbol{\psi}(\mathbf{y})=\begin{pmatrix}y_3 & 0 & y_1\\ 0 & 2+\cos y_2 & 0\\ 1 & -1 & 0\end{pmatrix}$. The inverse matrix $\mathbf{D}^{-1}\boldsymbol{\psi}(\mathbf{y})$ is

$$\mathbf{D}^{-1}\boldsymbol{\psi}(\mathbf{y})=\begin{pmatrix}0 & (2+\cos y_2)^{-1} & 1\\ 0 & (2+\cos y_2)^{-1} & 0\\ y_1^{-1} & -y_3y_1^{-1}(2+\cos y_2)^{-1} & -y_3y_1^{-1}\end{pmatrix}.$$

The drive system is

$$d\mathbf{x}/dt=\mathbf{A}\mathbf{x}+\mathbf{h}(x_1,dx_1/dt,d^2x_1/dt^2) \tag{6}$$

where $\mathbf{A}=\begin{pmatrix}-25\alpha-10 & 25\alpha+10 & 0\\ 28-35\alpha & 29\alpha-1 & 0\\ 0 & 0 & -\alpha/3-8/3\end{pmatrix}$, $\mathbf{h}(x_1,dx_1/dt,d^2x_1/dt)=\begin{pmatrix}0\\ -x_1g_3(x_1,dx_1/dt,d^2x_1/dt^2)\\ x_1g_2(x_1,dx_1/dt,d^2x_1/dt^2)\end{pmatrix}$.

So, the response system is described as:

$$\dot{\mathbf{y}}=\mathbf{D}^{-1}\boldsymbol{\psi}(\mathbf{y})[\mathbf{A}\boldsymbol{\psi}(\mathbf{y})+\begin{pmatrix}0\\ -x_1g_3(x_1,dx_1/dt,d^2x_1/dt^2)\\ x_1g_2(x_1,dx_1/dt,d^2x_1/dt^2)\end{pmatrix}+\mathbf{K}\begin{pmatrix}\psi_1(\mathbf{y})-x_1\\ \psi_2(\mathbf{y})-g_2(x_1,dx_1/dt,d^2x_1/dt^2)\\ \psi_3(\mathbf{y})-g_3(x_1,dx_1/dt,d^2x_1/dt^2)\end{pmatrix}]. \tag{7}$$

Choose $\mathbf{K}=\mathrm{diag}(0,-25,0)$, we can obtain that the eigenvalues of the matrix $\mathbf{A}+\mathbf{k}$ are -0.6752, -33.3248 and -2.8333, respectively. Therefore, all eigenvalues of the matrix $\mathbf{A}+\mathbf{k}$ have negative real parts.

According to the Theorem 1 the GPS between (6) and (7) can be achieved. For example, let the initial values of drive system (6) and response system (7) as follows: $(x_1(0), x_2(0), x_3(0)) = (1,2,3)$ and $(y_1(0), y_2(0), y_3(0)) = (0.1,0.2,0.3)$, respectively. The simulation result of GPS between systems (6) and (7) is shown in Fig. 2.

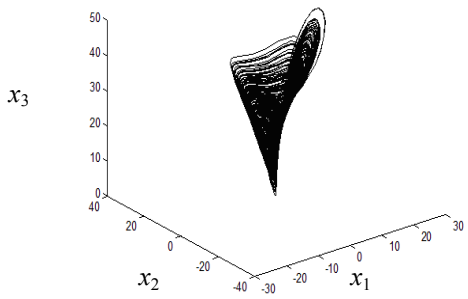


Fig.1 Chaotic attractor of unified chaotic system for $\alpha = 0.5$

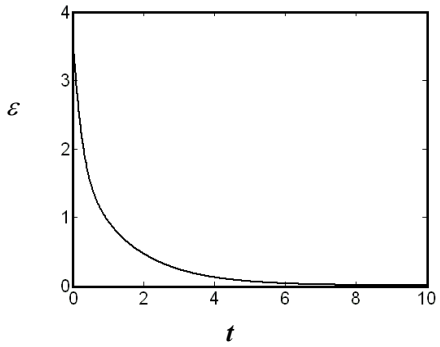


Fig.2 The generalized projective synchronized result of unified chaotic system for $\alpha = 0.5$

5. Conclusion

In this paper, we present that all the state variables for a class of continuous chaotic systems can be yielded via a specific sate variable and its time derivatives. Furthermore, we investigate the generalized projective synchronization for the chaotic systems and develop a new synchronization scheme. This synchronization approach is theoretically and numerically studied. One example is used to illustrate the effectiveness of the proposed synchronization method.

References

[1] L.M. Pecora, T.L. Carroll: Phys. Rev. Lett. Vol.64(1990), p. 821.

[2] T.L. Carroll, L.M. Pecora: IEEE Trans. Circuits Syst. 38(1991), p.453

[3] A.N. Miliou, I.P. Antoniades, S.G. Stavrinides: Nonlinear Anal. Vol.8(2007), p.1003

[4] C. Aguirre, D. Campos, P. Pascual, E. Serrano: Neurocomputing Vol.(69)2006, p.1116

[5] P. Zhou, X. F. Cheng, N.Y. Zhang: Chinese Physics Vol.17 (2008), p.3252

[6] P. Zhou, L. J. Wei, X.F.Cheng: Chinese Physics B Vol.18 (2009), p.2674

[7] P. Zhou, Y.X. Cao, X.F.Cheng: Chinese Physics B Vol.18 (2009), p.1394

[8] P. Zhou, L. J. Wei, X.F.Cheng: Acta Physica Sinica Vol.58 (2009), p.5201

[9] P. Zhou, X.H. Luo, H. Y.A. Chen: Acta Physica Sinica Vol.54(2005), p.5048

[10] R. Mainieri, J. Rehacek: Phys. Rev. Lett. Vol.82(1999), p.3042

[11] C.Y. Chee, D.L. Xu: Chaos, Solitons & Fractals Vol.23(2005), p.1063.

[12] X.Y. Wang, Y.J. He: Physics Letters A Vol.372(2008), p.435

[13] X.S. Yang: Physics Letters A Vol.260 (1999), p.340

[14] X.S. Yang: International Journal of Bifurcation and Chaos Vol.(12)2002, p.1159

[15] J. Lu, G. Chen, S. Zhang: International Journal of Bifurcation and Chaos Vol.(12) 2002,p.1001

[16] C.X. Liu, T. Liu, K. Liu and L. Liu: Chaos, Solitons & Fractals Vol.(22)2004, p.1031

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.262

Solving two-person zero-sum game by Matlab

Yanmei Yang[1, a], Yan Guo[2, b], Lichao Feng[1,c], and Jianyong Di[3,d]

[1] College of Science, Heibei Polytechnic University, Tangshan, 063009, China;

[2]Department of Mathematics & Physics, North China Electric Power University, Baoding, 071003, China

[3] College of light industry, Heibei Polytechnic University , Tangshan, 063009 , China ;

[a]yym04@163.com, [b]guoyan12@126.com, [c] flch19820520@yahoo.cn, [d]dijianyong60@126.com

Keywords: game theory; two-person zero-sum games; linear programming; Matlab;

Abstract: In this article we present an overview on two-person zero-sum games, which play a central role in the development of the theory of games. Two-person zero-sum games is a special class of game theory in which one player wins what the other player loses with only two players. It is difficult to solve 2-person matrix game with the order m×n(m≥3,n≥3). The aim of the article is to determine the method on how to solve a 2-person matrix game by linear programming function linprog() in matlab. With linear programming techniques in the Matlab software, we present effective method for solving large zero-sum games problems.

Introduction

Game theory is a branch of applied mathematics that is used in the social sciences, most notably in economics [1]. Two-person zero-sum games refer to games of pure conflict, which is the simplest case of game theory. In zero-sum games the total benefit to all players in the game, for every combination of strategies, always adds to zero, the payoff of one player is the negative of the payoff of the other player [2].

Two-person zero-sum games play a central role in the development of the theory of games, but it is difficult to solve 2-person matrix game with the order m×n (m≥3,n≥3). The connection with linear programming was discovered by von Neumann in 1947. Mixed strategies must exist for two-player zero-sum games had been proven by John von Neumann and Oskar Morgenstern. The Minimax Theorem is a simple consequence of the Duality Theorem of linear programming [2,3]. So using linear programming techniques can solve two-person zero-sum games. In this paper we discuss how to convert a two-person zero-sum game into a linear programming problem

Linprog command in Matlab

Linear programming is one of the most widespread methods used to solve management and economic problems. The command linprog from the optimization toolbox implements the simplex algorithm to solve a linear programming problem. Matlab uses the following format for linear programs:

$$\min z = f^T x$$
$$s.t. \quad \begin{cases} A \cdot x \le b \\ Aeq \cdot x = beq \\ lb \le x \le ub \end{cases}$$

Where f, x, b, beq, lb, and ub are vectors, and A and Aeq are matrices.

x = linprog(f, A, b, Aeq, beq, lb, ub) solves the above linear programming min $f^T x$ such that $A \cdot x \le b$. While additionally satisfying the equality constraints $Aeq \cdot x = beq$, set A = [] and b = [] if no inequalities exist. Defines a set of lower and upper bounds on the design variables, x, so that the solution is always in the range $lb \le x \le ub$, set Aeq = [] and beq = [] if no equalities exist [4].

For example, our original problem is translated into the format:

$$\min Z = 4x_1 + 3x_2$$
$$\begin{cases} x_1 + x_2 \le 5 \\ -6 \le x_1 \le 10 \\ -1 \le x_2 \le 4 \end{cases}$$

Solve the problem by Matlab, first we set up the vectors and matrices:

```
c=[4,3];
a=[1,1];
b=[5];
lb=[-6;-1];
ub=[10;4];
```

Finally, input the command x=linprog(c,a,b,[],[],lb,ub). Applying with the command linprog in optimization toolbox can be solved the above problem.

Solving two-person zero-sum game by Matlab

It is still difficult to solve the matrix game for high order, although there are many solving ways. The Minimax Theorem is the fundamental theorem in the theory of two-person zero-sum game. The Minimax Theorem as follow: For every two-person, zero-sum game with finite strategies, there exists a value V and a mixed strategy for each player, such that (a) Given player 2's strategy, the best payoff possible for player 1 is V, and (b) Given player 1's strategy, the best payoff possible for player 2 is −V. It can be proved by techniques from linear programming [5-7]. Seeing the relationship of the Minimax Theorem and Duality Theorem, linear programming can be used to solve two-person zero-sum game.

We have an m×n payoff matrix $A=(a_{ij})_{m\times n,}$ The entry a_{ij} represents the payoff to the row player when she picks the i^{th} strategy and column player picks his j^{th} strategy. The two-person zero-sum game can be solved by solving thus two linear programming:

$$\min Z=\sum_i x_i' \qquad \max w=\sum_i y_j'$$

$$\begin{cases}\sum_i a_{ij}x_i' \geq 1 & (j=1,2,\cdots,n)\\ x_i' \geq 0 & (i=1,2,\cdots,m)\end{cases} \qquad \begin{cases}\sum_j a_{ij}y_j' \leq 1 & (i=1,2,\cdots,m)\\ y_i' \geq 0 & (j=1,2,\cdots,n)\end{cases}$$

Suppose we want to solve the two-person zero-sum game, which payoff matrix as follow,

$$\begin{pmatrix}3 & -2 & 1\\ -1 & 4 & 2\\ 2 & 2 & 3\end{pmatrix}$$

Step 1 Reduce the payoff matrix by dominance. The above payoff matrix is a simplest form.
Step 2 Convert to a payoff matrix with no negative entries by adding a suitable fixed number to all the entries. Adding 3 to every element of the above payoff matrix, then the new matrix is

$$\begin{pmatrix}6 & 1 & 4\\ 2 & 7 & 5\\ 5 & 5 & 6\end{pmatrix}$$

Step 3 Solve the associated standard linear programming problem. Solving thus two linear programming:

$$\min x_1+x_2+x_3 \qquad \max y_1+y_2+y_3$$

$$\begin{cases}6x_1+2x_2+5x_3 \geq 1\\ x_1+7x_2+5x_3 \geq 1\\ 4x_1+5x_2+6x_3 \geq 1\\ x_i \geq 0 \quad i=1,2,3\end{cases} \qquad \begin{cases}6y_1+y_2+4y_3 \leq 1\\ 2y_1+7y_2+5y_3 \leq 1\\ 5y_1+5y_2+6y_3 \leq 1\\ y_i \geq 0 \quad i=1,2,3\end{cases}$$

Step 4 Calculate the optimal strategies by Matlab

Translated LP $\min x_1+x_2+x_3$ $\begin{cases}6x_1+2x_2+5x_3 \geq 1\\ x_1+7x_2+5x_3 \geq 1\\ 4x_1+5x_2+6x_3 \geq 1\\ x_i \geq 0 \quad i=1,2,3\end{cases}$ into the format $\min x_1+x_2+x_3$ $\begin{cases}-6x_1-2x_2-5x_3 \leq -1\\ -x_1-7x_2-5x_3 \leq -1\\ -4x_1-5x_2-6x_3 \leq -1\\ x_i \geq 0 \quad i=1,2,3\end{cases}$

set up the vectors and matrices:

```
c=[1,1,1];
a=[-6,-2,-5;-1,-7,-5;-4,-5,-6];
b=[-1,-1,-1];
lb=[0;0;0];
```

To find the optimal solution, the optimization toolbox has the command linprog: x=linprog(c,a,b,[],[],lb)

$$\max\ y_1+y_2+y_3 \qquad \min\ -y_1-y_2-y_3$$

Translated LP $\begin{cases} 6y_1+y_2+4y_3\le 1 \\ 2y_1+7y_2+5y_3\le 1 \\ 5y_1+5y_2+6y_3\le 1 \\ y_i\ge 0 \quad i=1,2,3 \end{cases}$ into the format $\begin{cases} 6y_1+y_2+4y_3\le 1 \\ 2y_1+7y_2+5y_3\le 1 \\ 5y_1+5y_2+6y_3\le 1 \\ y_i\ge 0 \quad i=1,2,3 \end{cases}$

Solved command in Matlab is

```
clear
c=[-1,-1,-1];
a=[6,1,4;2,7,5;5,5,6];
b=[1,1,1];
lb=[0;0;0];
y=linprog(c,a,b,[],[],lb)
```

The solution of above two LP is $x'=\left(0,0,\frac{1}{5}\right), y'=\left(\frac{4}{25},\frac{1}{25},0\right), v'=\frac{1}{5}$

The final result of the two-person zero-sum game is

$$x^*=(0,0,1), y^*=\left(\frac{4}{5},\frac{1}{5},0\right), v=5-3=2$$

Step by step, we show how to use this linear programming method solve the two- person zero-sum game in Matlab. The help of computer tools, we can greatly reduce the solving work time.

Conclusion

Matlab is a high-level technical computing language. In optimization toolbox of Matlab, linear programming problem can be solved simply by the command linprog. Many algorithms have been proposed for solving two-player zero-sum games, but it is not easy to solve for large matrix game. Using linear programming is efficient method. Especially, it uses the linear programming function linprog() in Matlab, calculates the two-person zero-sum game. This method is efficient and effective for large matrix game.

References

[1] Fernandez, L F.; Bierman, H S., Game theory with economic applications, Addison-Wesley, (1998)
[2] Dutta. Prajit K, Strategies and games: theory and practice, MIT Press, (1999)
[3] Osborne, Martin J. (2004), An introduction to game theory, Oxford University Press
[4] B.R. Hunt, R.L. Lipsman, and J.M. Rosenberg. A Guide to MATLAB, for beginners and experienced users. Cambridge University Press, 2001
[5] Philip D.Straffin. Game theory and strategy, Mathematical association of America, 1993.
[6] Andrey Garnaev, Search games and other applications of game theory, Lecture notes in economics and mathematical systems 485, Springer, 2000.
[7] Mehrotra, S., On the Implementation of a Primal-Dual Interior Point Method, SIAM Journal on Optimization, Vol. 2, pp. 575–601, 1992.

©(2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.266

Effects of Overburden on Deformation and Failure of Anti-dip Slopes under the Action of Seismic Load

Fayou A [1,2,3,a], Jiming Kong [1,3,b], Yu Tang [4,c], Zhenqiang Ni [1,2,3,d], Yun Cui [1,2,3,e], Shujun Tian [1,2,3,f]

[1]Institute of mountain Hazards and Environment， Chinese Academy of Sciences, Chengdu Sichuan 610041, China

[2]Graduate University of Chinese Academy of Science , Beijing， 100101, China

[3]Key Laboratory of mountain Hazards and Surface Process， Chinese Academy of Sciences, Chengdu Sichuan 610041, China

[4]Sichuan Agricultural University，Yaan Sichuan 625014, China

[a]afayou@163.com, [b] jimingk@imde.ac.cn

Keywords: anti-dip slope; overburden; seismic load; deformation and failure

Abstract. A large number of mountain disasters were induced by Wenchuan Earthquake, and as one type of them, the deformation and failure modes of anti-dip slopes were influenced by overburden. Under the condition of no overburden, the collapse will mainly occur; while under the overburden condition, it is the shallow slump mainly occured. Taking the landslide of Wangjia Mountain in Longmenshan Town, Pengzhou City as an example, the anti-dip slopes were stimulated by using the finite element method in this study, where the influences of overburden on deformation and failure of anti-dip slopes were analyzed under the static condition with seismic load effect. The related analysis showed that significant influences of overburden existed on the deformation and failure of slopes, resulting into great increments of slope displacement, stress and sliding distance. The analytical results were consistent with the characteristics that the large-scale deformation and failure of anti-dip slopes mainly occurred on the slopes with overburden in the Wenchuan Earthquake.

Introduction

At 14:28 on May 12, 2008, Ms 8.0 earthquake occurred in Wenchuan, Sichuan Province of China, inducing a large number of geological disasters [1-7]. It was found through the investigation that a great many mountain disasters were related to the shallow overburden of slopes. As one type of mountain disasters induced by Wenchuan earthquake, the deformation and failure of anti-dip slopes were affected significantly by overburden, leading to different modes of deformation and failure [8-9]. Without the overburden, a relatively small number of disasters occurred to anti-dip slopes with a small scale, and the collapse was the main type. While under the overburden condition, a relatively large number of disasters usually occurred to anti-dip slopes with a large scale, and the shallow slump was the main type. In conclusion, the deformation and failure of anti-dip slopes was greatly affected by shallow overburden. Wenchuan earthquake occurred in the Longmen Mountain area, where anti-dip slopes were distributed widely. The effects of overburden on anti-dip slopes could not be ignored as one of concerns in the restoration and reconstruction of disaster areas after the earthquake. Therefore, the study on the effects of overburden on deformation and failure mechanism of anti-dip slopes was of a great practical and theoretical significance to the current restoration and reconstruction of disaster areas.

In the past few decades, a large number of studies were conducted on slope response and its stability under the action of seismic load by researchers at home and abroad, of which two aspects

Funded by Knowledge Innovation and Critical Direction Project of Chinese Academy of Sciences (KKCX1-YW-03).
Author introduction: Mr. Fayou A, born in Yunnan in 1979, is a doctoral candidate mainly studying the mitigation project.
Corresponding author: Mr. Jiming Kong is a researcher mainly conducting the studies on mountain disaster formation mechanism theory and mitigation project, born in Chongqing in 1956.

were the main focuses, i.e., post-earthquake field investigation and numerical simulation. Additionally, more studies were carried out on rock slopes aiming at the seismic response and the stability of high slopes, jointed rock mass and bedding rock mass [10-17]. However, the effects of overburden on deformation and failure of slopes were not considered in such analyses, and scarce studies were conducted on anti-dip slopes.

In conclusion, very scarce studies were conducted on the effects of overburden on deformation and failure mechanism of anti-dip slopes. Therefore, using ANSYS finite element method, the numerical simulation model of slopes was established with the landslide in Wangjiaping as an example in this study, where comparative analysis was carried out on the deformation and failure of anti-dip slopes under static condition and the action of seismic load with or without the overburden. In addition, the effects of overburden on deformation and failure of anti-dip slopes were studied.

Main Deformation and Failure Modes of Anti-dip Slopes in the Disaster Area after Wenchuan Earthquake

In the disaster area after Wenchuan earthquake, the disaster type of anti-dip slopes mainly displayed as collapse, while shallow landslide was mainly generated on the anti-dip slopes with surface overburden. According to the differences under the condition with or without the overburden, the deformation and failure modes of anti-dip slopes in Wenchuan earthquake were divided into two categories, as shown in Figure 1:

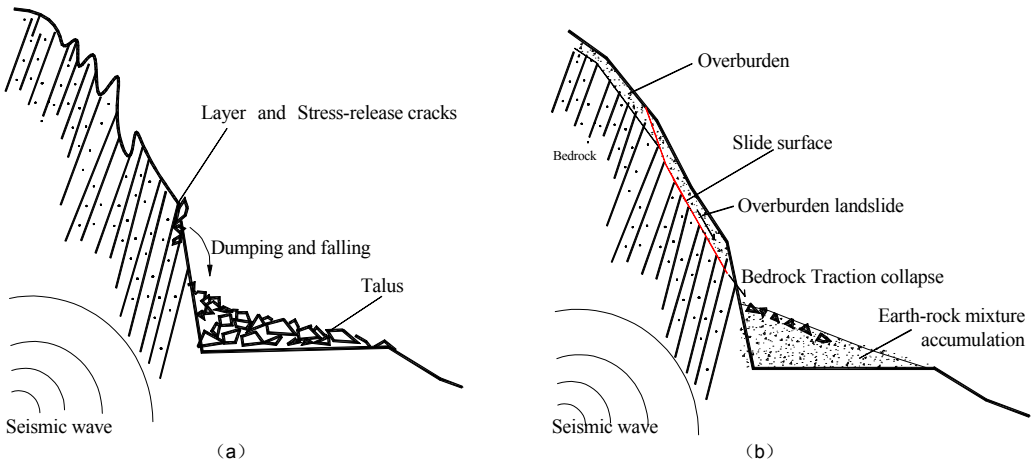


Fig. 1 Deformation and failure modes of anti-dip slopes in Wenchuan earthquake (a. collapse-fall type; b. shallow-slump type)

The first type was the collapse-fall type mainly generated on the slopes without overburden, where anti-dip rock mass was exposed. The collapse occurred to rock mass along the excavation face or adjacent to free face, leading to the formation of talus at the base of slope after rock falling. The second type was the shallow-slump type mainly generated on the slope covered by accumulative formation, where the free face of slope showed a binary structure (the upper section was overburden, while the lower section was bedrock). Shallow slump was induced under the seismic effect, leading to the collapse of the bedrock at the lower section as well as the mixed accumulation of deposit rocks.

Brief Introduction of Wangjiaping Landslide

The landslide occurred in Wangjiaping, Longmenshan Town, Pengzhou City in Sichuan with the geographic coordinates of N31 ° 17′ 9.3″ and E103 ° 49′ 48.7″. More than 50 km away from Wenchuan epicenter, this landslide was classified as the shallow-slump type. The landslide body was 120m in length and 150m in width with a scale of about more than 50,000 m^3, causing 150m Pengzhou-Baishuihe Highway to be buried, as shown in Figure 2.

Fig. 2 Overall picture of the landslide

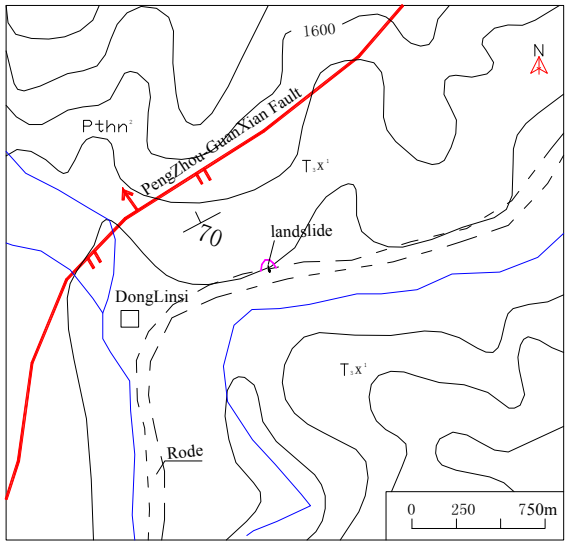


Fig. 3 Topographic and geologic map of landslide area

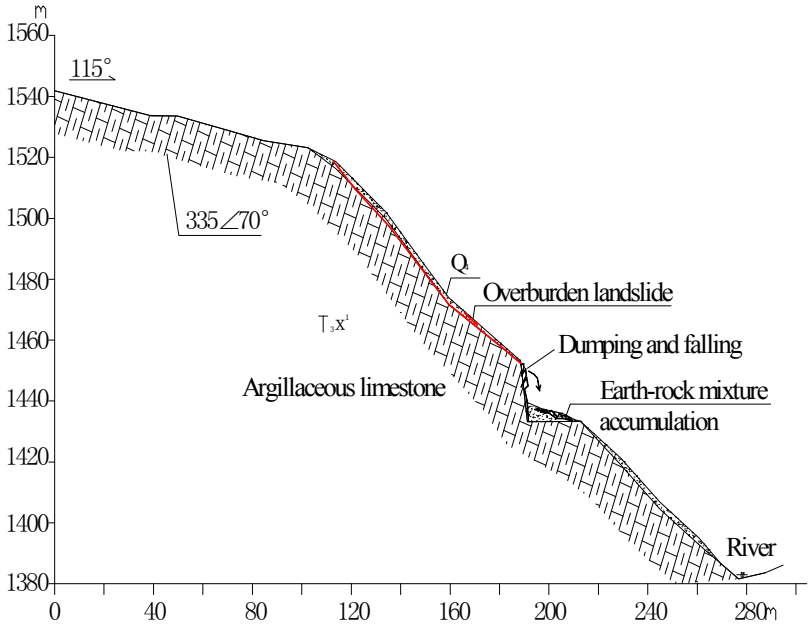


Fig. 4 The geological profile of slope in landslide area

The area where the landslide occurred was classified as a valley geomorphic unit with middle mountainous terrain and medium cutting. Dabaoshan in the northwest was the location with the highest altitude of about 2,000 m, while the bottom of Baishuihe valley had the lowest altitude of 1,350 m, leading to a cutting depth above 500 m. The landslide occurred on the right bank of Baishuihe with the posterior elevation of about 1,520 m and the terrain slope of about 35 ~ 40 °. The highway crossed the middle section of slope with the excavated slope height of 15 ~ 20 m, where luxuriant vegetation existed. Consisting of medium-thick bedded gray argillaceous limestones, bedrocks on the slope were classified as Silurian Maoxian Group (T_3mx^1) and mainly distributed in the lower section of both valley banks with the occurrence of 335 ° ∠ 70 ° and overall N-W orientation. The landslide occurred on the anaclinal slope with the orientation of 150 °. Composed of plastic gravelly silty clay with a small quantity of breakstones, Quaternary eluvium and diluvium (Q_4^{el+dl}) were widely distributed on the slope with the thickness between 1 and 3 m. The landslide area was located to the southeast side of Peng-Guan Fault with a distance of 0.8 km away. In geology, this area was classified as Longmenshan fold-fault zone with intensive uplifts. Figure 3 shows the topographic and geologic map of landslide area, and Figure 4 shows the longitudinal profile of the slope on which the landslide occurred.

Establishment of Numerical Model for Slope Analysis

Based on the geological profile of slope engineering, as shown in Figure 4, the boundary size of the geometric model was appropriately extended on the premise of constant structure and thickness of strata, in order to achieve simulation accuracy. The model was divided into two layers, i.e., bedrocks, eluvium and diluvium. Physical and mechanical parameters of slope rocks were determined by the tests with reference to relevant experience. Physical and mechanical parameters of the slope are shown in Table 1.

Table 1 Physical and mechanical parameters of the slope

Accumulative formation	Density(kg.m^{-3})	Elastic modulus (Gpa)	Poisson's ratio	Cohesion (kpa)	Internal friction angle（° ）
	1,980	**0.001**	**0.3**	**16**	**26**
Bedrock	Density(kg.m^{-3})	Elastic modulus (Gpa)	Poisson's ratio	Interlayer cohesion (kpa)	Interlayer friction coefficient（° ）
	2,500	**25**	**0.25**	**20**	**0.6**

In order to ensure the simulation accuracy, the upper and lower boundaries as well as left and right boundaries were set 2.5-times more than excavated slope height. The elastic material was adopted as the rock material for the calculation; soil mass and rock aspects were plastic materials. D-P yield criterion was used in the calculation as well as ANSYS large deformation analysis. Boundary conditions were set as follows: vertical and horizontal displacement constraints were imposed at the left and lower boundary, and no constraint existed for the slope body as free face of seismic response. Through birth-death element, the calculation models were converted under the condition with or without the overburden.

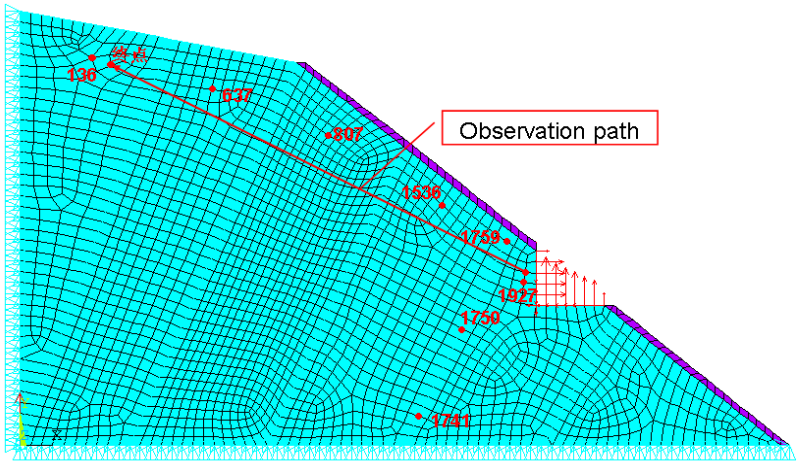


Fig. 5 Finite element calculation model of the slope on which Wangjiaping landslide occurred

The transient dynamic analysis technology was adopted for dynamic analysis of structure. The used seismic waves were referred to “Earthquake Records in Tianjin, China (1975)” with interception record duration of 5 s and time interval of 0.01 s. Figure 5 shows the finite element calculation model of the slope on which Wangjiaping landslide occurred. Two observation nodes (located on the layer most adjacent to excavated free face with serial number of 1759 and the internal rock mass above excavated free face with serial number of 1536 respectively) were taken in the figure as well as an observation path (pointing to the upper slope from excavated free face).

Effects of Overburden on Deformation and Failure of Slopes

Effects on Slopes under the Static Condition. Figure 6 (a) shows the displacement variations of anti-dip slopes along the observation path under the conditions with or without the overburden. It can be seen from the figure that the slope displacement was significantly increased with the overburden, and relatively obvious increase was shown in the vertical displacement.

Figure 6 (b) shows the stress variations of anti-dip slopes along the observation path under the conditions with or without the overburden. It can be seen from the figure that the slope stress was significantly increased with the overburden, and relatively obvious increase was shown in the vertical stress.

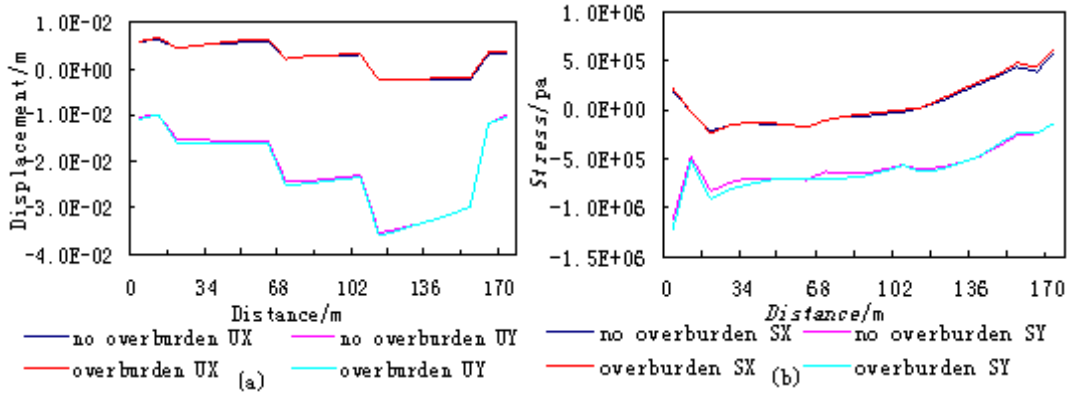


Fig. 6 Under the static condition, the variations of displacement (a) and stress (b) of slopes with or without the overburden along the observation path

Overburdens also had great influences on interlayer friction stress and sliding distance between strata. It can be seen from the analysis that the magnitude of both positive friction stress and negative one was significantly increased with the overburden.

Effects on Slopes under the Action of Seismic Load. Figure 7 shows the displacement and stress time-history curves of anti-dip slopes at the observation node 1536 with or without the overburden. It can be seen from the figure that the displacement and stress of slopes were significantly increased with the overburden. With the continuity of seismic time-history, the slope displacement showed a gradually increasing trend with the overburden.

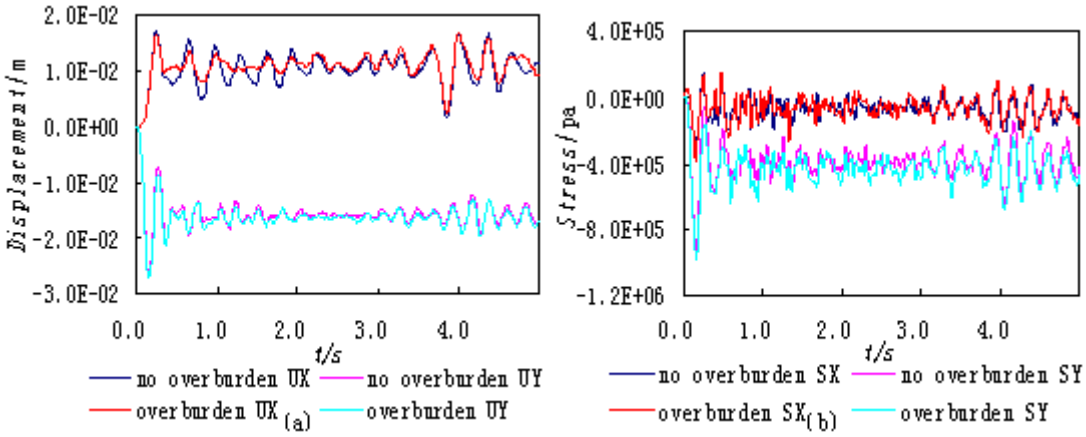


Fig. 7 Time-history curves of displacement (a) and stress (b) at slope node 1536 with or without the overburden

Figure 8 shows the time-history curves of sliding distance and friction stress at observation node 1759 with or without the overburden. It can be seen from the figure that the interlayer friction stress

between strata was decreased significantly with the overburden and sliding distance between strata was increased significantly with the overburden.

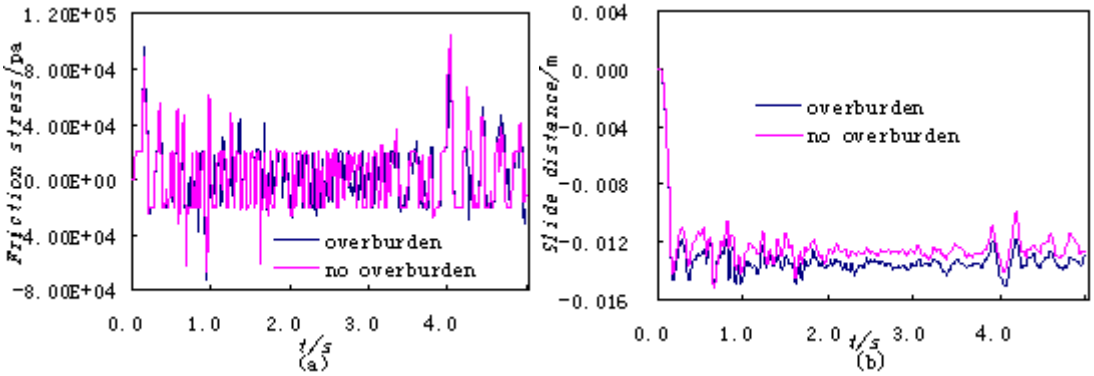


Fig. 8 Time-history curves of friction stress(a) and sliding distance(b) at observation node 1759 with or without the overburden

Conclusions and Recommendations

Based on analyses and conclusions on the deformation and failure of anti-dip slopes with excavation damages in disaster areas after Wenchuan earthquake, the main modes of deformation and failure of anti-dip slopes were obtained. Taking Wangjiaping landslide in Longmenshan Town, Pengzhou City as a typical example, the ANSYS software was used to simulate the slopes in this study; meanwhile, the effects of overburden on deformation and failure of anti-dip slopes were analyzed. The following conclusions can be obtained from analyses:

(1) Overburdens had significant influences on deformation and failure of anti-dip slopes. With the overburden, the slope displacement, stress and interlayer sliding distance were significantly increased and friction stress was decreased either under static condition or under the action of seismic load.

(2) The analysis results were consistent with the investigation results that deformation and failure of anti-dip slopes occurred in a small quantity with a small scale without the overburden, while the situation was right on the contrary with the overburden.

(3) It is recommended to pay more attention to the anti-dip slopes with shallow overburden in the restoration and reconstruction of disaster areas.

References

[1] Y.F.Li, Z.Y.Wang, W.J.Shi, X.Z: Journal of Mountain Science. Vol. 7(2010), p. 226-233.

[2] P Cui, Y Zhu, Y Han et al: Landslides. Vol. 6(2009), p. 209-223.

[3] R.Q.Huang, W.L.Li: Science in China Series E(Technological Sciences).Vol. 52(2009), p. 810-819.

[4] Q.Xu, X.M.Fan, R.Q.Huang, Cee Van Westen: Bulletin. Vol. 68(2009), p. 373-386.

[5] C.Z.Wang, G.J.Chen, R.Z.Tan, L.K.Wen: Journal of Sichuan University (Engineering Science Edition). Vol. 41(2009), p. 84-88.

[6] F.Y.A, J.M.Kong, S.J.Tian, Y.Cun, W.P.Wu: Geology and Exploration, Vol. 45(2009), p.312-320.

[7] P.Cui, Y.S.Han, X.Q.Chen: Journal of Sichuan University (Engineering Science Edition). Vol. 41(2009), p. 35-42.

[8] J.M.Kong, F.Y.A, W.P.Wu: Journal of Soil and Water Conservation. Vol. 23(2009), p. 66-70.

[9] J.M.Kong, Y.Cun, S.J.Tian, F.Y.A: Journal of Sichuan University (Engineering Science Edition). Vol. 41(2009), p. 119-124.

[10] J.L.J.Bommer,Carlos E.Rodry'guez: Engineering Geology. Vol.63(2002), p. 189-220.

[11] Y.L.He,S.Y.Lu,Y.H.Duan: Journal of Yangtze River Scientific Research Institute. Vol.15(1998), p. 35-38.

[12] Y.L.He, S.Y.Lu: Chinese Journal of Geotechnical Engineering. Vol.20(1998), p.66-68.

[13] D.Peng,L.L.Cheng: Journal of Yangtze River Scientific Research Institute. Vol.18(2001), p.47-49.

[14] L.J.Tao, S.R.Su, Z.Y.Zhang: Journal of Engineering Geology. Vol.9 (2001), p.32-38.

[15] W.Y.Xu, S.Y.Xie, Z.C.Qi: Journal of University of Hydraulic and Electric Engineering/Yichang. Vol.20 (1998), p.11-15.

[16] K.Q.Xiao, H.B.Li,etal: Rock and Soil Mechanics. Vol.28 (2007), p.1557-1564.

[17] D.Meng, X.C.Meng, Y.Zhao: Coal Engineering. Vol.1 (2009), p.77-79.

Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.273

Preliminary Discussion on the Applications of Cloud Computing in the Bank System

Zhiqin Huang [1, a], Jinli Zhang[2,b] and Huizhen Zhou[3,c]

[1] Science of school, University of Jinan, Jinan, China

[2] College of mathematics and information science, Qinghai normal university, Xining, China

[3]Xinzhou subsidiary teachers college of foreign languages school, Xinzhou, China

[a] ss_huangzq@ujn.edu.cn, [b] zhangjl_happy@126.com, [c] Wsliu482@sohu.com

Keywords: cloud computing. banking industry. virtualization. index.

Abstract. Cloud computing with its powerful feature has been used preliminary in some commercial areas in the world, cloud computing also is used in IT industry, and it is a new revolution in IT industry. Cloud computing is of great concerned for its characteristics of the security and reliability of data mass storages and the low requirement on the customer premise equipment, especially in the financial industry. This paper describes the basic concept, characteristics and development status of cloud computing, and on the basis, the application of cloud computing in the banking industry is discussed, and the possible problems are pointed out.

Introduction

About cloud computing：

Network cloud is a symbol, then the cloud, the most liberal interpretation is "Network Computing". Cloud computing can be seen as the developing of parallel computing, distributed computing and grid computing, or can be say that the cloud computing is the commercial realization of computer science concepts [1] . This is a common vision for cloud computing at present. The basic principle of cloud computing refers to the gathering and sharing the computing, storage, service components, network software, IT hardware and software, and providing a convenient, fast network technology equipped with on-demand service for the users through an open, unified way [2] .

Cloud computing and grid computing are very similar, but the cloud computing is more power and superior than the grid in computing power, integrated and security. Cloud computing can be taken as a group of virtual servers working together on the internet, applications is very easy to implement through internet. In cloud computing, completion of operations is the large data centers and a powerful main server as a network application, network services. In grid computing, large task is divided into small tasks, then these small tasks are put into many parallel systems to run, the difference between cloud computing and grid computing is that the structure of cloud computing is a collection of resources, management of these resources can be dynamically, at any point in time these resources can be monitored and maintenance.

Cloud computing includes the following three basic elements:

(1) the gathering and sharing of resources: achieve a variety of statistical multiplexing of resources.

(2) the opening and uniform access to resources: providing with an open, unified access interface for the users who can access to various resources just like using water and electricity, and the users do not need to concern about the details of various resources.

(3) Simulation and isolation of resources: it can provide services equipped with on-demand for the users, so that each user has the feeling of using resources in the form of exclusive (on-demand service). Virtualization is the essence of cloud computing, or the fundamental property of cloud

computing is virtualization. As mentioned above, cloud computing at this stage is only achieved by virtualization of resources. With the further research and development to the technology and the relevant concept of cloud computing, in the near future, it will gradually realize the virtualization operations and virtualization management, virtualization is the vitality of cloud computing. printer.

The characteristics of cloud computing

Cloud computing has the following main features:

(1)Huge amounts of data, very large scale: “cloud” with considerable scale, for example, Google cloud computing already has more than 100 million servers.

(2) Virtualization: Cloud computing allows users at any location, using a variety of terminal access to application services. The requested resources come from the "cloud" rather than fixed physical entity. Application is running in the "cloud" somewhere, but in fact users do not need to know and do not have to worry about the specific location of running applications, the only need is a laptop computer or a mobile phone which can achieve all we need through network services, even supercomputing.

(3) High reliability: there are many data in “cloud”, and in order to protect service high reliability, some measures have been used in “cloud”, such as fault tolerance with multiple copies, interchange with the structure of computing nodes, the use of cloud computing are more reliable than using the local computer

(4) Universal property: cloud computing is not for specific applications, the ever-changing application can be constructed under the support of “cloud” computing, a “cloud” can support different applications to run simultaneously.

(5) High expansibility: the scale of “cloud” can be dynamic stretched, which can meet the needs of the application and the users’ scale of growth.

(6) On-demand service: "cloud" is a huge resource pool, users can buy it by demand; Cloud can be billed just as water supply, electricity, LPG.

(7) low cost: as the special fault-tolerant measures of "cloud", very low-cost nodes can be used to form cloud, automated centralized management of "cloud" makes a large number of companies without the burden of increasingly high cost of data center management, and the universal of "cloud" also makes the utilization of resources greatly enhance than the traditional systems, So users can fully enjoy the low-cost advantage of "cloud".

Preliminary discussion on Cloud computing applications in the banking system

Cloud computing and the banking system

Now, many companies in the IT building is at a crossroads, the banking industry is no exception. IT complexity and low efficiency of the crisis is obvious and the situation became worse in the global economic crisis situation, and customers have encountered serious challenges in the growth of IT infrastructure. Faced with this situation, we can create a virtual version of the complex environment. Virtualization which is coupled with service management automation, system image, the standardization of version and procedures, will achieve significant operating advantages. This model may be understood as the dynamic infrastructure.

Cloud computing is in the purposes of improving a computing infrastructure and services. Because cloud computing has the following characteristics, it can be applied to the banking industry.

Cloud computing offers standards and determines the location and how to calculate delivery options: security, control, flexibility and SLA (service level agreements), this can help enterprises to improve competitive advantages. In addition, users can seamlessly insert into the existing service of the existing infrastructure at any time, this is the best options which are to shorten time and reduce the risk. For risk and time, companies need to consider many factors, but the most important is the IT organizations which are responsible for managing IT services. As more and more services can be

used, management systems and the mixed infrastructure management model will change. This is essential for the enterprises successfully integrating cloud calculation into the enterprises will be integrated into the internal of the enterprises.

Cloud computing embodies the economies of scale: the influence of the scale on the unit cost is the same as traditional infrastructure costs which will decrease with the increase of the size. Clearly, cloud service providers can significantly reduce infrastructure costs. If we use the virtualization, standardization, automation, and other basic cloud computing technology, the users can also reduce costs and create value, even in the implementation of IT procurement before the outside is no exception. Therefore, if the basic cloud computing is deployed, enterprises will be able to use cloud computing to take precautions to create a cost advantage. The cost model may differ because of work load and will gradually change, but it clearly shows the development of proprietary technology cloud or public clouds can have a huge advantage in creating value and decreasing cost.

The advantages of cloud computing combined with the banking industry:

The development of cloud computing will have a profound impact on the banking industry. The bank which is a resource-rich, large data, requires a lot of scientific computing, to serve the public sector. By the development of the current status and needs of the banking sector, banking and cloud computing have more entry points. We must balance the interests of maximizing and minimizing risk in the application of cloud computing in the bank. Therefore, we can transfer part of the businesses and resources to the public cloud, and this part should obtain benefits from the dynamic resource management. At the same time, most exclusive business and private resources (in particular customers' data and information) are on the internal cloud.

Public cloud:

(1) Calculation of banking data, cloud computing can bring the staff of banks powerful, flexible and low cost platform for collaboration and innovation, and cloud computing is a clear advantage of reducing utility computing costs. Using cloud computing, users can avoid the local construction, expensive computing systems. Through the payment of low-cost services, clients can complete the same calculations or processing. In the October 2007, Google 101 plans have facilitated Google and a IBM great cooperation - - - integrating global many universities to the computation "the cloud" which is similar with Google Google's "cloud" programs provide Google-style computing infinite capacity for students, researchers and entrepreneurs. In the near future, with the expansion of the "cloud" effects, we can enjoy powerful computing capability without the need to purchase a large number of infrastructure.

(2) Bank data storage, at present, banks stored large amounts of data in the respective terminal equipment. When combined with cloud computing, bank get the service only through a terminal tool access system, by a third-party service providers providing computing and storage resources, operation and maintenance. Banks do not need purchase equipment for lack of storage resources, just to the service providers to purchase storage services.

(3) Resources sharing among banks, web-based servers, storage, databases, and other cloud computing architecture, software and services provide an attractive platform for cooperation to major banks around the world. Putting these advanced application software resources on the clouds for the global bank workers to use, can save significant costs for banks.

(4) Bank and other enterprise's relation, enterprise is a very important part of the banks' customers originates. If the bank's information resources and business information resources place in the same cloud, they can use each other.

Inner cloud

(5) Personnel management, financial control, as the chart 1 shown, what the platform layer provides is PaaS (Platform-as-a-Service) serves. Its value lies in pre-defined and options. PaaS similar IaaS, but it includes operating systems and applications around specific essential services. PaaS can be described as a complete virtual platform, which includes one or more servers, operating systems and specific applications. Application procedure is provided by the application layer. Application layer is called SaaS (Software-as-a-Service) which deploys software from a centralized system, so that it can remotely run a model in a local computer or from the cloud. Salesforce is the

pioneers of a cloud computing software the same as a service company. 800 APP provides consummation PaaS software products in China. 800 APP have developed in almost all SaaS applications on the PaaS platform, including human resources management, financial management and accounting systems. These facilitates which are easy to use, full-featured management software will bring to the bank workers more efficient.

(6) Business model, the bottom of banks' inner cloud is the physical layer, including the most basic of hardware resources, such as computing resources, storage resources and network resources [6] . It provides HaaS (Hardware as a Service) services. For hardware resource virtualization, IT technology automation and how much to use of resources by the inevitable consequence of the pricing, banking customers can purchase or rent an entire data center to order services. HaaS is the distribution according to need, flexible, scalable, and is to reasonably control. This can not only save the cost of the inner bank, but also reduces the utilization of hardware resources idle and low risk. Uniform Resource layer can be decomposed into two small levels, a virtual layer and infrastructure layer.

Services provided by the virtual layer are SaaS (data-Storage-as-a-Service), and its value is stored in the form of metering services. Infrastructure layer provides services for the IaaS (Infrastructure-as-a-Service). IaaS only provides the virtual hardware and it has not software stack, so cloud computing allows banking staff upload pictures and music, which automatically generate web-based video presentations and be shared with others. So long as a computer or a mobile phone can make real-time, remote business communication, reporting that becomes very simple. Cloud computing is not only a combination of new technology,

It is also an innovation of business model. Many public areas, such as banking business computing power, processing capability, and the service capacity will be improved.

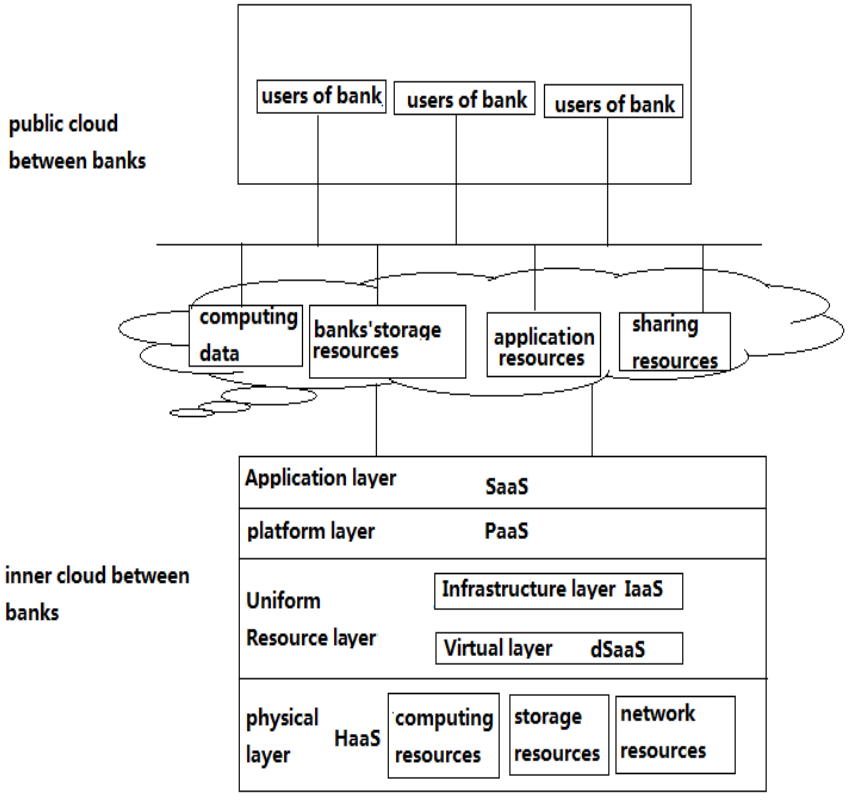


Fig. 1 **The internal structure of the banks' cloud and their external structure chart**

Issues requiring attention of cloud computing applied to banking

Classification of bank resources

The classification of banking resources will be a problem confronting the future: What should be put on public clouds, which should be placed on private cloud; what resources should be shared, what resources should be kept confidential.

Procurement of hardware changing into purchase storage services

In the cloud computing environment, storage is no longer a hardware, but a service. Will this change the user's memory for future procurement method to purchase from the procurement of hardware into storage services?

The banking sector is facing greater competition

As the impact of cloud computing expands in depth, this leads to maximize the sharing of resources, regional differences tend to minimize. So our level of business will face greater competition and challenges.

Summary

Cloud computing is the connotation of computing, storage, service components and network software, IT software, hardware virtualization. It is the integration of grid computing, distributed computing, utility computing, Web 2.0, SaaS and many other new technologies. It provides new business models by hiring an in-demand IT resources as the core. The development of technology and business models will undoubtedly have a major impact for the future of electronic information in the field of software industry and hardware industry. This will affect many areas, such as servers, storage, network infrastructure and middleware, operating systems, application software, and network services, and create a new IT applications. Banking system is based on data storage reliability and stability as the premise, taking full account of technology, nature, scientific management, ease of maintenance, operation efficiency, timely backup and restore reliability. These can be answered in the cloud. Therefore, cloud computing and the banking industry combined with great practical significance.

References

[1]Zhang Yu Jiang. Cloud opened to the school gate [J]. China Information Technology Education, 2008 (12) :93-94.

[2]Zhang Jian, Cao Ji-guang. Cloud Computing Technology of the Internet [J]. Telecommunication network technology ,2009-10-10 :1-3.

[3]ErichCiementi. Using of Cloud computing. [EB / OL]. [2009-7-23]. Http://www.ccu.com.cn/html/2009/0713/4884.html.

[4] Hu Ying. Discusses the cloud computation shallowly in the meteorological enterprise's forecast [J]. Computer Knowledge and Technology, 2009 (7) :5876-5879.

[5]Cloud computing application type classification [EB / OL]. [2008-11-12]. Http://www.cncloudcomputing.com/.

[6]Ian Foster, Yong Zhao, Ioan Raicu, Shiyong Lu. Cloud Computing and Grid Computing 360 -Degree C ompared. [EB/OL] . [2009-07-01] . http://download.csdn.net/source/1186804.

doi:10.4028/www.scientific.net/AMM.50-51.278

A Novel Heat Sink for High Power Light Emitting Diode—Phase Change Heat Sink

J. H. Xiang [1a], X. C. Liu [1b], C. L. Zhang [1c], Y. Tang [2d]

[1.] School of Mechanical and Electric Engineering, Guangzhou University, Guangzhou 510006, China

[2.] School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China

[a.] Email: furongbridge@yahoo.com.cn, [b.] Email: gdliuxiaochu@163.com, [c.] Email: nhzcl@163.com, [d.] Email: ytang@scut.edu.cn

Key words: High power LED; Phase change heat sink; Enhanced boiling; Wick; Ploughing-extrusion

Abstract: A novel heat sink was fabricated for packaging cooling of high power LED. Enhanced boiling structures in the evaporation surface were processed by ploughing-extrusion (P-E) and stamping methods, respectively. The cycle power of refrigerant was supplied by wick of sintered copper powder on internal surface of phase change heat sink. The experimental results showed that phase change heat sink was provided with a good heat transfer capability and the temperature of phase change heat sink reached 86.8℃ under input power of 10W LED at ambient temperature of 20℃.

1 Introduction

High power light emitting diode (LED) as a new type of solid-state light source possessed a number of advantages compared to traditional light sources such as the life time, environmental friendly and energy saving et al [1-2]. However, thermal management had become one of the main factors for application of high power LED due to greater spreading resistance of solid heat sink [3-5]. Efficiency of heat transfer can increase 10-100 time compared to the traditional solid heat sink at the same operating conditions [6]. So phase change heat transfer technology can be widely applied in the heat dissipation of optoelectronic and micro-electronics fields.

Sheng Liu[7] optimized a kind of closed micro-jet array cooling system of high power LED. The result showed that temperature of chip surface reached 44.2℃ when input power was 16.4W. Lu Xiang-you [8] made a kind of LHP as cooling system of high-power LED, and junction temperature of LED can be controlled below 100℃ under 100W input power. Lan Kim[9] set up a cooling system for high power LED array and the temperature of cooling system was decreased 24.3 ℃ after heat pipes were used. But these cooling system which were large size and complex structure, were difficult to be used as heat dissipation in packaging of LED.

In this study, the principle and structure of phase change heat sink as a heat transfer component for high power LED package was analyzed. Enhanced boiling structures in evaporation surface were processed by P-E and stamping, respectively. And the relationships between forming parameters were analyzed. Furthermore, copper powder was sintered as wick on internal surface of phase change heat sink. In addition, the performances were investigated under different power inputs and working fluids.

2 Fabrication of phase change heat sink

2.1 Structure and principle of phase change heat sink

Phase change heat sink of high power LED make up of vapor chamber, wick and boiling structure as shown in Fig.1. When heat is applied, the fluid at the location immediately vaporizes, and the vapor rushes to fill the vacuum. When vapor comes into contact with condenser surface, vapor began to release its latent heat of vaporization and turn into liquid. Then the condensed liquid returns to the heat source via capillary action or natural circulation. Higher effective thermal conductivity is obtained due to the high latent heat of a liquid during phase change.

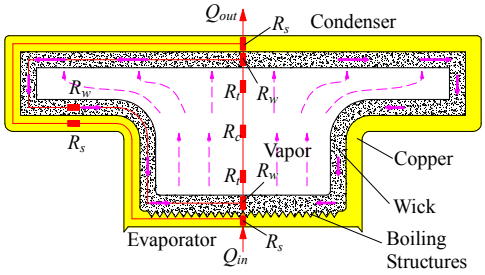


Fig.1 Schematic diagram of LED high-power phase-change heat sink structure and working principle

(R_s, R_w, R_l, R_c are thermal resistance of solid wall, wick, liquid vapor interface, vapor core, respectively)

2.2 Fabrication of 3D boiling structure

2.2.1 Forming of spiral circumferential grooves

Nitesh D. Nimkar [10] found that micro-grooves, cracks and fins could speed up the vaporization of refrigerant, so the boiling performance could be improved significantly. Tang [11] investigated the processing methods of micro-grooves, cracks and fins to improve the heat transfer performance. In this study, spiral micro-grooves along circumferential direction were fabricated by P-E.

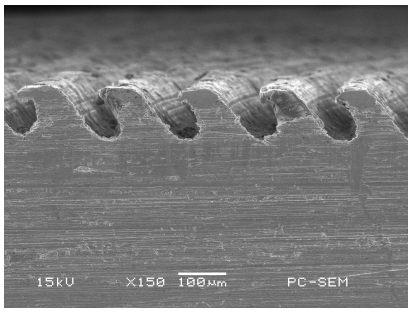


Fig.3 Boiling structure by P-E processing

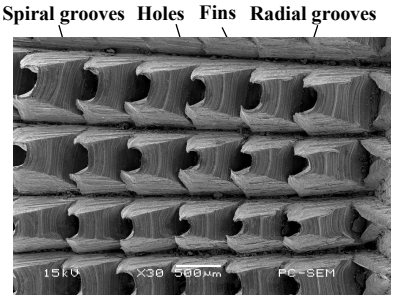


Fig.4 3D boiling structure

The experiment was carried out on the lathe C6132A1, and the red copper column with diameter of 10 mm, thickness of 1 mm was used as the work-piece. P-E tool was installed by way of parallel with copper columnar axis. The section SEM of spiral micro-grooves which was processed when P-E depth a_p was 0.3 mm, P-E interval d_p was 1.24 mm was shown in Fig.3.

P-E speed was continuously changed during forming process of boiling structure. The value of P-E speed was maximum when process was beginning. Along with P-E tool was closed to the center of the work piece gradually, the P-E speed reduced synchronous. The relations of the d, f and n must be calculated as following:

$$V_{P-E} = \frac{n\pi\sqrt{(\frac{f}{n})^2 + (2d - \frac{fnt}{15})^2}}{60} \tag{1}$$

Where $V_{P\text{-}E}$ is processing speed, t is time, n is rotational speed, f is amount of feed.

2.2.2 Forming of radial grooves

Radial grooves were processed by punch method based on the circumferential grooves. Firstly, workpiece was fixed on the plate, and installation direction of stamping tool was parallel with axial direction of workpiece. Secondly, stamping tool moved downward to punch the workpiece to achieve a radial micro-groove after stamping depth a_c was confirmed. Thirdly, stamping tool moved upward while the angle θ_c was rotated by plate, and another radial micro-groove was achieved when stamping tool punched the workpiece again. So, the angle between the two radial grooves was equal to the angle of the plate rotation, and one end of the two grooves was coincidence on the center of workpiece.

When stamping depth a_c was 0.3mm and angle θ_c was 2°, 3D boiling structure with micro-grooves and fins on the surface which could improve the boiling performance of working fluid significantly was obtained after workpiece was processed (shown in Fig.4). The structure interconnected the radial and circumferential grooves and was conducive to flow of liquid from radius to center in evaporating surface.

Interference occur between two adjacent grooves near the center of workpiece in stamping process because there was a certain width of grooves, and the length of interference was relation to the angle θ_c. Defining L_i is the length of interference, so the relations of the L_i and θ_c must be met as following:

$$L_i = \frac{a_c \sin \alpha}{(1 + \cos \alpha)\sqrt{\frac{1 - \cos \theta_c}{2}}} \tag{2}$$

Where a_c is stamping depth, α is blade angle.

2.3 Fabrication of wick

Wick which provided cycle power of working fluid was a sintering layer on internal face of heat sink by metal powder or fiber [12], and there were a promising application due to the porous structure, high porosity and large specific surface area. Experiment was carried out under the protective atmosphere of hydrogen. Wick was fabricated at the temperature range of 900℃-950℃ and sintered time was 45 min. Phase change heat sink was obtained eventually (shown in Fig.5). The porosity of wick can be calculated by the following equation:

$$E = (1 - \frac{M}{\rho V}) \times 100\% \tag{3}$$

Where M is quality of wick, ρ is density of copper powder, V is volume of wick.

3 Testing system of high power LED

Fig.5 The sample of phase change heat sink

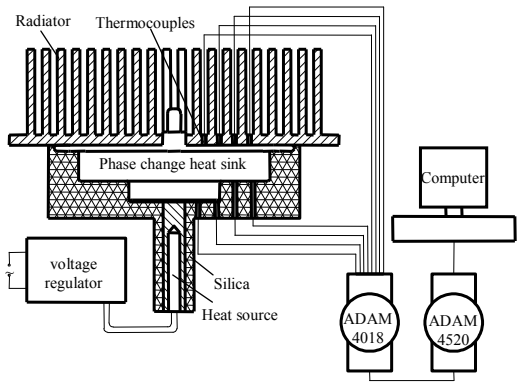


Fig.6 Performance Test System of phase change heat sink

The testing system included heater, cooler and temperature data collection. The solid copper embedded one resistors of 100 Ω was used as heat source. The heat generated from resistors could

be regulated by voltage regulator. To improve the accuracy of experimental data, the solid copper and heat sink were wrapped with asbestos tightly. Meanwhile, an aluminum radiator was installed on condenser for cooling by natural convection in order to improve the reliability of LED. Two contact surfaces was polished and coated with thermal silica and then radiator was fastened on the condenser by mechanically to reduce contact resistance between the radiator and condenser.

The eight thermocouples were distributed on the evaporator (T1, T2, T3, T4) and the condenser (T5, T6, T7, T8). Temperature signal was transmitted to ADAM-4520 from ADAM-4018 which was connected with the eight K type thermocouples, and then to computer whose signal sampling frequency was 1 datum s^{-1}. The testing system and location of thermocouples was shown in Fig.6.

4 Results and discussion

4.1 Effect of power input on the performance

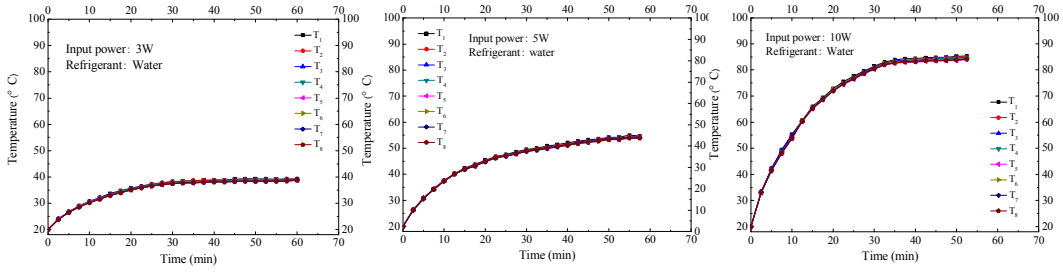
The heat transfer performance of phase change heat sink was tested under the condition of 3W, 5W and 10W as input power respectively and water as refrigerant (shown in Fig.7). The balance time was short, the rise speed of temperature was quick and balance temperature was increased with input power added. The balanced time and the lowest temperature on condenser were 60 min, 48 min, 30 min and 38.6℃, 53.8℃, 84℃, respectively. Less heat was generated under small input power. Meanwhile, some of heat was absorbed by testing system, so the rate of heat exchange was slowly between radiator and air when heat was transferred to radiator by phase change heat sink. After input power was increased, more heat was generated to accelerate the heat exchange between radiator and air, so the balance time was short.

4.2 Effect of refrigerant on the performance

Heat transfer characteristics of heat sink can be reflected by thermal resistance. Thermal resistance can be calculated according to average temperature of evaporator and condenser:

$$R_{ave} = \frac{T_{evp-ave} - T_{cond-ave}}{Q_{in}} \tag{4}$$

Where $T_{evp\text{-}ave}$ and $T_{cond\text{-}ave}$ are average temperatures of evaporator and condenser, respectively.



(a) 3W (b) 5W (c) 10W

Fig.7 Heat transfer performance of phase-change heat sink under different thermal load

Thermal resistance of heat sink was shown as Fig.8 under different refrigerant and input power. Thermal resistance decreased and temperature of evaporator rise when input power increased. Thermal resistance with water as refrigerant was greater than that with alcohol as refrigerant when input power was 3W or 5W, and was equal when input power rise to 7.5W under different refrigerants. The reverse was the case in 10W input power. Because the specific heat capacity and boiling point in vacuum of water were higher than that of alcohol, so more heat should be absorbed by water to boiling and more latent heat was carried under the same conditions. The highest temperature was 85.4℃ and 86.8℃with water and alcohol as refrigerant

respectively under the condition of input power for 10W, ambient temperature 20℃, and the minimum thermal resistance was 0.09℃/W. So the low thermal resistance could be achieved in case of alcohol as refrigerant under the low input power, but water as refrigerant was more suitable for large heat density of heat transfer.

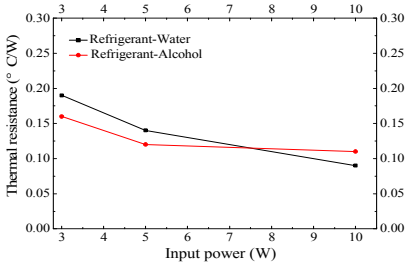


Fig.8 Thermal resistance of phase-change heat sink under different refrigerant

5 Conclusions

(1) A novel phase change heat sink was designed and fabricated for packaging cooling of high power LED in this study. It provided a new way to resolve the thermal problem of packaging level for high power LED.

(2) 3D boiling structures with micro-grooves, cracks and fins on the evaporator surface were processed after spiral and radial micro-grooves were formed by P-E and stamping respectively. Wick was fabricated by sintering method to provide the capillary force of refrigerant cycle.

(3) A good performance of heat transfer was obtained by phase change heat sink had when alcohol as refrigerant under the low input power, or water as refrigerant under the high input power. The minimum thermal resistance and the highest temperature were 0.09℃/W and 86.8℃ under input power of 10W, ambient temperature of 20℃.

Conference

[1] D. A. Steigerwald, J. C. Bhat, D. Collins, R. M. Fletcher, M. O. Holcomb, M. J. Ludowise, P. S. Marin, S. L. Rudaz. IEEE journal on selected topics in Quantum Electronics Vol. 8 (2002), p. 310-320

[2] A. Y. Kim, W. Gotz, D. A. Steigerwald, J. J. Wierer, N. F. Gardner, J. Sun, S. A. Stockman, P. S. Martin, M. R. Krames, R. S. Kern, F. M. Steranka. Physica Status Solidi Vol. 188 (2001), p. 15-21

[3] N. Narendran, Y. Gu, J. P. Freyssinier, H. Yu, L. Deng. Journal of Crystal Growth Vol. 268 (2004), p. 449-456

[4] C. Adam, G. Samuel. Applied Thermal Engineering Vol. 29 (2009), p. 364-371

[5] N. Narendran, Y. Gu. IEEE Journal of Display Technology Vol. 1 (2005), p. 167-171

[6] L. L. Vasiliev. Applied Thermal Engineering Vol. 28 (2008), p. 266-273

[7] S. Liu, J. Yang, Z. Gan, X. Luo. International Journal of Thermal Sciences Vol. 47 (2008), p. 1086-1095

[8] X. Lu, T. C. Hua, M. Liu, Y. Cheng. Thermochimica Acta Vol. 493 (2009), p. 25-29

[9] K. Lan, C. H. Jong, H. J. Sun, W. S. Moo. Thermochimica Acta Vol. 455 (2007), p.21-25

[10] N. D. Nitesh, H. B. Sushil, C. J. Richard. International Journal of Heat and Mass Transfer Vol. 49 (2006), p. 2829-2839

[11] J. H. Xiang, Y. Tang, B. Y. Ye, W. Zhou, H. Yan, Z. H. Hu. Transactions of Nonferrous Metals Society of China Vol. 19 (2009), p. 335-340

[12] Y. M. Z. Ahmed, M. I. Riad, M. K. Ahlam, M. E. H. Shalabi. Powder Technology Vol. 175 (2007), p. 48-54

doi:10.4028/www.scientific.net/AMM.50-51.283

Aggregate Constraint Homotopy Method for Nonlinear Programming on Unbounded Set

Xiao Yu[1,a] , Xiong Huijuan[2,b] and Yan Zhigang[3,c]

[1]School of Basic Science, East China Jiaotong University, Nanchang, Jiangxi 330013, China

[2]Department of Science, Hua-Zhong Agricultural University, Wuhan, Hubei 430070, China

[3]Accountancy College, Dongbei University of Finance and Economics, Dalian, Liaoning 116025

[a]fisherxiao@gmail.com, [b]hj_xiong@mail.hzau.edu.cn, [c]zhigangyan@126.com

Keywords: nonlinear programming; nonconvex programming; unbounded set; aggregate constraint homotopy method.

Abstract. In [1], an aggregate constraint aggregate (ACH) method for nonconvex nonlinear programming problems was presented and global convergence result was obtained when the feasible set is bounded and satisfies a weak normal cone condition with some standard constraint qualifications. In this paper, without assuming the boundedness of feasible set, the global convergence of ACH method is proven under a suitable additional assumption.

Introduction

Consider the following nonlinear programming problem:

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } x \in \Omega = \{x \in R^n \mid g_j(x) \le 0, j \in \overline{q}\}, \end{aligned} \tag{1}$$

where $f, g_j \in C^r (r > 2)$ and $\overline{q} = \{1, \cdots, q\}$. Let $g(x) = \max\{g_1(x), \cdots, g_q(x)\}$. It is well know that if $x^* \in \Omega$ is a solution of (1), then there exists $\xi^* = \{\xi_1^*, \cdots, \xi_q^*\} \in R^q$, such that (x^*, ξ^*) is a solution of the KKT system

$$\begin{aligned} &\nabla f(x) + \sum\nolimits_{j \in \overline{q}} \xi_j \nabla g_j(x) = 0, \\ &\xi_j g_j(x) = 0, \xi_j \ge 0, g_j(x) \le 0. \end{aligned} \tag{2}$$

If (x^*, ξ^*) is a solution of (2), then x^* is called a KKT point of (1), and ξ^* is called the Lagrangian multiplier vector corresponding to x^*.

In [2]--[4], a combined homotopy interior point (CHIP) method was presented for (1). Assuming Ω is bounded, under a normal cone condition of the constraints, it was proven that a smooth interior path from an interior point of the feasible set to a KKT point of (1) exists. In [5], without assuming the boundedness of Ω but adding some suitable additional assumptions, the CHIP method was proven to be still global convergent. In [1], by using the aggregate function, which is induced from the Jayne's maximum entropy principle by Li in [6], an aggregate constraint homotopy (ACH) method was given for (1). When Ω is bounded and satisfies a weak normal cone condition with some standard constraint qualifications, it was proven that the ACH determines a smooth interior path from the given interior point $(x^0, y^0, 1)$ to a KKT point of (1). Because the dimension of the linear systems arising in the process of numerically tracing the ACH path is $n+2$ and not $n+q+1$ as in CHIP methods, the ACH method is more efficient for solving nonlinear programming problems with large amount of constraints.

In this paper, we consider to remove the boundedness of Ω for ACH method. The paper is organized as follows. we give the convergence result of the ACH method for problems (1) on unbounded Ω in Section 2. In Section 3, some test results are given.

Aggregate Constraint Homotopy Method for Nonconvex Nonlinear Problems on Unbounded Sets

Given the following assumptions:

Assumption 1 $f(x)$ and $g_j(x)$ are three times continuously differential functions.

Assumption 2 $\Omega^0 = \{x \in R^n : g(x) < 0\}$ is nonempty (Slater condition).

Assumption 3 For any $x \in \partial\Omega$, $\{\nabla g_j(x) : j \in J(x) = \{j : g_j(x) = g(x)\}\}$ are positively independent, i.e. $\sum_{j\in J(x)} \xi_j \nabla g_j(x) = 0$, $\xi_j \geq 0 \Rightarrow \xi_j = 0$.

Assumption 4 There exists a closed subset $\hat{\Omega} \subset \Omega^0$ with nonempty interior $\hat{\Omega}^0$, such that for any given $x \in \partial\Omega$, $\{x + \sum_{j\in J(x)} \xi_j \nabla g_j(x) : \xi_j \geq 0, \sum_{j\in J(x)} \xi_j > 0\} \bigcap \hat{\Omega} = \varnothing$.

Assumption 5 For any sequence $\{x^k\} \subset \Omega$, $\|x^k\| \to \infty$, there exists $x \in \Omega$ such that for any $\{\xi^k\} \subset R_+^q$, it has $\lim_{k\to\infty}(x - x^k)^T(\nabla f(x^k) + \sum_{j\in\bar{q}} \xi_j^k \nabla g_j(x^k)) < 0$.

For a given constant $\theta \in (0,1]$, denote $\Omega_\theta(t) = \{x \in R^n : g(x, \theta t) \leq 0\}$, $\Omega_\theta^0(t) = \{x \in R^n : g(x, \theta t) < 0\}$. The aggregate homotopy equation was constructed as follows (see[1]):

$$H(x,y,t) = \begin{pmatrix} (1-t)(\nabla_x f(x) + y\nabla_x g(x,\theta t)) + t(x - x^0) \\ yg(x,\theta t) - ty^0 g(x^0,\theta) \end{pmatrix} = 0, \tag{3}$$

where $x^0 \in \hat{\Omega}^0$, $y^0 > 0$, and

$$g(x,t) = t \ln\left(\sum_{j\in\bar{q}} \exp\left(g_j(x)/t\right)\right), \tag{4}$$

is the aggregate function of $g(x)$, and

$$\begin{aligned} &\nabla_x g(x,t) = \sum_{j\in\bar{q}} \lambda_j(x,t) \nabla g_j(x), \\ &\lambda_j(x,t) = \frac{\exp\left(g_j(x)/t\right)}{\sum_{j\in\bar{q}} \exp\left(g_j(x)/t\right)} \in (0,1), \sum_{j\in\bar{q}} \lambda_j(x,t) = 1. \end{aligned} \tag{5}$$

It can be proven that (3) determines a smooth curve starting from the initial point $(x^0, y^0) \in \hat{\Omega}^0 \times R_{++}^1$ and approaching to a point whose x-component is a KKT point of (1). We first give some necessary propositions. The proofs for Propositions 1-4 are similar to that in [1], and omitted here.

Proposition 1 $\lambda_j(x,t)$ in (3) satisfies that

(i) $0 < \lambda_j(x,t)$ and $\sum_{j\in\bar{q}} \lambda_j(x,t) = 1$; (6)

(ii) $\lambda_j(x,t) \to 0$ as $t \to 0+$ and $x \to \bar{x}(\|x\| < \infty)$ for $j \notin J(\bar{x})$.

Proposition 2 If Assumptions 1-2 hold, then

(i) for any $\theta \in [0,1]$ and $t \in (0,1]$ $\Omega_\theta(t) \subset \Omega$;

(ii) for any bounded closed subset $N \subset \Omega^0$, there exists a $\theta \in (0,1]$ such that $N \subset \Omega_\theta^0(1)$.

Proposition 3 If Assumptions 1-3 hold, then there exists a $\theta \in (0,1]$ such that for any $t \in (0,1]$, $\nabla_x g(x, \theta t) \neq 0$ for any $x \in \partial\Omega_\theta(t)$, $\|x\| < +\infty$.

Proposition 4 If Assumptions 1, 2, 4 hold, then for any bounded closed subset $N \subset \hat{\Omega}$, there exists a $\theta \in (0,1]$ such that for any $t \in (0,1]$, $x \in \partial\Omega_\theta(t)$ and $\|x\| < \infty$, it has

$$\left\{x + \sum_{j \in J(x)} \xi_j \nabla g_j(x) : \xi_j \geq 0, \sum_{j \in J(x)} \xi_j > 0\right\} \cap \hat{\Omega} = \varnothing. \tag{7}$$

Theorem 1 Suppose that Assumptions1-5 hold. Then for any $\tilde{x} \in \hat{\Omega}^0$ and its any neighborhood $N(\tilde{x})$ such that $N(\tilde{x}) \subset \hat{\Omega}^0$, there exists $\theta \in (0,1]$, for almost all $x^0 \in N(\tilde{x})$ and $y^0 > 0$, 0 is a regular value of the mapping $H(x, y, t)$ defined as (3) and the set

$$H^{-1}(0) = \left\{(x, y, t) \in \Omega \times R^1_{++} \times (0,1] : H(x, y, t) = 0\right\}$$

contains a smooth curve Γ, which starts from $(x^0, y^0, 1)$ and approaches to the hyperplane $t = 0$,. Moreover, if $(x^*, y^*, 0)$ is a limit point of on the hyperplane $t = 0$, then x^* is a KKT point of (1)

Proof From (i) of Proposition 1, we can easily see that, for any $\tilde{x} \in \hat{\Omega}^0$, there exists a bounded closed neighborhood $N(\tilde{x})$ of $\tilde{x}$ such that $N(\tilde{x}) \subset \hat{\Omega}^0$. By Propositions 1-4, there exists a θ such that (i) $N(\tilde{x}) \subset \Omega_\theta^0(1)$, (ii) for any $x \in \partial\Omega_\theta(t)$, $\|x\| < \infty$, it has $\nabla_x g(x, \theta t) \neq 0$ and (7) holds for $N = N(\tilde{x})$. Using the parameterized Sard theorem, similar to the proof of Lemma 3.1 in [3], we can prove that for almost all $x^0 \in N(\tilde{x})$ and $y^0 > 0$, 0 is a regular value of H. By the implicit functional theorem, the inverse image theorem (see [7]) and the fact that $H(x^0, y^0, 1) = 0$, we know that $H^{-1}(0)$ contains a smooth curve Γ, which starts from $(x^0, y^0, 1)$. And it will return to $(x^0, y^0, 1)$ terminates in or approaches to the boundary of $\Omega \times R^1_+ \times [0,1]$ or goes to infinity. In fact, x-component of Γ is bounded. If not, there exists sequence $\left\{(x^k, y^k, t_k)\right\} \subset \Gamma$, i.e.,

$$(1 - t_k)(\nabla_x f(x^k) + y^k \nabla_x g(x^k, \theta t_k)) + t_k(x^k - x^0) = 0,\ y^k g(x^k, \theta t_k) - t_k y^0 g(x^0, \theta) = 0, \tag{8}$$

such that $\|x^k\| \to \infty$. Then from the first equation in (8) we obtain that for any given $x \in \Omega$,

$$(x - x^k)^T (\nabla f(x^k) + y^k \nabla g(x^k, \theta t_k)) = \frac{t_k}{1 - t_k}(x^k - x)^T (x^k - x^0). \tag{9}$$

From (5), the left hand of (9) is equal to

$$(x - x^k)^T (\nabla f(x^k) + y^k \sum_{j \in \bar{q}} \lambda_j(x^k, \theta t_k) \nabla g_j(x^k)), \tag{10}$$

where $\lambda_j(x^k, \theta t_k) \in (0,1]$, $\sum_{j \in \bar{q}} \lambda_j(x^k, \theta t_k) = 1$. When $\|x^k\| \to \infty$, it has $(x^k - x)^T (x^k - x^0) \to +\infty$, and hence (10) is bigger than 0 as $k \to +\infty$. Hence there exists a subsequence of $\left\{x^k\right\}$ such that (10) has a nonnegative limit (finite or infinite) which contradicts Assumption 5. Hence x-component of Γ is bounded.

Since

$$\left[\frac{\partial H}{\partial x}, \frac{\partial H}{\partial y}\right]_{(x^0,y^0,1)} = \begin{bmatrix} I & 0 \\ y^0 \nabla_x g(x^0,\theta) & g(x^0,\theta) \end{bmatrix},$$

is nonsingular, we know that (x^0, y^0) is not a multiple solution of $H(x, y, 1) = 0$. Therefore, Γ can not return to $(x^0, y^0, 1)$. Let $(\bar{x}, \bar{y}, \bar{t})$ be any extreme point of Γ other than $(x^0, y^0, 1)$. Then similar to the proof of Theorem 2.9 in [1], we can prove that $(\bar{x}, \bar{y}, \bar{t}) \in \Omega \times R_+^1 \times \{0\}$ and $\bar{y} < +\infty$, that is, Γ must terminate in or approach to the hyperplane at $t = 0$.

By the boundedness of x, y for $H(x, y, t) = 0$ and $t \in (0,1]$ and $\lambda_j(x, \theta t) \in (0,1]$, where $\lambda_j(x, \theta t)$ is defined as (5), we know $\left(x, y, \lambda_1(x, \theta t), \cdots, \lambda_q(x, \theta t)\right)$ (where $H(x, y, t) = 0$), has at least one accumulation point as $t \to 0+$. Let $\left(x^*, y^*, \lambda_1^*, \cdots, \lambda_q^*\right)$ be an accumulation point of $\left(x, y, \lambda_1(x, \theta t), \cdots, \lambda_q(x, \theta t)\right)$ and $\xi_j^* = y^* \lambda_j^*$, by (3) we have

$$\nabla f(x^*) + \sum_{j \in \bar{q}} \xi_j^* \nabla g_j(x^*) = 0, \; y^* g(x^*) = 0. \tag{11}$$

From (ii) of Proposition 1, it has $\lambda_j(x^*) = 0$ for $j \notin J(x^*)$, and hence $\nabla f(x^*) + \sum_{j \in J(x^*)} \xi_j^* \nabla g_j(x^*) = 0$, $\xi_j^* g_j(x^*) = 0$ for $j \notin J(x^*)$. For $j \in J(x^*)$, from the second equation of (11), it also has $\xi_j^* g_j(x^*) = 0$. Thus $\xi_j^* g_j(x^*) = 0$ for any $j \in \bar{q}$.

Summing up, $\left(x^*, y^*, \lambda_1^*, \cdots, \lambda_q^*\right)$ is a solution of (2), which means that x^* is a KKT point of (1).

Numerical Test

In this section, we give some numerical results to trace the homotopy curve Γ using the predictor-corrector algorithm (see [8] for detail). The example 1 is from [9] and example 2 is artificial. Both of their feasible sets are unbounded, and the first example satisfies assumptions 1--5, example 2 does not satisfy assumption 5 at least. Parameter $\theta = 1$ for example 1 and $\theta = 0.5$ for example 2. All the computations are done by running MATLAB 7.6.0 on a laptop with AMD Turion (tm) 64×2 Mobile Technology TL-58 CPU 1.9 GHz and 896M memory. In the following table, x^* denotes the final approximate solution point, f^* is the value of the objective function at x^*, g^* is the constraint value, N_{iter} is the corrector times, and time is the CPU time in seconds.

Example 1 ([9], Problem 4) Let $x = (x_1, x_2)^T \in R^2$,

$$\begin{aligned} \min \quad & \frac{1}{3}(x_1 + 1)^3 + x_2, \\ \text{s.t.} \quad & 1 - x_1 \le 0, \\ & -x_2 \le 0. \end{aligned}$$

Example 2 Let $x = (x_1, x_2, \cdots, x_{50})^T \in R^{50}$,

$$\min \quad \sum_{1\le k\le n} (x_k - 1)^2 / 50,$$

$$\text{s.t.} \quad \prod_{1\le k\le n} \cos(s_j x_k) + s_j \sum_{1\le k\le n} x_k^3 \le 0,\ j = 1,\cdots,100,$$

where $s_j = 0.5 + (\pi - 0.5)(j-1)/99$.

Table 1 The numerical results

example	x^*	f^*	g^*	N_{iter}	time
1	(1.000000,0.000000)	2.666666	-2.033e-9	22	0.148
2	(-0.287666, -0.287666,⋯, -0.287666)	1.658084	-7.565e-7	43	3.656

$x^0 = (3,1)^T$ for example 1 and $x^0 = -(1,1,\cdots,1)^T$ for example 2

References

[1] B. Yu, G.C. Feng and S.L. Zhang, The aggregate constraint homotopy method for nonconvex nonlinear programming, Nonlinear Analysis TMA 45 (2001) 839--847.

[2] G.C. Feng, B. Yu, Combined homotopy interior point method for nonlinear programming problems, in: H. Fujita, M. Yamaguti (Eds.), Advances in Numerial Mathematics; Proceedings of The Second Japan-China Seminar on Numerical Mathematics, Lecture Notes in Numerical and Applied Analysis, Vol. 14, Kinokuniya, Tokyo, 1995, pp. 9--16.

[3]G.C. Feng, Z.H. Lin and B. Yu, Existence of an interior pathway to a Karush-Kuhn-Tucker point of a nonconvex programming problem, Nonlinear Anal. 32 (1998) 761--768.

[4] Z.H. Lin, B. Yu and G.C. Feng, A combined homotopy interior point method for convex nonlinear programming, Appl. Math. Comput. 84 (1997) 193--211.

[5] Q. Xu, B. Yu, Solving the Karush-Kuhn-Tucker system of nonlinear programming problem on an unbounded set, Nonlinear Analysis TMA 70 (2009) 757--763.

[6] X.S. Li, An aggregate function method for nonlinear programming, J. Science In China A 34 (1991) 1467--1473.

[7] G.L. Naber, Topological Methods in Euclidean Space, Cambridge University Press, London, 1980.

[8] E.L. Allgower, K. Georg, Numerical Path Following, in: P. G. Ciarlet and J. L. Lions (Eds.), Handbook of Numerical Analysis, North-Holland, 1997, Vol. 5, 3--207.

[9] W.Hock, K. Schittkowski, Test examples for nonlinear programming codes, Springer, Berlin, 1981.

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.288

Some Properties of BSDEs Driven by A Simple Lévy Process with Continuous Coeffcient

Shiqiu Zheng [a], Dianchuan Jin, Shuai Zhang, Yanmei Yang, Jinpeng Wang

College of Science, Hebei United University, Tangshan 063009, China

[a] shiqiumath@163.com

Keywords: Backward stochastic differential equation; Lévy process; comparison theorem

Abstract: In this paper, we mainly study the properties of solutions of backward stochastic differential equations (BSDEs) driven by a simple Lévy process, whose coefficient coeffcient is continuous with linear growth. A comparison theorem for solutions of the equations are obtained, we also show the equation has either one or uncountably many solutions.

Introduction

It is well-known that Pardoux and Peng [1] proves the existence and uniqueness of solutions of backward stochastic differential equations (BSDEs for short), when the coeffcient g satisfies Lipschitz condition. After that, the theory of BSDEs has been developed rapidly because of its connections with a wide of fields, such as PDE, mathematical finance and stochastic control problems and so on. Many papers have been published devoted to weaken the Lipschitz condition on the coeffcient g, Lepeltier and San Martin [2] have proved that there exists at least one solution, when the coeffcient is continuous with linear growth. More recently, under the similar condition restricted on the coeffcient g as [2], El Otmani [3] show the existence of solution of BSDE driven by a simple Lévy process (BSDESL in short) with continuous coeffcient. Comparison theorem of one-dimensional BSDEs is an important and powerful technique in the theory of BSDEs (see Peng [4] or El Karoui, Peng and Quenez M.C.[5]). El Otmani [3] establishes a comparison theorem BSDESLs where the coeffcients g satisfy Lipschitz condition. The main results in this paper is that, we generalize the comparison theorems for BSDESLs to the case where the coeffcients are only continuous with linear growth, we also show the BSDESL has either one or uncountably many solutions.

Preliminaries

Let (Ω, F, P), be a complete probability space, $T > 0$ be a given real number, $L = (L_t)_{0\le t\le T}$ be a simple Lévy process defined by

$$L_t = \mu t + \sigma W_t + \sum_{i=1}^{n} \alpha_i \tilde{N}_t^{(i)},$$

where W_t is a standard Brownian motion in R and $N_t^{(i)}$ is a family of independent Poisson processes with parameters $\left(\lambda^{(i)}\right)_{i=1}^{n}$ such that the process

$$\left(N_t^{(i)}\right)_{t\in[0,T]} = \left(\lambda^{(i)}\right)^{\frac{1}{2}} (N_t^{(i)} - \lambda^{(i)} t)$$

is a martingale for all i. Note that μ, σ and $\left(\alpha^{(i)}\right)_{i=1}^{n}$ are real constants. Let $(F_t^0)_{t\ge 0}$ denote the natural filtration generated by $(L_t)_{t\ge 0}$. Let $(F_t)_{t\ge 0}$ be the completed filtration of $(F_t^0)_{t\ge 0}$ with the P-null sets of F, then it satisfies the usual hypothesis.

For any positive integer n, if $z \in R^n$, let $\|z\|$ denote its Euclidean norm, for $u = \left(u^{(i)}\right)_{i=1}^{n}$ and $v = \left(v^{(i)}\right)_{i=1}^{n} \in R^n$, we denote $u \ge v$ if $u(i) \ge v(i)$, for all $i = 1, \cdots, n$. We define the following usual spaces:

Let $L^2(F_t)$ denote the space of F_t-measurable random variable ξ such that

$$\|\xi\|^2 = E|\xi|^2 < \infty;$$

Let $S^2(R)$ denote the space of right continuous with left limit predictable process ψ such that

$$\|\psi\|_{S^2}^2 = E\left[\sup_{0\le t\le T}|\psi|^2\right] < \infty;$$

Let $H^2\left(R^n\right)$ denote the space of predictable process ψ such that

$$\|\psi\|_{H^2}^2 = E\left[\int_0^T \|\psi\|^2\, dt\right] < \infty.$$

Let us consider a function $g:\Omega\times[0,T]\times R\times R\times R^n \to R$, such that $\left(g\left(t,y,z,u\right)\right)_{0\le t\le T}$ is progressively measurable for each $\left(y,z,u\right)\in R\times R\times R^n$, For the function g, in this paper, we will use the following assumptions:

(A1). There exists a constant $K>0$, such that $dP\times dt-a.s.$, $\forall\left(y_i,z_i,u_i\right)\in R\times R\times R^n$, $i=1,2$,

$$\left|g\left(t,y_1,z_1,u_1\right)-g\left(t,y_2,z_2,u_1\right)\right| \le k\left(\left|y_1-y_2\right|+\left|z_1-z_2\right|+\left\|u_1-u_2\right\|\right).$$

(A2). The process $\left(g\left(t,0,0,0\right)\right)_{t\in[0,T]}\in H^2\left(R\right)$.

(A3). There exists a constant $K>0$, such that $dP\times dt-a.s.$, $\forall\left(y,z,u\right)\in R\times R\times R^n$, $i=1,2$,

$$\left|g\left(t,y,z,u\right)\right| \le k\left(1+\left|y\right|+\left|z\right|+\left\|u\right\|\right).$$

(A4). $P-a.s.$, $\forall t\in[0,T]$, $g(t,y,z,u)$ is continuous in (y,z,u).

(A5). If $u\ge v$ then $P-a.s.$, $\forall t\in[0,T]$, $g(t,y,z,u)\ge g(t,y,z,v)$.

From Lepeltier, San Martin [2] or Liu, Ren [6], we can get the following lemma 2.1.

Lemma 2.1 Let g satisfies the assumptions (A3) and (A4), for $\forall n\ge K$, we consider the sequence $\left(g_n^1\right)_{n\ge K}$ and $\left(g_n^2\right)_{n\ge K}$ defined by

$$g_n^1(y,z,u) = \inf_{(y',z',u')\in Q\times Q\times Q^d}\left\{g(y',z',u')+n\left|y-y'\right|+n\left|z-z'\right|+n\left\|u-u'\right\|\right\},$$

$$g_n^2(y,z,u) = -\inf_{(y',z',u')\in Q\times Q\times Q^d}\left\{-g(y',z',u')+n\left|y-y'\right|+n\left|z-z'\right|+n\left\|u-u'\right\|\right\}.$$

respectively, then we have

(*i*) $\left|g_n^i\left(t,y,z,u\right)\right| \le K\left(1+\left|y\right|+\left|z\right|+\left\|u\right\|\right)$, $i=1,2$;

(*ii*) $g_n^1\left(t,y,z,u\right)\nearrow$, $g_n^2\left(t,y,z,u\right)\searrow$, as $n\to\infty$;

(*iii*) $\left|g_n^i\left(t,y_1,z_1,u_1\right)-g_n^i\left(t,y_2,z_2,u_1\right)\right| \le K\left(\left|y_1-y_2\right|+\left|z_1-z_2\right|+\left\|u_1-u_2\right\|\right)$, $i=1,2$;

(*iv*) If $(y_n,z_n,u_n)\xrightarrow{n\to\infty}(y,z,u)$, then $g_n^i(y_n,z_n,u_n)\xrightarrow{n\to\infty}g(y,z,u)$, $i=1,2$.

Main results

We consider the following BSDESL:

$$y_t = \xi+\int_t^T g\left(s,y_s,z_s,u_s\right)ds-\int_t^T z_s dW_s-\sum_{i=1}^n\int_t^T u_s^{(i)}d\tilde{N}_s^{(i)}, \quad \forall t\in[0,T], \tag{1}$$

where g and (ξ,g,T) are called coeffcient and standard parameters, respectively. For convenient, we denote (1) by BSDESL(ξ,g,T). When coeffcient g satisfies the assumptions $(A1)$ and $(A2)$, for any $\xi\in L^2\left(F_T\right)$, El Otmani [3, Theorem 3.1] showed that BSDESL(ξ,g,T) has a unique solution $\left(y,z,u\right)$ belonging to $S^2(R)\times H^2(R^{1+n})$.

Remark 1. Under (A3) (A4) and (A5), El Otmani [3, Theorem 4.1] shows BSDESL(ξ,g,T) has a minimal solution $(\underline{y},\underline{z},\underline{u})$ in the sense that for any solution (y,z,u) of BSDESL(ξ,g,T), we have $\underline{y}\le y$, for all $t\in[0,T]$, since the solution is obtained through approximating the coeffcient g by the sequence $(g_n^1)_{n\ge K}$ constructed in lemma 2.1 from below.

In fact, if we approximate the coeffcient g by the sequence $(g_n^2)_{n\ge K}$ constructed in lemma 2.1 from above, by the similar method, we can show that BSDESL(ξ,g,T) have a maximal solution $(\overline{y},\overline{z},\overline{u})$ in the sense that for any solution (y,z,u) of BSDESL(ξ,g,T), we have $y\le\overline{y}$, for all $t\in[0,T]$.

We now establish a comparison theorem for the BSDESLs with continuous coeffcient.

Theorem 3.1 Suppose that $\xi_i\in L^2(F_T)$ and g_i satisfies the assumptions (A3)-(A5), let $(\underline{y}_t^i,\underline{z}_t^i,\underline{u}_t^i)$ and $(\overline{y}_t^i,\overline{z}_t^i,\overline{u}_t^i)$ be the minimal and maximal solution of the BSDESL (ξ_i,g_i,T), $i=1,2$. If $\xi_1\le\xi_2$, $g_1\le g_2$, then for all $t\in[0,T]$, we have

$$\underline{y}_t^1\le\underline{y}_t^2,\quad \overline{y}_t^1\le\overline{y}_t^2,\quad a.s.$$

Proof. For all $t\in[0,T]$, let $g_n^1(t,y,z,u)$ be the sequence constructed in Lemma 2.1. By lemma 2.1 and El Otmani [3, Theorem 3.1], we know that the following BSDESLs have a unique adapted solution $(y_t^{1,n},z_t^{1,n},u_t^{1,n})$

$$y_t^{1,n}=\xi+\int_t^T g\left(s,y_s^{1,n},z_s^{1,n},u_s^{1,n}\right)ds-\int_t^T z_s^{1,n}dW_s-\sum\nolimits_{i=1}^n\int_t^T u_s^{1,n,(i)}d\tilde{N}_s^{(i)},\quad\forall t\in[0,T],$$

By lemma 2.1 again, we have $g_n^1\le g_1\le g_2$, then using El Otmani [3, Theorem 3.2], we have $y_t^{1,n}\le y_t^1$ and $y_t^{1,n}\le y_t^2$. Since $y_t^{1,n}\to\underline{y}_t^1$, as $n\to\infty$,, we have $\underline{y}_t^1\le y_t^2$, then $\underline{y}_t^1\le\underline{y}_t^2$, $a.s.$

For all $t\in[0,T]$, let $g_n^2(t,y,z,u)$ be the sequence c onstructed in Lemma 2.1. By lemma 2.1 and El Otmani [3, Theorem 3.1], the following BSDESLs have a unique adapted solution $(y_t^{2,n},z_t^{2,n},u_t^{2,n})$

$$y_t^{2,n}=\xi+\int_t^T g\left(s,y_s^{2,n},z_s^{2,n},u_s^{2,n}\right)ds-\int_t^T z_s^{2,n}dW_s-\sum\nolimits_{i=1}^n\int_t^T u_s^{2,n,(i)}d\tilde{N}_s^{(i)},\quad\forall t\in[0,T],$$

By lemma 2.1 again, we have $g_1\le g_2\le g_n^2$, then using El Otmani [3, Theorem 3.2], we have $y_t^2\le y_t^{2,n}$ and $y_t^1\le y_t^{2,n}$. Since $\overline{y}_t^{2,n}\to\overline{y}_t^2$, as $n\to\infty$, we have $y_t^1\le\overline{y}_t^2$, then $\overline{y}_t^1\le\overline{y}_t^2$, $a.s.$ The proof is completed.

Corllary 3.2 Under the condition of Theorem 3.1, we have the following two statements:

(i) If BSDESL(ξ_1,g_1,T) have a unique solution (y_t^1,z_t^1,u_t^1), then for any solution (y_t^2,z_t^2,u_t^2) of BSDESL(ξ_2,g_2,T), we have for all $t\in[0,T]$, $y_t^1\le y_t^2$, $a.s.$

(ii) If BSDESL(ξ_2,g_2,T), have a unique solution (y_t^2,z_t^2,u_t^2) then for any solution (y_t^1,z_t^1,u_t^1) of BSDESL(ξ_1,g_1,T), we have for all $t\in[0,T]$, $y_t^1\le y_t^2$, $a.s.$

Theorem 3.3 Suppose that $\xi\in L^2(F_T)$, and g satisfies the assumptions (A3)-(A5). Let $(\underline{y}_t,\underline{z}_t,\underline{u}_t)$ and $(\overline{y}_t,\overline{z}_t,\overline{u}_t)$ be the minimal and maximal solution of the BSDESL(ξ,g,T) respectively. For all $t_0\in[0,T]$, and $\eta\in L^2(F_{t_0})$, if $\underline{y}_t\le\eta\le\overline{y}_t$, $a.s.$, then there exist at least a solution (y_t,z_t,u_t) of BSDESL(ξ,g,T) such that

$$y_{t_0} = \eta, a.s.$$

Proof. We prove the theorem using the method from Jia and Peng [7]. Let $\left(y'_t, z'_t, u'_t\right)$ be a solution of the following BSDESL(η, g, T)

$$y'_t = \eta + \int_t^{t_0} g\left(s, y'_s, z'_s, u'_s\right)ds - \int_t^{t_0} z'_s dW_s - \sum_{i=1}^{n}\int_t^{t_0} u'^{(i)}_s d\tilde{N}^{(i)}_s, \quad \forall t \in [0, t_0],$$

For any fixed $(\breve{z}, \breve{u}) \in H^2(R^{1+n})$, we consider the following SDE :

$$\breve{y}_t = \eta - \int_{t_0}^{t} g\left(s, \breve{y}_s, \breve{z}_s, \breve{u}_s\right)ds + \int_{t_0}^{t} \breve{z}_s dW_s + \sum_{i=1}^{n}\int_{t_0}^{t} \breve{u}^{(i)}_s d\tilde{N}^{(i)}_s.$$

By the proof of El Otmani [3, Theorem 3.1] and Tang [8, Theorem 2.1], the SDE has a unique solution $\breve{y}_t$. Then we define the following stopping time

$$\tau = \inf\{t \geq t_0; \breve{y}_t \geq \overline{y}_t, \breve{y}_t \leq \underline{y}_t\} \wedge T,$$

For all $t_0 \in [0, T]$, we set

$$y_t = 1_{[0,t_0)} y'_t + 1_{[t_0,\tau)} \breve{y}_t + 1_{[\tau,T]} \overline{y}_t 1_{\{y_\tau > \underline{y}_\tau\}} + 1_{[\tau,T]} \underline{y}_t 1_{\{y_\tau = \underline{y}_\tau\}};$$
$$z_t = 1_{[0,t_0)} z'_t + 1_{[t_0,\tau)} \breve{z}_t + 1_{[\tau,T]} \overline{z}_t 1_{\{y_\tau > \underline{y}_\tau\}} + 1_{[\tau,T]} \underline{z}_t 1_{\{y_\tau = \underline{y}_\tau\}};$$
$$u_t = 1_{[0,t_0)} u'_t + 1_{[t_0,\tau)} \breve{u}_t + 1_{[\tau,T]} \overline{u}_t 1_{\{y_\tau > \underline{y}_\tau\}} + 1_{[\tau,T]} \underline{u}_t 1_{\{y_\tau = \underline{y}_\tau\}};$$

Then we can check that $\left(y_t, z_t, u_t\right)$ is a solution of BSDESL(ξ, g, T) such that $y_{t_0} = \eta, a.s.$ The proof is completed.

By the proof method of Theorem 3.2, we can get the following Corllary 3.4.

Corllary 3.4 Suppose that $\xi \in L^2\left(F_T\right)$, and g satisfies the assumptions (A3)-(A5). Let $\left(\underline{y}_t, \underline{z}_t, \underline{u}_t\right)$ and $\left(\overline{y}_t, \overline{z}_t, \overline{u}_t\right)$ be the minimal and maximal solution of the BSDESL(ξ, g, T) respectively. For $t_1, t_2 \in [0, T], t_1 < t_2, (z', u') \in H^2(R^{1+n})$, if x_t is a process such that

$$x_t = x_{t_2} + \int_t^{t_2} g\left(s, x_s, z'_s, u'_s\right)ds - \int_t^{t_2} z'_s dW_s - \sum_{i=1}^{n}\int_t^{t_2} u'^{(i)}_s d\tilde{N}^{(i)}_s, \quad t \in [t_1, t_2],$$

and

$$\underline{y}_t \leq x_t \leq \overline{y}_t, \ a.s., \quad t \in [t_1, t_2],$$

then there exist at least a solution $\left(y_t, z_t, u_t\right)$ is a solution of BSDESL(ξ, g, T), such that

$$y_t = x_t, \ a.s., \quad t \in [t_1, t_2].$$

Remark 2. Under the condition of theorem 3.1, if the solution of BSDESL (ξ_1, g_1, T) and BSDESL(ξ_2, g_2, T) are both not unique, then by theorem 3.3, a comparison theorem for all solutions of BSDESL(ξ_1, g_1, T) and BSDESL(ξ_2, g_2, T) like Corllary 3.2 may be not true.

Remark 3. Motivated by Jia and Peng [7, Remark 2], we also get if $\xi \in L^2\left(F_T\right)$, and g satisfies the assumptions (A3)-(A5), then BSDESL(ξ, g, T) has either one or uncountable many solutions, whose cardinality is at least continuum since we can take $\eta = a\underline{y}_{t_0} + (1-a)\overline{y}_{t_0}$ for any $a \in [0,1]$, from theorem 3.2.

Acknowledgments.

The work was supported by the Science Foundation of Hebei Polytechnic University (No.z201018), the work was also supported partically by National Natural Science Foundation of China (No. 30971611) and Natural Science Foundation of Hebei Province (No.A2010000905).

References

[1] E. Pardoux, S. Peng: Adapted Solution of Backward Stochastic Differential Equations. Systems Control Lett. (1990), 14, 51-61

[2] J.P. Lepeltier, J. San Martin: Backward Stochastic Differential Equations with Continuous Coeffcient. Statist. Probab. Lett. (1997), 32(4), 425-430

[3] M. El Otmani: BSDE Driven by a Simple Lévy Process with Continuous Coeffcient. Statistics and Probability Letters. (2008), 78, 1259-1265

[4] S. Peng: A Generalized Dynamic Programming Principle and Hamilton-Jacobi-Bellman Equation. Stochastics Stochastic Rep. (1992), 38(2), 119-134

[5] N. El Karoui, S. Peng, and M.C. Quenz: Backward Stochastic Differential Equations in Finance. Math. Finance (1997), 7, 1-71

[6] J. Liu, J. Ren: Comparison Theorem for Solutions of Backward Stochastic Differential Equations with Continuous Coeffcient. Statist. Probab. lett. (2002), 56, 93-100

[7] G. Jia, S. Peng: On the Set of Solutions of a BSDE with Continuous Coeffcient. C. R. Acad. Sci. Paris, Ser. I, (2007), 344, 395-397

[8] H. Tang: Stochastic linear quadratic control problem with Lévy process and BSDEs with markov chains. Ph.D Thesis. Shandong University, China.(2008)

doi:10.4028/www.scientific.net/AMM.50-51.293

Infinite Time Interval BSDEs Driven by A Lévy Process

Shiqiu Zheng [a], Aimin Yang, Dianxuan Gong, Qiumei Liu, Yamian Peng

College of Science, Hebei United University, Tangshan 063009, China

[a] shiqiumath@163.com

Keywords: Backward stochastic differential equation; Lévy process; Teugels martingale

Abstract: In this paper, we study the infinite time interval backward stochastic differential equations (BSDEs) driven by a Lévy process. A existence and uniqueness theorem for solution of the BSDEs is established, which can be considered a generalization of existence and uniqueness theorem of BSDEs. A continuous dependence theorem for solutions of the BSDEs is also given.

Introduction

It is well-known that Pardoux and Peng [1] proves the existence and uniqueness of solutions of backward stochastic differential equations (BSDEs for short), when the coeffcient g satisfies Lipschitz condition. After that, the theory of BSDEs has been developed rapidly because of its connections with a wide of fields, such as PDE, mathematical finance and stochastic control problems and so on. Chen and Wang [2] generalized the result of Pardoux and Peng (1990), firstly give a condition on coeffcient of infinite time interval BSDEs under which the infinite time interval BSDEs have a unique solution. Nualart and Schoutens [3] generalized the result of Pardoux and Peng [1], established the existence and uniqueness of solutions for BSDEs with Lipschitz coeffcients driven by Lévy process of the kind considered in Nualart and Schoutens[4]. After that, this kind of BSDE is study many papers such as [5] and [6] and so on. Motivated by Chen and Wang [2] and Nualart, Schoutens [3], in this paper, we consider infinite time interval BSDEs driven by a Lévy process introduced by Nualart and Schoutens [4], a existence and uniqueness theorem and a continuous dependence theorem for solution of this kind of BSDEs are established.

Preliminaries

Let $(X_t)_{t\geq 0}$ be a Levy process defined on a complete probability space (Ω, F, P), corresponding to a standard Levy measure υ such that $\int_R (1\wedge y^2)\upsilon(dy) < \infty$ and $\int_{(-\varepsilon,\varepsilon)} e^{\lambda|y|}\upsilon(dy) < \infty$ for every $\varepsilon > 0$ and for some $\lambda > 0$. Let $(F_t^0)_{t\geq 0}$ denote the natural filtration generated by $(X_t)_{t\geq 0}$. Let $(F_t)_{t\geq 0}$ be the completed filtration of $(F_t^0)_{t\geq 0}$ with the P-null sets of F and $F = F_\infty = \vee_{t\geq 0}(F_t)$.

Let $L^2(F_t)$ denote the space of F_t-measurable random variable ξ such that

$$\|\xi\|^2 = E|\xi|^2 < \infty;$$

Let $S^2(R)$ denote the space of right continuous with left limit predictable process ψ such that

$$\|\psi\|_{S^2}^2 = E\left[\sup_{0\leq t\leq\infty}|\psi|^2\right] < \infty;$$

Let l^2 denote the space of real valued sequences $(x_i)_{i\geq 1}$ such that $\sum_{i=1}^{\infty}|x_i| < \infty$, $H^2(l^2)$ denote the space of l^2-valued predictable process ψ such that

$$\|\psi\|_{H^2}^2 = E\left[\int_0^\infty \sum_{i=1}^{\infty}\left|\psi_t^{(i)}\right|^2 dt\right] < \infty.$$

We denote by $(Y^{(i)})_{i\geq 1}$ the Teugels martingales associated to the Lévy process $(x_i)_{i\geq 1}$. An orthonormalization procedure applied to the martingales $(Y^{(i)})_{i\geq 1}$ can obtain a set of pairwise strongly orthonormal martingales $(H^{(i)})_{i\geq 1}$ such that each $Y^{(i)}$ is a linear combination of the $Y^{(j)}$, $j=1,\cdots,i$, we refer to Nualart and Schoutens [3,4] for details.

From Nualart and Schoutens [4,Theorem 2], we get the following Predictable Representation Property:

Lemma 2.1 Every square integrable random variable $F\in L^2(F)$ has a representation of the form

$$F=E[F]+\sum_{i=1}^{\infty}\int_0^{\infty}\psi_t^{(i)}dH_t^{(i)}.$$

Define a function $g:\Omega\times[0,T]\times R\times l^2\to R$, such that $\left(g(t,y,z)\right)_{0\leq t\leq T}$ is progressively measurable for each $(y,z)\in R\times l^2$, . For the function g, in this paper, we will use the following assumptions:

(A1). There exist two positive non-random functions $u(t)$ and $v(t)$ such that $dP\times dt-a.s.$, $\forall(y_i,z_i)\in R\times R^n$, $i=1,2$,

$$\left|g(t,y_1,z_1)-g(t,y_2,z_2)\right|\leq u(t)\left|y_1-y_2\right|+v(t)\left|z_1-z_2\right|.$$

where $u(t)$ and $v(t)$ satisfying $\int_0^{\infty}v(t)dt<\infty$ and $\int_0^{\infty}u(t)^2dt<\infty$.

(A2). The process $E\left(\int_0^{\infty}g(t,0,0)dt\right)^2<+\infty$.

In the paper, we will consider the following infnite time interval BSDE:

$$y_t=\xi+\int_t^{\infty}g(s,y_{s-},z_s)ds-\sum_{i=1}^{\infty}\int_t^{\infty}z_s^{(i)}dH_s^{(i)},\quad\forall t\in[0,\infty],\tag{1}$$

Defnition 2.1 A solution of BSDE(1) is a pair (y,z) belonging to $S^2(R)\times H^2(l^2)$ and satisfying (1).

Main results

In this section, we will present existence and uniqueness of solution of BSDE(1).

Theorem 3.1 Suppose that $\xi\in L^2(F)$ and g satisfies the assumptions (A1) and (A2), then BSDE(1) has a unique solution.

Proof. Given $(U,V)\in S^2(R)\times H^2(l^2)$, we set

$$y_t=E[\xi+\int_t^{\infty}g(s,U_{s-},V_s)\,ds|F_t],\quad t\in[0,\infty],\tag{2}$$

then by Doob's inequality, the assumptions (A1) and (A2), we have

$$\begin{aligned}
E\left[\sup_{0\leq t\leq\infty}|y_t|^2\right]&\leq E\left[\sup_{0\leq t\leq\infty}\left(E\left[|\xi|+\int_0^{\infty}\left|g(s,U_{s-},V_s)\right|ds|F_t\right]\right)^2\right]\\
&\leq CE\left[|\xi|+\int_0^{\infty}\left|g(s,U_{s-},V_s)\right|ds\right]^2\\
&\leq CE\left[|\xi|^2+\left(\int_0^{\infty}|g(s,0,0)|ds\right)^2+\left(\int_0^{\infty}u(s)|U_{s-}|ds\right)^2+\left(\int_0^{\infty}v(s)|V|_s\,ds\right)^2\right]\\
&\leq C\left[E|\xi|^2+E\left(\int_0^{\infty}|g(s,0,0)|ds\right)^2+\|U_{s-}\|_{S^2}^2\left(\int_0^{\infty}u(s)ds\right)^2+\|V_s\|_{H^2}^2\left(\int_0^{\infty}v(s)^2ds\right)\right]\\
&<\infty.
\end{aligned}$$

Applying Lemma 2.1 to

$$\xi + \int_0^{\infty} g(s, U_{s-}, V_s)\, ds,$$

we get that there exists a process $z \in H^2(l^2)$ such that

$$\xi + \int_0^{\infty} g(s, U_{s-}, V_s)\, ds = E[\xi + \int_0^{\infty} g(s, U_{s-}, V_s)\, ds] + \sum_{i=1}^{\infty} \int_0^{\infty} z_s^{(i)} dH_s^{(i)}, \tag{3}$$

from (2) and (3) above, we can deduce that

$$y_t = \xi + \int_t^{\infty} g(s, U_{s-}, V_s) ds - \sum_{i=1}^{\infty} \int_t^{\infty} z_s^{(i)} dH_s^{(i)}, \quad \forall t \in [0, \infty],$$

Then we can define a mapping ϕ from $S^2(R) \times H^2(l^2)$ into itself, that is,

$$(y,\ z) := \phi(U, V),$$

where y and z is defined as (2) and (3), respectively. Then we only need show ϕ is a contractive mapping.

We define a norm $\|\cdot\|_{S^2 \times H^2}$ of $S^2(R) \times H^2(l^2)$ as follows:

$$\|(y,\ z)\|^2_{S^2 \times H^2} = E\left[\sup_{0 \le t \le \infty} |\psi_t|^2 + \int_0^{\infty} \sum_{i=1}^{\infty} |\psi_t^{(i)}|^2 dt \right].$$

Let

$$(y^i,\ z^i) := \phi(U^i, V^i), \quad i = 1, 2.$$

We set $\hat{\eta} = \eta^1 - \eta^2$, for any process $\eta^i, i = 1, 2,$ then by Doob's inequality,

$$\begin{aligned} E\left[\sup_{0 \le t \le \infty} |\hat{y}_t|^2 \right] &\le E\left[\sup_{0 \le t \le \infty} \left(E\left[\int_0^{\infty} \left| g(s, U_{s-}^1, V_s^1) - g(s, U_{s-}^2, V_s^2) \right| ds \,\middle|\, F_t \right]^2 \right) \right] \\ &\le 4E\left[\int_0^{\infty} \left| g(s, U_{s-}^1, V_s^1) - g(s, U_{s-}^2, V_s^2) \right| ds \right]^2, \end{aligned} \tag{4}$$

Since

$$-d\hat{y}_t = \left(g(s, U_{s-}^1, V_s^1) - g(s, U_{s-}^2, V_s^2) \right) ds - \sum_{i=1}^{\infty} \int_t^{\infty} \hat{z}_s^{(i)} dH_s^{(i)},$$

$$d[\hat{y}_s, \hat{y}_s] = |\hat{z}_s|^2 ds + \sum_{i=1}^{\infty} \sum_{i=1}^{\infty} \hat{z}_s^{(i)} \hat{z}_s^{(j)} d[H^{(i)}, H^{(j)}]_s,$$

$$< H^{(i)}, H^{(j)} >_t = \delta_{i,j} t.$$

Applying Ito formula to $|\hat{y}_t|^2$ for $s \in [t, \infty]$, we have

$$\begin{aligned} &(\hat{y}_t)^2 + \int_t^{\infty} \sum_{i=1}^{\infty} |\hat{z}_s^{(i)}|^2 ds \\ &= 2\int_t^{\infty} \hat{y}_{s-} \left(g(s, U_{s-}^1, V_s^1) - g(s, U_{s-}^2, V_s^2) \right) ds - 2\int_t^{\infty} \sum_{i=1}^{\infty} \int_0^{\infty} \hat{y}_{s-} \hat{z}_s^{(i)} dH_s^{(i)} \\ &\quad - \sum_{i=1}^{n} \sum_{i=1}^{n} \int_t^{\infty} \hat{z}_s^{(i)} \hat{z}_s^{(j)} d([H^{(i)}, H^{(j)}]_s - < H^{(i)}, H^{(j)} >_s). \end{aligned}$$

Taking expectation in both sides and using the fact that

$$([H^{(i)}, H^{(j)}] - < H^{(i)}, H^{(j)} >)_s, \quad s \in [0, \infty]$$

is a martingale and (4), we can deduce

$$\begin{aligned} &E(\hat{y}_t)^2 + E\int_t^{\infty} \sum_{i=1}^{\infty} |\hat{z}_s^{(i)}|^2 ds \\ &= 2E\int_t^{\infty} \hat{y}_{s-} \left(g(s, U_{s-}^1, V_s^1) - g(s, U_{s-}^2, V_s^2) \right) ds \\ &\le E\left[\sup_{0 \le t \le \infty} |y_t|^2 \right] + E\left[\int_t^{\infty} \left(g(s, U_{s-}^1, V_s^1) - g(s, U_{s-}^2, V_s^2) \right) ds \right]^2 \\ &\le 5E\left[\int_t^{\infty} \left(g(s, U_{s-}^1, V_s^1) - g(s, U_{s-}^2, V_s^2) \right) ds \right]^2. \end{aligned} \tag{5}$$

Indeed, we have

$$
\begin{aligned}
&E\left[\int_t^\infty \left(g\left(s,U_{s-}^1,V_s^1\right)-g\left(s,U_{s-}^2,V_s^2\right)\right)ds\right]^2 \\
&\le \|U_{s-}\|_{S^2}^2\left(\int_0^\infty u(s)ds\right)^2+\|V_s\|_{H^2}^2\left(\int_0^\infty v(s)^2ds\right) \\
&\le 2\left[\left(\int_0^\infty u(s)ds\right)^2+\int_0^\infty v(s)^2ds\right]\left\|(\hat{U},\ \hat{V})\right\|_{S^2\times H^2}^2
\end{aligned}
\tag{6}
$$

Then from (4), (5) and (6), we can deduce that

$$
\left\|(\hat{y},\hat{z})\right\|_{S^2\times H^2}^2 \le 18\left[\left(\int_0^\infty u(s)ds\right)^2+\int_0^\infty v(s)^2ds\right]\left\|(\hat{U},\hat{V})\right\|_{S^2\times H^2}^2.
$$

Case 1: If $\left[\left(\int_0^\infty u(s)ds\right)^2+\int_0^\infty v(s)^2ds\right]<18,$ then ϕ is a contraction mapping, we can complete the proof.

Case 2: If $\left[\left(\int_0^\infty u(s)ds\right)^2+\int_0^\infty v(s)^2ds\right]\ge 18,$ we can find $T>0$ large enough, such that

$$
\left[\left(\int_T^\infty u(s)ds\right)^2+\int_T^\infty v(s)^2ds\right]<18.
$$

Set

$$
\bar{g}(t,y_t,z_t)\ =\ 1_{[T;\infty]}(t)g(t,y_t,z_t),
$$

then $\bar{g}$ satisfies the assumptions (A1) and (A2), by Case 1, we know that

$$
\bar{y}_t=\xi+\int_t^\infty \bar{g}\left(s,\bar{y}_{s-},\bar{z}_s\right)ds-\sum_{i=1}^\infty\int_t^\infty \bar{z}_s^{(i)}dH_s^{(i)},\quad \forall t\in[0,\infty],
$$

has a unique solution $\left(\bar{y},\bar{z}\right)$ Then we consider the following BSDE:

$$
\tilde{y}_t=\bar{y}_T+\int_t^T g\left(s,\tilde{y}_{s-},\tilde{z}_s\right)ds-\sum_{i=1}^\infty\int_t^\infty \tilde{z}_s^{(i)}dH_s^{(i)},\quad \forall t\in[0,T], \tag{7}
$$

by Nualart and Schoutens [3,Theorem 1], we know that BSDE(7) has a unique solution $\left(\tilde{y},\tilde{z}\right)$. We set $\left(\tilde{y},\tilde{z}\right)=(0,0),\ \forall t\in(T,\infty],$

Let

$$
y=\bar{y}+\tilde{y},\quad z=\bar{z}+\tilde{z},
$$

then it is easy to check that $\left(y,z\right)$ is a unique solution of BSDE(1). The proof is completed.

We now give a continuous dependence result for solution of BSDE(1). we can complete its proof by the method of Chen and Wang [2, Theorem 1.3], we omit it here.

Theorem 3.2 Suppose that $\xi^i\in L^2\left(F\right)$ and g satisfies the assumptions (A1)-(A2), let $\left(y_t^i,z_t^i\right)$ be the solution of the BSDE(1) with data $\xi^i,\ i=1,2,$ then

$$
E\left[\sup_{0\le t\le\infty}\left|y_t^1-y_t^2\right|^2\right]+E\left[\int_t^\infty\sum_{i=1}^\infty\left|z_s^{1,(i)}-z_s^{2,(i)}\right|^2ds\right]\le CE\left[\left|\xi^1-\xi^2\right|^2\right],
$$

where C is a constant.

Acknowledgments

The work was supported by the Science Foundation of Hebei Polytechnic University (No.z201018), the work was also supported partically by National Natural Science Foundation of China (No. 30971611) and Natural Science Foundation of Hebei Province (No.A2010000905).

References

[1] E. Pardoux, S. Peng: Adapted solution of backward stochastic differential equations. Systems Control Lett. (1990), 14, 51-61.

[2] Z. Chen, B. Wang: Infinite time interval Backword SDEs and the convergence of *g*-martingales. J. Austral. Math. Soc.(Series A), (2000), 69, 187-211.

[3] D. Nualart, W. Schoutens: Backward stochastic differential equations and Feynman-Kac formula for Lévy processes, with applications in finance. Bernoulli (2001),7(5), 761-776.

[4] D. Nualart, W. Schoutens: Chaotic and predictable representations for Lévy processes. Stochastic Process. Appl. (2000), 90 (1), 109-122.

[5] M. El Otmani: BSDE Driven by a Simple Lévy Process with Continuous Coeffcient. Statistics and Probability Letters. (2008), 78, 1259-1265.

[6] M. El Otmani: Reflected BSDE Driven by a Lévy Process. J. Theor Probab (2009) 22: 601–619

doi:10.4028/www.scientific.net/AMM.50-51.298

Study on Simulation and Optimization of the Road Rush-repair Model after Disaster

Tiezhu Zhang[1,a], Yongmin Lu[1,b]

[1]Zhejiang Gongshang University, Hangzhou , China

[a]zhangtz@mail.zjgsu.edu.cn, [b]lym-8709@163.com

Keywords: disaster relief; road rush-repair; schedule model; genetic algorithm; eM-plant

Abstract. In view of the facts that the road network is damaged after disaster, this study made an intensive study of how to make the road repair scheduling in a short time. Based on the actual situation of disaster, and the emergency repairs at the time of uncertainty, established a fuzzy multi-objective mathematical model for road rush-repair scheduling. This model based on VRP, containing time window, take the sum time of the repair work, the time of the work group in transit and the risk of the rush-repair. In the model solution, combined the simulation with optimization methods, take eM-plant as a platform, design a genetic algorithm method which is more suitable in solving scheduling model of the transport emergency supplies, Through a numerical example of empirical analysis, proves the validity of models and algorithms, provided a scientific foundation for the government to make reasonable rush-repair scheduling when the disasters occur.

Introduction

China is a natural disaster-prone country, disasters not only cause the direct injury of people and property loss, but also cause great damage to the road network so that traffic can't maintain smooth flow in the affected areas. The outside world has not been able to rescue the victims and the victims can not flee to a safe shelter, which cause irreparable secondary loss of life and property. Therefore, whether the network transport system is smooth has a decisive impact in large-scale natural disaster relief after the evacuation, so how the decision-making units take the environment and the efficiency into account, formulate the repair work, including rush-repair work and rescue scheduling order as soon as possible has become a important study of China's disaster relief.

The basic question of the road repair is: how to find out the repair scheduling with the maximum benefit within a short time, and under limited resources. Relevant scholars do their research: Fiedrich et al. had developed a decision support system with the best allocation of resources for search and rescue of post-earthquake[1]. Liao built double planning model road repairs and emergency logistics distribution, the upper layer focused on the problem of the emergency rehabilitation of engineering units road network, the lower problem is revised Hsueh et al. emergency logistics problem[2,3]. It can be seen that most domestic and international scholars regard the repair problem of road network as a combinatorial optimization problem, mostly use travel time, traffic and other information to performance the assessment goal of the scheduling procedure for post-disaster road network rehabilitation, and repair time as a specific value. However, there are many uncertainty elements in road repairs, uncertainty of repair time, strong dynamic, large scale of the problem and other issues, meanwhile there are some difficulties for the previous studies in solving the problem.

In view of the facts that the road network is damaged after disaster, and the emergency repairs

at the time of uncertainty, established a fuzzy multi-objective mathematical model for road rush-repair scheduling. This model based on VRP[4], containing time window, take the sum time of the repair work, the time of the work group in transit and the risk of the rush-repair. In the model solution, combined the simulation with optimization methods, take eM-plant as a platform, design a genetic algorithm method which is more suitable in solving scheduling model of the transport emergency supplies, the genetic algorithm solves the model by using simtalk language programming, it provides a scientific basis for government departments when carrying out repairs scheduling reasonably in the disaster[5].

Construction of Road Repair and Disaster Relief Model

Problem description. Assumptions of this study are as follows:

1) Information of the disaster area has been fully collected, to make repair decision under consideration in the overall.

2) Emergency response units should know saving energy of engineering units, and no producing wrong dispatch and energy shortage.

3) Bridge is also a part of road, regard bridge that is damaged as a disaster point, so for the bridge problem, this research regard it as a high urgency of disaster point.

4) The principle with disaster points identified is that vehicles are completely blocked and impassable, therefore, that vehicles can't pass is regarded as a disaster point. And the identification with repair completed, under the consideration of emergency, is that it can meet single lane access for a small vehicle through repairs.

5) Assumption that the road network in the model is a complete road network with section line connected.

6) In the case of an emergency repair, the road network is no directivity.

7) Each disaster point must be exactly serviced by one team.

8) Assumption that the travel time between various sections of the road network is a set of known values, with no impact of traffic flow changes.

9) Repair time of the disaster point obeys normal distribution, the relevant parameters of normal distribution have been collected in advance.

Construction of post-disaster road repair model. Before construction of mode, first make a note about symbol as follows:

D: expressed as task force set;

S : expressed as disaster point set;

Oij: restriction of pre-conditions about the disaster point i=1…m+n, j=m+1…m+n;

$$xij: xij = \begin{cases} 1, \text{ point i supports point j;} \\ 0, \text{ point i doesn't support point j.} \end{cases}$$

Tij: records cumulative time when point i supports point j, i=1…m+n, j=m+1…m+n;

Tj: records the time when point j is supported, j=m+1…m+n;

m: that there are m task forces, m=1…m;

n: that there are n disaster points, n=m+1…m+n;

dij: expressed as time distance from point i to point j, i=1…m+n, j=m+1…m+n;

Ci : the repair energy of all the task forces, i=1…m;

Dj: the repair time of disaster point j required, j=m+1…m+n;

Pj: the urgency of the disaster point j, j=m+1…m+n;

Rj: the risk value of repair assessment about disaster point j, j=m+1…m+n;

Ki: the number of task forces from project station, i=1…m;

Timej: the latest time when disaster point j should be repaired, j=m+1…m+n;

According to the definition of the problem and assumptions, construct the relief model of road repair are as follows:

OBJ

$$MinZ_1 = \sum_{i=m+1}^{m+n} \sum_{j=m+1}^{m+n} (x_{ij} \times D_j + T_{ij}) \tag{1}$$

$$MinZ_2 = \sum_{j=m+1}^{m+n} \left(P_j \times T_j\right) \tag{2}$$

$$MinZ_3 = \sum_{i=1}^{m} \sum_{j=m+1}^{m+n} \left(T_{ij} \times R_j\right) \tag{3}$$

ST

$$\sum_{j=m+1}^{m+n} x_{ij} = \begin{cases} K_i & (i = 1,2,\ldots, \ m) \\ 1 \end{cases} \tag{4}$$

$$\sum_{i=1}^{m+n} x_{ij} = 1 \quad (j = m+1,\ldots, \ m+n) \tag{5}$$

$$T_{ij} = d_{ij} \times x_{ij} \tag{6}$$

$$\sum_{k=1}^{m+n} T_{ik} - \sum_{i=1}^{m+n} T_{ij} - \sum_{k=1}^{m+n} \left(d_{jk} \times x_{jk}\right) = D_i \tag{7}$$

$$T_j = \sum_{j=1}^{m+n} T_{ij} \quad (j = m+1,\ldots, \ m+n) \tag{8}$$

$$T_j \le Time_j \quad (j = m+1,\ldots, \ m+n) \tag{9}$$

$$\sum_{i=1}^{m+n} \sum_{j=m+1}^{m+n} O_{ij} \ge 1 \qquad i, j \in D \tag{10}$$

$$\sum_{i=1}^{m} C_i - \sum_{j=m+1}^{m+n} T_j \ge 0 \tag{11}$$

$$D_j \sim N(d, r) \tag{12}$$

$$x_{ij} \in \{0,1\} \tag{13}$$

Explain the above mathematical model as follows:

Formula (1): the total time cost is that the travel time of the team to the disaster point plus the repair estimated time of all disaster points, seek its minimum.

Formula (2): the supported time of all disaster points T_j multiply weight of emergency about disaster point P_j, seek its minimum.

Formula (3): risk value of assessment of all disaster points multiply the travel time distance of each teams, seek the minimum of its objectives.

Formula (4): restriction the number of support groups from the workstation to the disaster point as K_k.

Formula (5): any disaster point is just assigned once.

Formula (6): the time T_{ij} from the workstation i to the disaster point j is equal to the distance d_{ij} multiplied by x_{ij}.

Formula (7): the time from point j to the next point is equal to that the time of reaching point j plus the needed time of point j and the distance d_{ij} from point j to the next point k multiplied by x_{ij}.

Formula (8): the time of reaching point j is equal to the sum of time when all of its nodes i arrive at point j.

Formula (9): the time of arriving point j must be less than the latest time when point j is supported

Formula (10): restriction of pre-conditions, that's to say, before some disaster point is assigned, its pre-disaster point should be completed.

Formula (11): the sum of repair energy of all the task forces is greater than repair time of all disaster points.

Formula (12): the repair time of the disaster point obeys normal distribution.

Formula (13): x_{ij}=1, expressed that the path between point i and point j is selected; x_{ij}=0, opposite.

Genetic Algorithm for Road Repair Scheduling Based on eM-plant

With the expansion of the scale of road repair issues, the traditional accuracy solution and some heuristic algorithms can't solve effectively, in this study, the simulation platform based on eM-plant, realize improved genetic algorithm to achieve the purpose of the model solution. The solution steps of model of road repair by genetic algorithm[6] as follows:

1) In the initial population generation, this study randomly generated the entire group solutions, however, some nodes have priority limit, then assess their priorities and restrictions of time window, produce a better initial solution through such cycle[7].

2) However, this study is to consider the adjustment of multiple objective functions, Therefore, the value of maximum of satisfaction among the various groups that is solved by fuzzy multi-objective approach.

3) Operations of genetic algorithms mainly have selection, crossover and mutation, in the operation of selection, this paper used the roulette selection method. In the operation of crossover, use the way of chromosome system within groups to exchange in the way of the whole bit, produce a new offspring by the exchange and combination between the parent and the mother, using two groups to exchange information to find a better solution. In the operation of mutation, in the beginning of the study, set the mutation rate of 1%, and then decreased by 0.1% in each generation to avoid falling into local optimal solution in search process.

4) According to the specific circumstances, termination condition of algorithm is set as follows: the iterative algorithm attains 200 generations, average fitness of some generation chromosome achieve 0.98 times of the best chromosome fitness of this generation. Parameter range: the group size is 20~40, the crossover rate is 0.7~0.9, the mutation rate is 0.05~0.1.

In specific implementations of the algorithm, use object-oriented simulation software eM-plant as a platform, and use genetic algorithm that is achieved by programming language simtalk to solve the model. Specific parameters of genetic algorithm are set as follows: the crossover rate is 0.8, the mutation rate is 0.1, implementation algebra is 200, groups are 20, the way of crossover chooses the roulette method.

Analysis of the Example

This section will use an example to analyze effectiveness of model and algorithm that is previously proposed. Case mainly as follows: there are 25 disaster points within the disaster area, there are four engineering repair teams to be sent emergency repairs, each task force can repair four disaster points, and that's a complete road network between each disaster points, in addition, it has the limits of priority between a part of disaster points.

System input data, including the time distance between disaster points, repair time limit of

disaster points, repair time of disaster points, emergency degree about repair of disaster points, repair risk of disaster points and repair restrictions of disaster points. Among, constraints for the disaster point is disaster point: to repair the disaster point 41, the disaster point 13 must be first repaired; repair time obeys normal distribution, in this study set $D_j \sim N(0,2)$, repair time of the disaster point and other parameters are specifically set in the table below as follows:

TABLE Ⅰ. REPAIR TIME TABLE OF DISASTER POINTS

Disaster Point No.	**1**	**2**	**3**	**4**	**5**
0~	10	9	12	2	6
1~	3	11	6	6	10
2~	5	8	5	6	10
3~	11	7	8	5	11
4~	10	2	9	8	2

TABLE Ⅱ. REPAIR TIME LIMIT TABLE OF DISASTER POINTS

Disaster Point No.	**1**	**2**	**3**	**4**	**5**
0~	100	100	100	100	20
1~	30	30	30	30	30
2~	40	40	40	40	40
3~	50	50	50	50	50
4~	60	60	60	60	60

TABLE Ⅲ. REPAIR EMERGENCY DEGREE TABLE OF DISASTER POINTS

Disaster Point No.	**1**	**2**	**3**	**4**	**5**
0~	1	3	2	2	5
1~	5	7	6	7	4
2~	6	7	8	6	5
3~	3	5	7	8	7
4~	6	5	9	7	8

TABLE Ⅳ. THE VALUE OF REPAIR RISK ASSESSMENT TABLE ABOUT DISASTER POINTS

Disaster Point No.	**1**	**2**	**3**	**4**	**5**
0~	5	2	5	5	3
1~	5	3	3	2	3
2~	5	5	4	2	3
3~	2	1	3	2	1
4~	2	3	1	3	4

The optimization information of road repair schedule of model is shown as follows in Table 6, take the third team for example, analyzing the results solved. As can be seen from the table, the repair order of the third team is 14-45-13-41.

TABLE Ⅴ. THE RESULT TABLE OF ROAD REPAIR SCHEDULE

	First Team				Second Team			
	1	2	3	4	1	2	3	4
Disaster Point No.	35	32	15		34	12	43	33
Travel time	1-5	16-25	32-46		1-8	13-19	30-37	46-50
Repair time	5-16	25-32	46-50		8-13	19-30	37-46	50-58
	Third Team				Fourth Team			
	1	2	3	4	1	2	3	4
Disaster Point No.	14	45	13	41	25	24	44	
Travel time	1-6	12-23	26-32	38-41	1-5	15-21	27-39	
Repair time	6-12	23-26	32-38	41-51	5-15	21-27	39-47	

Summary

This paper mainly studies the genetic algorithm of model of post-disaster road repair and solution. As the post-disaster repair model is different from the general VRP model, therefore, in the repair order, firstly, meet the conditions of repair order limited; secondly, observed that these points all belong to disaster points that have relatively large emergency degree of repairs and relatively low risk of repairs; Finally, repair time and travel time of the task force in each disaster points also meet the time window constraints of the model. So you can believe that the solution of genetic algorithm is effective.

In addition, in dealing with conditions of the priority and restrictions of the time window, this study makes some special design on producing the initial group of the genetic algorithm in order to ensure the initial group to fully meet all restriction conditions. By example simulation, verify effectiveness and practicality of model and algorithm, and get good results.

References

[1] Fiedrich, F., Gehbauer, F., Rickers, U: Optimized Resource Allocation for Emergency Response after Earthquake Disasters, Safety Science, Vol.35 (2000), p. 41-57.

[2] Liao Jungli, H.K, in: Study on Network Reconstruction and Relief Logistics, Master Thesis at National Central University (2005).

[3] Kawashima, K., Miyaji, K. in: Seismic Performance Requirement of Highway Bridges, Proceedings of the 8th U.S. National Conference on Earthquake Engineering, Vol. 565(2006), p.1-10.

[4] Zhang Guoying, Lin Xianmao, in: Application of the Genetic Algorithm on VRP, Logistics Technology (2009).

[5] Liang Yana, Bao Jinsong, Hu Xiaofeng, Jin Ye, in: A heuristic project scheduling approach for quick response to maritime disaster rescue, International Journal of Project Management, Vol. 8(2009), p.620-628.

[6] J. Leaning, in: Disasters and Emergency Planning, International Encyclopedia of Public Health, Vol. 20(2008), p.204-215.

[7] Brysy O, Dullaert W, in: A Fast Evolutionary Metaheuristic for the Vehicle Routing Problem with Time Windows, International Journal of Artificial Intelligence Tools, Vol. 12(2002).

doi:10.4028/www.scientific.net/AMM.50-51.304

An intelligent feature selection method based on the Bacterial Foraging Algorithm

Dongying Liang, Weikun Zheng

Shenzhen Institute of Information Technology, Shenzhen 518029, P.R. China

liangdy@sziit.com.cn, zhengwk@sziit.com.cn

Keywords: bacteria foraging, agent, feature selection, genetic algorithm

Abstract. This paper puts forward an agent genetic algorithm based on bacteria foraging strategy (BFOA-L) as the feature selection method. The algorithm introduces the bacteria foraging (BF) behavior, and integrates the neural network and link agent structure to achieve fuzzy logic reasoning, so that the weights with no definite physical meaning in traditional neural network are endowed with the physical meaning of fuzzy logic reasoning parameters. The algorithm can maintain the diversity of the agent, so as to achieve satisfactory global optimization precision. The test result shows that this algorithm has good stability, little time complexity and high recognition accuracy.

Introduction

The issue of optimization widely exists in the scientific research and engineering practice. However, the optimization in global optimization (especially high-dimension complex function) is still an open problem. Normal optimization algorithm is slow in convergence and easy to fall into local optimal solution. In 2002, K.M.Passino brought forward the bacterial foraging algorithm (BFOA). In the BFA model, the solution to the optimization problem corresponds to the status of the bacterial in the search space, namely, optimize the fitness value of functions. The BFA algorithm includes three steps: chemotaxis, replication and dispelling. This algorithm has advantages such as parallel search of swarm intelligence algorithm and easy to jump out of local minimum.

Aiming at the dynamic neighborhood competition, this paper puts forward the agent genetic [3] evolutionary algorithm based on bacteria foraging link competition strategy (BFOA-LA). The bacterial foraging chemotaxis process can assure the local search ability of the bacteria. The replication process can accelerate the search speed of the bacteria. The BFOA-LA method adopts the dispelling process to strengthen the global optimization ability of the algorithm. It imitates and describes the human's thinking mode by using the fuzzy technology, integrates the fuzzy logic with the neural network to form the fuzzy neural network, and achieves mutual competition and intercross through link agent. This algorithm can not only raise the diversity of swarm, local search ability and search accuracy, but also control the number of selected features flexibly.

Description

Integration of fuzzy system and neural network

Construction method of the fuzzy neural network: here it adopts the integration of nerve cell and fuzzy model. This model is based on the fuzzy reasoning as main body, and adopts the nerve cell network to achieve the decision-making process of fuzzy reasoning. With the fuzzy method as sample, it trains the neural network to learn. Here it adopts the four-layer structure fuzzy neural network, as shown in Figure 1.

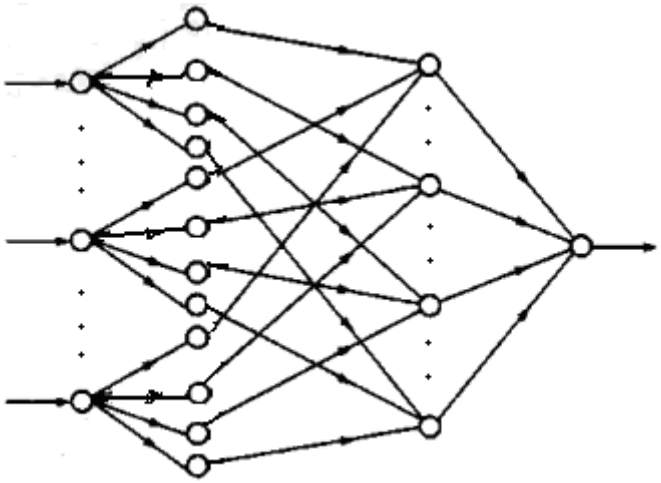


Figure 1 Structure of Fuzzy Neural Network

Layer 1 is the input layer, in which each nerve cell stands for an input variable. The number is equal to the number of variables appearing in the antecedent of fuzzy rules.

Layer 2 is the membership function layer, in which nerve cells are used to simulate membership function. Use the Sigmoid function's simulation to fuzz up signals. Its form is as below:

$$f(x)=1/(1+e-x) \quad (1)$$

Layer 3 is the rule layer, whose nerve cells are used to simulate the rule matching in the fuzzy reasoning process. The nerve cells of this layer completes the fuzzy " Add " operation.

Layer 4 is the output layer, in which the number of nerve cells is equal to the number of output variables. The layer is the linear combination of the above layer. The nerve cells of this layer have two functions: one is to simulate the membership function of conclusion; the other is to fuzz calculation and output the accurate value.

By constructing the fuzzy neural network, it can make good use of the advantages of fuzzy reasoning system and manual neural network for mutual complement, to prepare for the further work.

Fuzzy Neural Network Structure Based on Bacterial Foraging

Along with the increase of network complexity, traditional fuzzy neural network is very slow in learning process and easy to fall into local minimum. This paper integrates the adaptive link agent and bacterial foraging algorithm (BFA), and uses the three processes of BFA as reference. That is, it adds chemotaxis, replication and dispelling of BFA as learning strategy into the neural network training process. It builds the hybrid adaptive link agent network by adjusting neural network weights, so that the output of neural network converges quickly. This algorithm combines the coevolution idea and adopts the parallel searching mode of multi-sub swarm. The steps are: first, initialize the swarm to multiple sub swarms; then, each sub swarm adopts the link agent genetic algorithm mode to evolve and avoids falling into local optimum through multipoint searching. The sub swarms share and transfer genetic information by sharing agents, to reach the aim of finding satisfactory solution by the collaborative searching of multi-sub swarms. The optimization rate is thus effectively raised.

According to the methods in Refs [3], it constructs the circle link agent structure as shown in Figure 2. The agent competes and cooperates with agents in its neighborhood. The shared agent transmits the hereditary information of the previous sub swarm to the current sub swarm, which is beneficial for the diffusion of hereditary messages among the swarm.

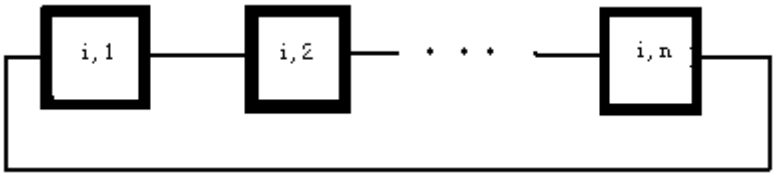


Figure 2 Structure of Link Agent

Make network weights training on the basis of the circle link agent structure (constructed in Figure2).

The BFOA-LA network model contains multiple comparatively independent four-layer structure agent fuzzy neural network BP network. Each individual occupies a grid. Being restrained by the local sensing ability, the agent can only interact with agents around it. Make network weights training on the basis of this networked agent.

The training of fuzzy neuron network is to search out the optimal structure and connection weights. It introduces the bacteria foraging method into the process, improves the agent network initial parameter algorithm, and then applies traditional BP algorithm to solve the iteration problem.

The construction method of the algorithm is as below:

1) Initialize the swarm and design the swarm size;

2) Design the fitness value of each individual in the current swarm. The fitness value decides the propagation and extinction of an individual. The formula of the fitness function is as below:

$$f(n)=\frac{1}{\frac{1}{2}\left[|x(n)|^2-\frac{G(|x(n)|^4)}{G(|x(n)|^2)}\right]^2} \tag{2}$$

Select the adaptive selection operator. The circle link agent adopts the neighborhood competition selection method; namely, the agent competes with agents around it. In the chemotaxis process of bacteria, the bacteria movement mode includes tumble and run. When the agent tumbles, it moves step-length in unit rightward or leftward. After one tumble, if the fitness value is improved, it keeps moving several steps in the same direction until the fitness value cannot be improved any longer; or it moves forward with the preset step value. By integrating multi-point searching mode, it can avoid falling into local optimum and be difficult to miss the excellent individual in the swarm.

Design the reproduction operator via the mixing proportion selection method; that is, integrating the replication actions of the tournament selection method and BFA. In view that the tournament selection method chooses the individual with the highest fitness value from some individuals each time, to avoid too strong randomness, based on the accumulation of each subnetwork's fitness value in chemotaxis process as standard, it releases half subnetworks which are poorer, and selects half subnetworks which are better for replication and splitting. The replicated subnetwork inherits the biological properties of the mother network and has the same structure and network weights as the mother network. It generates the next iterative swarm, so that the individuals with high fitness value have higher probability to be inherited to the next generation. The replication process accelerates the searching speed of the swarm.

Design crossover operator to maintain the diversity of the swarm. As for the complicated optimization situation, with the chemotaxis and replication process, the diversity of swarm becomes poor, and the swarm easily falls into local minimum. The introduction of dispelling process can strengthen the global optimization of the algorithm and maintain the diversity of network. It makes adaptive adjustment over crossover probability, and selects part network with certain probability to endow them with new initial weights. The dispelling of BFA is thus completed.

Design the adaptive mutation operator. Select individuals according to the mutation probability for mutation. Adopt uniform mutation. Randomly select a mutation point for mutation. If the fitness value of the newly-generated individual is higher than that of the original individual, replace the original individual with the new one.

Design the adaptive termination rule. If the swarm number reaching the requirement is above a threshold value, or if the hamming distance between the optimal individual of the current generation swarm and the optimal individual of previous generations is less than a given threshold value, the operation stops. Otherwise, it continues the selection of the next generation. Here, use the most fast grads descent method to obtain the iterative formula of network weight value.

Experimental Results and Analysis

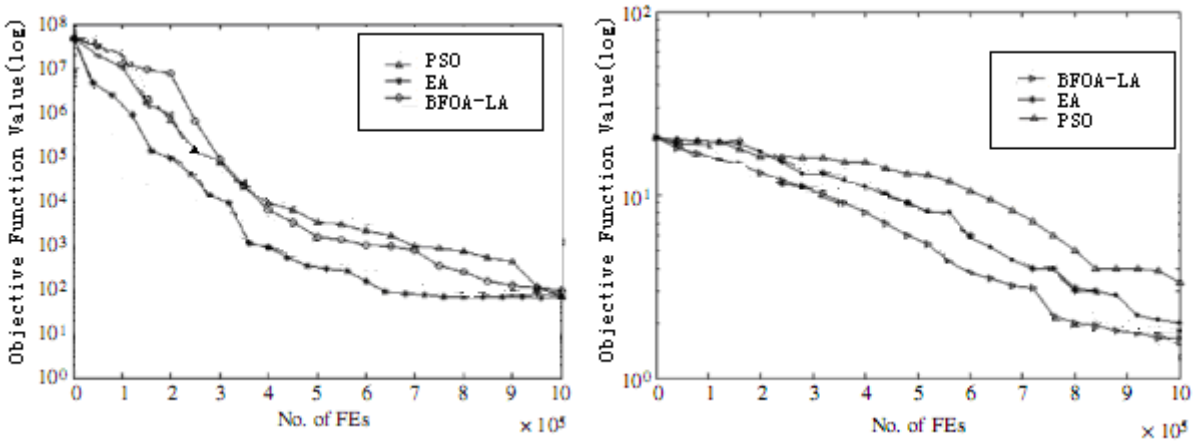
To validate the performance of this algorithm, compare with the particle swarm optimization (PSO) algorithm and Evolutionary Algorithms (EA). Select five Benchmark functions for numerical experiments. Among them, the former three functions are single-mode functions, and the latter two functions are multimode functions. The function description, value range and theoretical extreme value are shown as in Table 1.

Table 1 Test Functions

Function	Value Range	Theoretical xtreme value
$f_1(\bar{x})=\sum_{i=1}^{D-1}[100(x_{i+1}-x_i^2)^2+(x_i-1)^2]$	$(-100,100)^n$	$f_1(\bar{1})=0$
$f_2(\bar{x})=\sum_{i=1}^{D}[x_i^2-10\cos(2\pi x_i)+10]$	$(-10,10)^n$	$f_2(\bar{0})=0$
$f_3(\bar{x})=\frac{1}{4000}\sum_{i=1}^{D}x_i^2-\prod_{i=1}^{D}\cos(\frac{x_i}{\sqrt{i}})+1$	$(-600,600)^n$	$f_3(\bar{0})=0$
$f_4(\bar{x})=-20\exp(-0.2\sqrt{(\frac{1}{D}\sum_{i=1}^{D}x_i^2)}-\exp(\frac{1}{D}\sum_{i=1}^{D}\cos 2\pi x_i)+20+e$	$(-32,32)^n$	$f_4(\bar{0})=0$
$f_5(\bar{x})=[\frac{1}{500}+\sum_{j=1}^{25}\frac{1}{j+\sum_{i=1}^{2}(x_i-a_{ij})^6}]^{-1}$	$(-65.536, 65.536)^2$	$f_5(-31.95,-31.95)=0.998$

Table 2 Mean Value and Variance of Best-of-run Solution of Each Algorithm after 30 Times' Independent Tests under the Max. Evolutionary Generations

Function	D	BFOA-LA		PSO		EA	
		Mean value	Variance	Mean value	Variance	Mean value	Variance
f1	60	48. 215	13. 029	109. 875	31. 291	82. 251	24. 7642
f2	60	13. 225	2. 619	13. 214	5. 2631	20. 5041	10. 1842
f3	60	0. 1872	0. 0244	0. 7933	0. 2176	0. 2142	0. 1200
f4	60	1. 3729	0. 4512	1. 4632	1. 2531	1. 7695	1. 3563
f5	60	0. 9998	0. 0000	0. 9998	0. 0016	0. 9998	0. 0032



(a) Rosenbrock (b) Ackley

Figure 3 Benchmark Function Optimization

For the above algorithms, initialize with the same random seeds. Table 2 presents the mean value and variance of best-of-run solution of each algorithm after 30 times' independent tests under the

max. evolutionary generations. Table 2 and Figure 3 show that the BFOA-LA algorithm gets satisfactory optimization result over the high dimensional function: the mean value and variance of it obtained on the five test functions are not higher than those of PSO and EA; and the recognition accuracy of this algorithm is obviously superior to that of PSO and EA. Due to the decrease of the hereditary generation number and running time, the dimension complexity of the BP network reduces obviously, and the running time is also far fewer than that of the two other algorithms.

Conclusions

This paper integrates the BFA and fuzzy neural network into feature selection, and thus improves the global optimization efficiency. This algorithm introduces the circle link agent structure, and utilizes the neural network structure to achieve fuzzy logic reasoning. The combination of the bacteria foraging activity and the link competition strategy maintains the diversity of swarm individuals in the process of evolution, and makes in-depth search for high-dimension complex functions. It prevents the objective function value from falling into local optimum in a certain degree. The simulation experiment proves that the BFOA-LA algorithm is stable in performance and achieves good effect. It has enhanced the real-time ability and robustness of the detecting system.

Acknowledgment

The research work reported in this article has been supported in part by a Youth Scientific Funds of Shenzhen Institute of Technology (Grant No. QN-08008),and a Nature Scientific Funds of Shenzhen Institute of Technology (Grant No. LG-08004).

References

[1] A. Abraham, EvoNF: A framework for optimization of fuzzy reasoning systems using neural network learning and evolutionary computation, in: The 17th IEEE International Symposium on Intelligent Control, ISIC'02, IEEE Press, ISBN 0780376218, (2002), p. 327–332.

[2] P. Angelov, A fuzzy controller with evolving structure, Information Sciences 161 (1–2) (2004), p. 21–35.

[3] ZENG Xiao-Ping, LI Yong-Ming, WANG Jing, ZHANG Xiao-Juan, ZHENG Ya-Min .Link-like Agent Genetic Algorithm for Feature Selection Based on Competition Strategy [J]. Journal of System Simulation,Vol.20(8), (2008), p. 1973-1979.

[4] Y. Liu and K. M. Passino, "Biomimicry of social foraging bacteria for distributed optimization: Models, principles, and emergent behaviors,"J. Optimization Theory Applicat., Vol. 115, no. 3 , Dec.(2002),p. 603–628.

[5] D. H. Kim, A. Abraham, and J. H. Cho, "A hybrid genetic algorithm and bacterial foraging approach for global optimization," Inform. Sci., Vol. 177, no. 18, (2007), p. 3918–3937.

[6] Mishra, S.: A hybrid least square-fuzzy bacterial foraging strategy for harmonic estimation. IEEE Trans. on Evolutionary Computation, Vol. 9(1), (2005),p.61-73.

[7] Tripathy, M., Mishra, S., Lai, L.L. and Zhang, Q.P.: Transmission Loss Reduction Based on FACTS and Bacteria Foraging Algorithm. PPSN, (2006), p. 222-231.

[8] Ratnaweera, A., Halgamuge, K.S.: Self organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients, In IEEE Transactions on Evolutionary Computation 8(3) , (2004), p. 240-254.

[9] W. J. Tang, Q. H. Wu, and J. R. Saunders, "A novel model for bacteria foraging in varying environments," in Proc. ICCSA, LNCS vol. 3980.(2006), p. 556–565.

[10] Ding jian li,Chen zeng qiang,Yuan zhu zhi.On the Combination of Genetic Algorithm and Ant Algorithm[J].Journal of computer research and development, Vol.40 (9),(2003), p. 1351-1356.

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.309

Dynamic Instability Mechanism and Vibration Control of Radial Gate Arms

Jiliang Liu [1, a], Zhengzhong Wang [1, b], Xing Fang [2, c], and Hanmei Fang[3]

[1]Research Center of Water Engineering Safety and Disaster Prevention, Northwest A & F University, Yangling, Shaanxi, China, 712100

[2]Department of Civil Engineering, Auburn University, Auburn, Alabama, USA

[3]Northwest Investigation Design and Research Institute of CHECC, Xi'an, Shaanxi, China, 710065

[a]liujiliang430@163.com, [b]wangzz0910@163.com, [c]xing.fang@auburn.edu

Keywords: Radial gate arms, Failure, I-section, Dynamic stability, Vibration control measures

Abstract. The failures of the low head radial gates for hydraulic structures are commonly caused by the dynamic instability of arms (struts) under dynamic loads during gate opening and closing operations. For understanding the dynamic instability mechanism of radial gate arm, the dynamic stability of radial gate arm with I-section was analyzed with the consideration of bending, torsion and warping by using parametric resonance theory, and the boundaries of main dynamic instability regions were obtained. In order to avoid severe vibrations of radial gate arms, two vibration control measures were proposed. The study provides useful and valuable information for design and safe operation of radial gates.

Introduction

Radial gates for spillways are the most widely used gates in hydraulic engineering projects because they are the simplest, most reliable, and usually the least expensive type of crest gates. They require considerably smaller forces during maneuvering compared with other types of gates, have seating that is more reliable on the sill, require less space, and have better conditions for the passage of ice and wood. A conventional radial gate consists of a curved skinplate (panel) supported by a framework of horizontal and/or vertical girders and stiffeners and by two or more radial struts (arms). Gate arms converge downstream to trunnion bearings (pins) and transfer entire thrust of the water load (both static and/or dynamic loads) to the anchorage. Safe operation of radial gates is very important for any dams and spillways, by now, most of the existing radial gates have good operational performance. However, some of the radial gates with small design head in some countries such as China, Japan, USA and Portugal had operational failures that led heavy damage. According to investigations on the destroyed radial gates [1-3], there are two common characteristics on gate failures: (1) the destroyed arms lose dynamic stability and had buckling, and (2) the destroyed radial gates had severe vibrations under evident vibration loads. Fig. 1 shows two examples of destroyed radial gates: (a) flexural buckling failure, and (b) torsional-flexural buckling failure.Therefore, it is necessary and important to study the dynamic instability mechanism of radial gate arms in order to improve the design of radial gates to avoid failures.

(a) **(b)**

Fig. 1 Examples of Radial Gates for Failure

Under some special operational conditions, such as gate opening and closing process or partial opening status of radial gate, the dynamic load acting on radial gate may turn into exciting force

with certain period. According to dynamic stability theory [4], when the ratio of longitudinal exciting force frequency and the natural vibration frequency of the arm is a certain value, the amplitude of transverse plane vibration of the arm increases rapidly to lead resonance, and this type of the resonance is known as parametric resonance. In recent years, some researchers [5-9] have analyzed the dynamic stability of radial gate arms under periodical exciting force in order to find out the reason of failures of radial gates, however, in these studies, researchers only analyzed transverse plane vibration caused by parametric resonance and did not consider the effect of torsion and warping, so the results have some limitations.

I-section arms are commonly used in radial gates, and it is easy to undergo severe bending-torsion-warping coupling vibration under longitudinal exciting force that may cause failures of radial gates. Based on dynamic stability theory, this paper studied general calculation model for analyzing space dynamic stability of arms with the consideration of bending, torsion, and warping; and the boundaries of main dynamic instability regions were obtained. According to engineering experience, we proposed two vibration control measures in order to avoid severe vibrations of radial gate arms. This study provides useful information for dynamic design of radial gate arms.

Materials and Methods

Dynamic Instability Mechanism of Radial Gate Arms. Under some special operational conditions (e.g., gate opening and closing process), the dynamic load acting on radial gate arm can turn into exciting force with certain period. The longitudinal exciting force N with certain period acting on arms is defined as:

$$N=N_0+N_t\cos\theta t\ . \tag{1}$$

where N_0 is the time invariant component of N, i.e. hydrostatic pressure force; $N_t\cos\theta t$ is the periodical change of exciting force N, i.e. hydrodynamic pressure force; θ is the frequency of N; and t is the time variable. Assuming the periodical exciting force N is acting on the centroid of the cross section. Fig. 2 shows the schematic diagram of an I-section arm for studying dynamic stability.

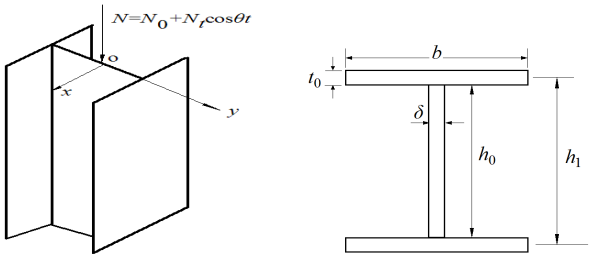


Fig. 2 Schematic Diagram of an I-section Arm under Periodical Exciting Force *N*

where o is the centroid of the cross-section; t_0 and b are the thickness and the width of the flange, respectively; δ and h_0 are the thickness and the height of the web, respectively; and h_1 is the total height of the cross section. The rectangular coordinate system is used for the study with the centriod o as its origin.

According to the thin-walled structure theory [4], when the arm is subject to longitudinal exciting force N, it undergoes parametric vibration, and the dynamic differential equation can be written as

$$\begin{cases} EI_x\dfrac{\mathrm{d}^4v}{\mathrm{d}z^4}+(N_0+N_t\cos\theta t)\dfrac{\mathrm{d}^2v}{\mathrm{d}z^2}+m\dfrac{\partial^2v}{\partial t^2}=0 \\ EI_y\dfrac{\mathrm{d}^4u}{\mathrm{d}z^4}+(N_0+N_t\cos\theta t)\dfrac{\mathrm{d}^2u}{\mathrm{d}z^2}+m\dfrac{\partial^2u}{\partial t^2}=0 \\ EI_\omega\dfrac{\mathrm{d}^4\varphi}{\mathrm{d}z^4}+\left[r^2(N_0+N_t\cos\theta t)-GI_t\right]\dfrac{\mathrm{d}^2\varphi}{\mathrm{d}z^2}+mr^2\dfrac{\partial^2\varphi}{\partial t^2}=0 \end{cases}\ . \tag{2}$$

where u and v are the kinetic deflections of the arm along x-axis and y-axis, respectively; φ is the torsion angle of cross section; z is the variable along the longitudinal exciting force N ; E and G are the elastic modulus of the arm and shear modulus of the arm, respectively; m is linear density of the arm; I_x and I_y denote the moments of inertia around x-axis and y-axis, respectively; I_ω is the warping constant of cross section; I_t is the torsion moment of inertia; and r denotes geometric feature of cross section. I_x, I_y, I_ω, I_t and r can be calculated by the equations below

$$\begin{cases} I_x = \dfrac{bt_0h_1^2}{2} + \dfrac{\delta h_0^3}{12}, I_y = \dfrac{t_0b^3}{6} \\ I_\omega = \dfrac{t_0b^3h_1^2}{24}, I_t = \dfrac{At_0^2}{3}, r = \dfrac{I_x + I_y}{2bt_0 + \delta h_0} \end{cases}. \tag{3}$$

Let l denote the length of the arm, and let

$$\begin{cases} u(z,t) = \sum\limits_{k=1}^{\infty} U_k(t)\sin\dfrac{k\pi z}{l} \\ v(z,t) = \sum\limits_{k=1}^{\infty} V_k(t)\sin\dfrac{k\pi z}{l} \qquad (k = 1,2,3,\cdots). \\ \varphi(z,t) = \sum\limits_{k=1}^{\infty} \Phi_k(t)\sin\dfrac{k\pi z}{l} \end{cases} \tag{4}$$

where $U_k(t)$, $V_k(t)$, and $\Phi_k(t)$ are the functions of time variable t. Substituting Eq. (4) into Eq. (2) yields

$$\begin{cases} m\dfrac{\mathrm{d}^2U_k}{\mathrm{d}t^2} + EI_y\dfrac{k^4\pi^4U_k}{l^4} - (N_0 + N_t\cos\theta t)\dfrac{k^2\pi^2U_k}{l^2} = 0 \\ m\dfrac{\mathrm{d}^2V_k}{\mathrm{d}t^2} + EI_x\dfrac{k^4\pi^4V_k}{l^4} - (N_0 + N_t\cos\theta t)\dfrac{k^2\pi^2V_k}{l^2} = 0 \\ mr^2\dfrac{\mathrm{d}^2\Phi_k}{\mathrm{d}t^2} + (EI_\omega\dfrac{k^4\pi^4}{l^4} + GI_t\dfrac{k^2\pi^2}{l^2})\Phi_k - r^2\dfrac{k^2\pi^2}{l^2}(N_0 + N_t\cos\theta t)\Phi_k = 0 (k = 1,2,3,\cdots) \end{cases}. \tag{5}$$

The main task of dynamic stability analysis of the arm is to determine the main dynamic instability regions. If we obtain the boundaries of dynamic instability, we can determine the main dynamic instability regions. According to the dynamic stability theory, when the amplitude of longitudinal exciting force is very small, the dynamic instability regions are near to $\theta = 2\Omega/i$ (i = 1, 2, 3, …), where Ω is the natural vibration frequency of the arm subjected to N_0. The region near $\theta = 2\Omega$ when $i = 1$ is the main dynamic instability region. Let

$$\begin{cases} \omega_x = \dfrac{k^2\pi^2}{l^2}\sqrt{\dfrac{EI_y}{m}}, \omega_y = \dfrac{k^2\pi^2}{l^2}\sqrt{\dfrac{EI_x}{m}}, \omega_\varphi = \dfrac{k\pi}{lr}\sqrt{\dfrac{EI_\omega\dfrac{k^2\pi^2}{l^2} + GI_t}{m}} \\ N_x = \dfrac{k^2\pi^2EI_y}{l^2}, N_y = \dfrac{k^2\pi^2EI_x}{l^2}, N_\varphi = \dfrac{1}{r^2}(EI_\omega\dfrac{k^2\pi^2}{l^2} + GI_t) \end{cases}. \tag{6}$$

Substituting Eq. (6) into Eq. (5) results in

$$\begin{cases} \dfrac{\mathrm{d}^2U}{\mathrm{d}t^2} + \omega_x^2\left(1 - \dfrac{N_0 + N_t\cos\theta t}{N_x}\right)U = 0 \\ \dfrac{\mathrm{d}^2V}{\mathrm{d}t^2} + \omega_y^2\left(1 - \dfrac{N_0 + N_t\cos\theta t}{N_x}\right)V = 0 \\ \dfrac{\mathrm{d}^2\Phi}{\mathrm{d}t^2} + \omega_\varphi^2\left(1 - \dfrac{N_0 + N_t\cos\theta t}{N_x}\right)\Phi = 0 \end{cases}. \tag{7}$$

Eq. (7) simplifies to

$$\begin{cases} U''+\Omega_x^2(1-2\mu_x\cos\theta t)U=0 \\ V''+\Omega_y^2(1-2\mu_y\cos\theta t)V=0 \\ \Phi''+\Omega_\varphi^2(1-2\mu_\varphi\cos\theta t)\Phi=0 \end{cases} . \tag{8}$$

where Ω_x, Ω_y, Ω_φ, μ_x, μ_y, and μ_φ can be calculated by the following equations:

$$\begin{cases} \Omega_x=\omega_x\sqrt{1-\dfrac{N_0}{N_x}},\Omega_y=\omega_y\sqrt{1-\dfrac{N_0}{N_y}},\Omega_\varphi=\omega_\varphi\sqrt{1-\dfrac{N_0}{N_\varphi}} \\ \mu_x=\dfrac{N_t}{2(N_x-N_0)},\mu_y=\dfrac{N_t}{2(N_y-N_0)},\mu_\varphi=\dfrac{N_t}{2(N_\varphi-N_0)} \end{cases} . \tag{9}$$

Eq. (8) has the same form with the Mathieu Equation. Using Eq. (8), we obtained the boundaries of the main dynamic instability regions of the I-section arm as follows:

$$\begin{cases} \theta=2\Omega_x\sqrt{1\pm\mu_x} \\ \theta=2\Omega_y\sqrt{1\pm\mu_y} \\ \theta=2\Omega_\varphi\sqrt{1\pm\mu_\varphi} \end{cases} . \tag{10}$$

After we obtained the boundary of main dynamic instability, we determined three main dynamic instability regions, and Fig. 3 shows one of them (Similarly, we have also determined the other two dynamic instability regions).

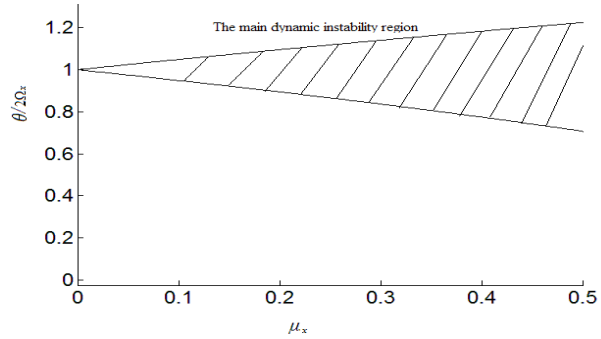


Fig. 3 The Main Dynamic Instability Region

The shadow region shown in Fig. 3 is the main dynamic instability region which is dangerous for radial gates that could undergo parameteric resonance under longitudinal exciting force. Therefore, the designers should make sure that ($\theta/2\Omega_x$, μ_x), ($\theta/2\Omega_y$, μ_y), and ($\theta/2\Omega_\varphi$, μ_φ) are far away from the main dynamic instabilities regions in order to avoid parametric resonance that may cause failures of raidal gates. For the above analysis procudure of dynamic stability of the I-section arm, becasue we considered bending, torsion, and warping that correspond to failure modes of radial gate arms, the calculation model of dynamic stability of the arm in this study is more reasonable and reliable than the previous studies and provide useful information for dynamic design of radial gate arms.

Vibration Control Measures: The severe vibration of radial gate arms caused by resonance is the main reason for the failures of radial gates, so vibration control measures should be considered in order to avoid the failures of raidal gates. Based on engineering experience[10,11], we proposed two vibration measures for engineers to use. The first measure is to add a rubber cushion at the junction of the main beam (welded on the skinplate of the radial gate) and the arm as shown in Fig. 4. The main beam and the arm are connected by prestressed bolts, and the prestress is equal to the pressure under the design head in order to assure the connection stiffness between the main beam and the arm. The second vibration control measure is to adopt viscoelastic dampers which are

frequently used in building structures and have good damping effect. The current viscoelastic damper is composied with steel splints and viscoelastic material connected by high-strength prestressed bolts, and the viscoelastic dampers are arranged at the end of the arm and are welded with the main beam as shown in Fig. 5. These two vibration control measures are proposed for designers and engineers to use in the design of radial gates.

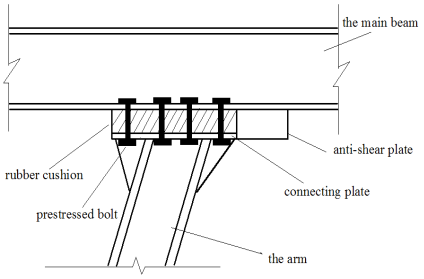


Fig. 4 The First Vibraion Control Measure

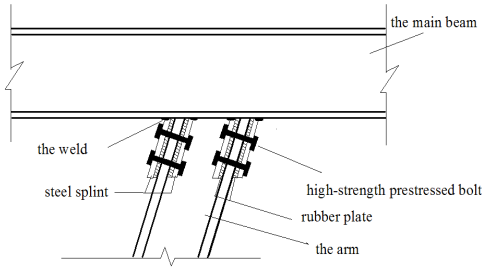


Fig. 5 The Second Vibraion Control Measure

Conclusions

This paper studied the dynamic stability of radial gate arm with I-section with the consideration of bending, torsion, and warping by using parametric resonance theory, and the boundaries of three main dynamic instability regions were obtained. In order to avoid serious vibration of radial gate arms, we proposed two vibration control measures: the first measure is to add a rubber cushion at the junction of the main beam and the arm, and the second measure is to adopt viscoelastic dampers arranged at the end of the arm. This study provides useful information for dynamic design of radial gates for safe operation.

Acknowledgements

This study was supported by a state 863 project (Grant No. 2002AA62Z3191) of the P.R. China.

References

[1] J.G. Zhang, and G.R. Liu: Journal of Hydroelectric Engineering Vol. 38 (1991), p. 49-57.

[2] T.P. Huang, and F. Wei: Journal of Metal Structure Vol. 5 (1986), p. 23-41.

[3] V. Robert: Waterpower Vo1. 1 (1999), p. 1-10.

[4] X.Z. Bolotin: *Dynamic Stability of Elastic System* (Higher Education Press, China 1960).

[5] J. G. Zhang: Water Power Vol. 4 (1985), p. 36-42.

[6] J.J. Lian, X.M. Peng, G.T. Cui, J.Y. Lin: Journal of Tianjin University Vol. 32 (1999), p. 171-176.

[7] L. Chen, Z.Z. Wang, and J.C. Pang: Yellow River Vol.28 (2006), p. 53-56.

[8] H.K. Li, and J.J. Lian: Journal of Nanchang University (Engineering & Technology). Vol. 28 (2006), p.72-78.

[9] Z.Q. Zhu, J.T. Zhuo, and G.L. Tao: Advances in Science and Technology of Water Vol. 19(1999), p. 27-29.

[10] Z.Z. Wang: Metal Structure Vol. 13(1994), p. 9.

[11] Z.Z. Wang: Metal Structure Vol. 14(1995), p. 1-2.

©(2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.314

Experimental Study On Collaborative Work Performance Of Multi-Functional Vibration-Absorption Structure To Enclosing Wall

Mingsheng He, Guiju Shi, Junwei Zheng

[1]College of Water Conservancy& Architectural Engineering, Shihezi University, Shihezi, 832003 P.R. China
Email: hms1971@sina.com

Keywords: Multi-functional Vibration-absorption Structure to Enclosing Wall, collaborative work performance, Experimental study

Abstract: Multi-functional Vibration-absorption Structure to Enclosing Wall (MVEW) is a new type of vibration absorption technology. In the structure, by the combination of frame filled wall, TMD and vibration absorption system, advantages of each part are used. To understand collaborative work performances of MVEW in minor earthquakes, 1/2 scale model of MVEW with two layers and one span is designed, and be carried out under three work situations(frame structure ,100mm and 200mm thickness enclosure wall), and monotonous lateral load tests were carried out. The result indicates that the shearing force bear by enclosing wall increase as the height increasing, similarly, it decrease by height increase for frame column. As the existing of enclosing wall, the stiffness and bearing capacity of MVEW are stronger obviously than frame structure. And when brittle compressive components are set at the top and bottom of the wall, the lateral force bear by the wall will be higher obviously than only setting (brittle compressive components) at the top of the wall, which showing that the wall will bear more lateral load by increasing touch point between enclosing wall and the column.

INTRODUCTION

Multi-functional Vibration-absorption Structure to Enclosing Wall (MVEW) is to a new type of vibration absorption technology by our project team proposing. The theory of the technique is that enclosure walls are connected with the main load-bearing structure by the brittle compressive and shear components, making it work together with the main load-bearing structure in minor earthquakes, and then the brittle components are destructed in medium and major earthquakes, enclosure wall as the mass block of tuned mass damper system (TMD) will work together with vibration reduction equipment around it to form a multi-functional vibration absorption system. therefore, the MVEW adopts to the earthquake resistant theory of multi-protect function and positive vibration absorption with obvious economic benefit and strong structural applicability. Article[1] showed that MVEW is converted into tuned mass dampers and can effectively control the vibration response. Therefore, studying force mechanism and collaborative work performance between main frame structure and enclosing wall of MVEW is very important to the structure. The force behavior of brittle components is put on in previous work by Liu and He [2]. Article[3,4]show that the structure is feasible to construction. Through recording the whole deforming process of the steel frames with and without infilled walls under the lateral forces , the effect of infillled walls on the load bearing capacity and stiffness are obtained [5]. To understand collaborative work performance of MVEW in minor earthquakes, 1/2 scale MVEW with two storey and one span is designed, three work situations which is frame structure ,100mm and 200mm thickness enclosure

wall are adopted , and monotonous lateral load tests were carried out.

Experimental design

Sample design

The frame model is designed to single span and two layers, which 1600mm tall and the span for 2500mm. Both the beam and the column are rectangular section, the size for columns section are 250mm x 250mm, for beams are 250mm x 200mm, for bressummers are 200mm x 150mm, and two size for the wall 2000mm×1000mm×100 and 2000mm×1000mm×200. All the frame columns and beams are poured as C30 concrete and the 28 days' cubic ideal compressive strength is 29.76MPa. Among the retaining wall is M5 mixed mortar laying aerated concrete block, aerated concrete masonry prismoidal compressive strength is 240 MPa. Moreover, the enclosing wall is built by aerated concrete blocks with M5 composite mortar, and the compressive strength of the prismoidal aerated concrete block is 2.4 MPa. The sample design can be seen in Fig.1, the material properties of the steel can be seen in Table 1.

Table 1 Material properties of steel

Steel Type	Rolling Shapes	Ultimate Tension [MPa]	Yield Strength [MPa]	Elastic Modulus [MPa]
ϕ 6	HPB235	538.9	458.0	2.1×10^5
ϕ 10	HPB235	456	415	2.4×10^5
ϕ 12	HPB235	432	395	2.3×10^5

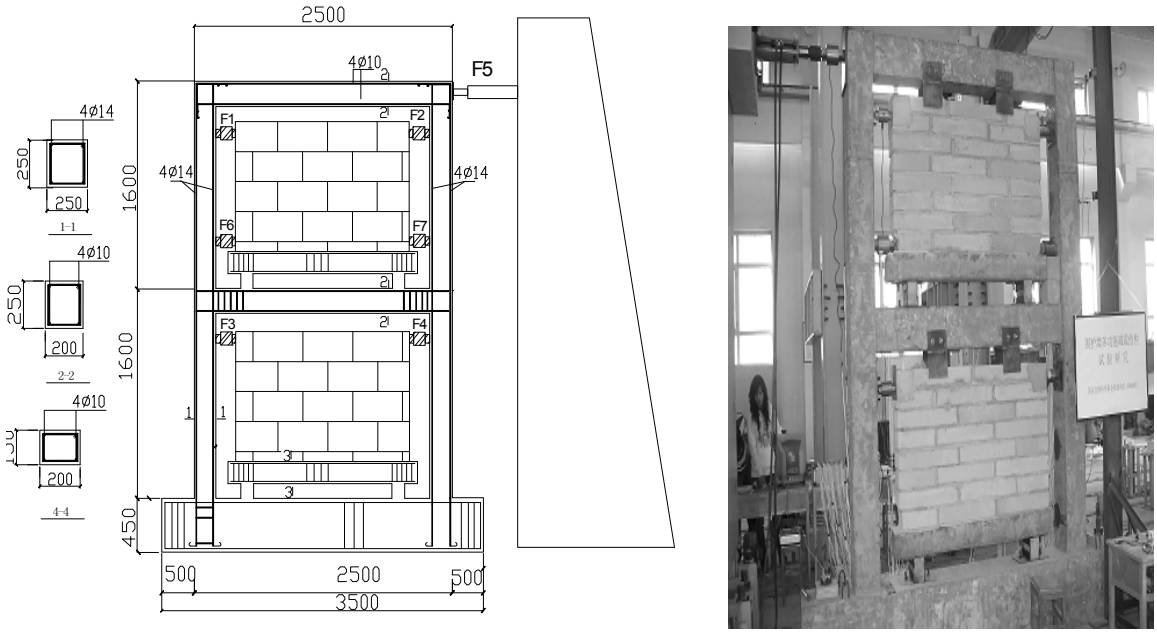


Fig.1 Dimension and reinforcement of specimens (mm)

Loading equipment and program

Loading process: Using level monotonic load, and the top displacement of the column is used as the control standard in loading process. For the pure frame, the displacement increment is 0.5mm and the test load divided is into 10 levels; For the MVEW, the displacement increment is 0.3mm and the test load is divided into 10 levels. Force sensor is set at the brittle resist-compression component to be joined to determine the loading ratio undertook by enclosing wall and the

framework. The setting of the force sensor use two working conditions, one is to set the force sensor at the top of enclosing wall, the other is to set it both at the top and the bottom of the enclosing wall.

Experimental results and analysis

Analysis of the sideway

The sideway refers to the top displacement of the structure. It is carried out main by the comparison of the different load-displacement curves of different structure forms.

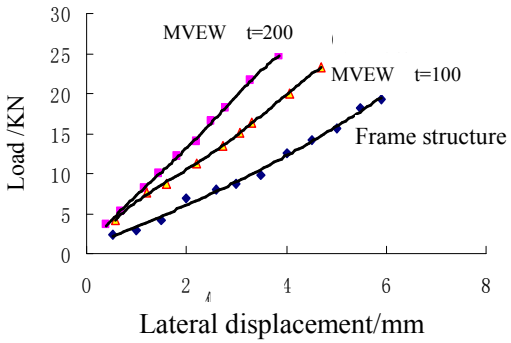


Fig.2 Load-displacement curves of different structure

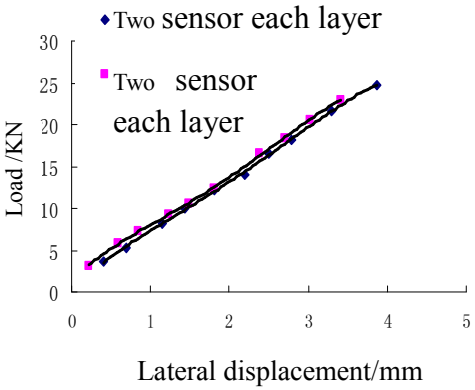


Fig.3 Load-displacement curves under different settings

It can be seen from Fig.2 that under the action of the same load, lateral stiffness of the structure larger as the increase of the thickness of the wall and the stiffness ratio between the wall and the column. The sideway decreases about 50% compared with the pure frame when the stiffness ratio is 4, and the decrease value is 25% when the stiffness ratio is 2. However, the lateral stiffness will decrease as the increase of the load in actual situation, but the inflection point does not appear because the experiment is taken under elastic range.

It can be seen from Fig.3 that when the structure has the same stiffness ratio, the reference between load and displacement is basic familiar rather than improving the lateral stiffness when one group and two groups of brittle resist-compression components are set respectively. This is mainly because the wall load delivered by the resist-compression component at the bottom is too small. The wall load mostly comes from the above resist-compression component, which receives confirmation in the analysis of this experiment.

The analysis of the load ratio bear by the wall and the frame

The brittle resist-compression component that replaced by the force sensor can obtain the wall load by the reading of the force sensor under level monotonic load. So the load ratio undertook by the wall can be worked out by comparing with apply the load directly to structure. According to formulas of the structure mechanics: the stiffness of the enclosing wall $D=\frac{12EI}{h^3}$; the stiffness of the frame column $K_1=\frac{EA}{3h}$; the stiffness ratio $\lambda_1=\frac{D}{K_1}$.

Table 2 The wall and column stiffness of different test model

Test Model	Frame stiffness[kN/m]	Enclosing Wall stiffness[kN/m]	Stiffness Ratio
Wall Thickness 100mm	47.6×103	193.94×103	4
Wall Thickness 200mm	47.6×103	96.97×103	2

Table 3 The load ratio bear by the wall

(a) Wall thickness t=200，one group of force sensor is set in each layer

F5 [kN]	F2 [kN]	F2/F5	F4/kN	F4/F5
3.624	1.380	0.381	0.789	0.218
5.285	2.086	0.395	1.192	0.226
8.154	3.374	0.414	1.946	0.239
9.966	3.702	0.371	2.150	0.216
12.231	4.554	0.372	2.732	0.223
14.043	6.132	0.437	3.616	0.258
16.459	7.017	0.426	4.243	0.258
18.271	6.876	0.376	4.210	0.230
21.593	7.303	0.338	4.913	0.228
24.764	9.587	0.387	6.335	0.256

(b) Wall thickness t=100，one group of force sensor is set in second layer

F5 [kN]	F2 [kN]	F2/F5	F4/kN	F4/F5
4.228	1.110	0.263	0.434	0.103
7.550	1.956	0.259	0.780	0.103
8.758	2.145	0.245	0.862	0.098
11.325	2.518	0.222	1.021	0.090
13.439	2.830	0.211	1.169	0.087
15.100	3.743	0.248	1.556	0.103
16.308	3.947	0.242	1.691	0.104
20.083	4.405	0.219	2.541	0.127
23.254	4.359	0.187	2.411	0.104

(c) Wall thickness t=100，two group of force sensor is set in each layer

F5 [kN]	F2 [kN]	F2/F5	F7/kN	F7/F5	F4 [kN]	F4/F5
3.171	0.000	0.000	0.246	0.078	0.000	0.000
5.738	1.609	0.280	0.395	0.069	0.806	0.141
7.248	2.228	0.307	0.529	0.073	1.141	0.157
9.211	2.777	0.302	0.670	0.073	1.634	0.177
10.570	3.574	0.338	0.893	0.085	1.787	0.169
12.382	4.712	0.381	0.942	0.076	2.356	0.190
16.517	7.593	0.460	2.088	0.126	4.176	0.253
18.422	8.640	0.469	2.442	0.133	4.695	0.255
20.536	8.465	0.412	2.201	0.107	4.740	0.231
22.952	7.759	0.338	1.770	0.077	4.220	0.184

It can be seen by the comparative analysis between (a) and (b) that the second layer enclosing wall bears about 0.4 of applied load when the thickness of the wall is 200mm and the stiffness ratio between the wall and the column is 4, however for the first layer enclosing wall is about 0.23;

Moreover, this two value change into about 0.22 and 0.1 when the wall thickness is 100mm and the stiffness ratio between the wall and the column is 2. It can be drawn that the load ratio bear by the wall increases as the increase of the stiffness ratio between the wall and the column. At the same time, it is also can be seen that the load ratio bear by the wall increase as the increase of the layer under the same stiffness ratio, and the second layer wall is two of the first layer. Thus it can be seen that the enclosing wall of the MVEW undertake quite a lot of lateral load.

It can be found that little influence on the load ratio bear by the wall when the brittle resist-compression component set at the bottom of the wall by comparison of the Table (a) and (c), and the load delivered by the column mostly bear by the brittle resist-compression component above. The cause of this phenomenon is that the separation of the enclosing wall and the top beam which result to big relative displacement between the wall and the column. However, the bottom of the enclosing wall is connected to the beam through the brittle resist-compression component whose relative displacement is little due to the wall offset as the beam, so the load ratio bear by the above wall become large yet little at the bottom.

From the Table (a), (b) and (c), it can be seen that the enclosing wall bear quite a lot lateral load relative to the frame, while this indicates that the enclosing wall has a large preliminary stiff and possesses deformation condition of itself to consume the building energy input by the earthquake. Afterwards, the frame gradually becomes the main anti-seismic component with the rigidity degeneration of the enclosing wall. That is to say, the enclosing wall acts as the main component in the first anti-seismic line in this progress which makes the frame withdraw the second defensive line. Therefore, the enclosing wall causes the increase of the defensive line for this part which achieves the purpose of this experiment.

Conclusion

Because of existing inclosing wall, the stiffness and bearing capacity of MVEW are stronger obviously than frame structure. So based on the advantage of MVEW , enclosing wall can be as first anti-seismic member.

When stiffness ration between wall and column is 4, bearing load to top brittle member in wall is about 40% , bearing load to bottom brittle member in wall is about 25% . when stiffness ration between wall and column is 2, bearing load to top brittle member in wall is about 24% , bearing load to bottom brittle member in wall is about 11% .Wall bear shearing force increase by height increase, similarly, it is decrease by height increase for frame column.

When top and bottom of wall is set brittle compressive components , lateral load of wall bearing is higher obviously than only setting brittle compressive components at top of wall, showing that the wall will bear more lateral load by touching point increase between enclosing wall and column.

References

[1] H.B.Zhang, M.S.He, C.G..Liu. Journal of Shihezi University, Vol.5 (2009), p637-641.
[2] C.G.Liu., M.S.Heng, H.B.Zhang. Journal of Shihezi University, Vol.3 (2010), p426-430.
[3] Z.F.Yuan, L.Yuan ; B.Liu . Building structure, Vol.5 (2010), P25-31
[4] Z.X.Guo; Y.B.Wu; Q.X.Huang**.** Journal of Earthquake Engineering and Engineering Vibration, Vol.6 (2008), P46-52
[5] Y.Z.Liu, G.Q.Li. Journal of Building Structures, Vol.3(2005), P78-84

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.319

Analysis On The Influence Factors Of Multi-Functional Vibration-Absorption Enclosure Wall Structures

Mingsheng He, Junwei Zheng, Gui-ju Shi

[1]College of Water Conservancy& Architectural Engineering, Shihezi University, Shihezi, 832003 P.R. China
Email: hms1971@sina.com

Keywords: multifunctional vibration-absorption, tune control, effectiveness of the shock absorption, optimized parameters

Abstract: Enclosure wall multi-functional vibration-absorption structures (MVEW) is a new style of damping structure, it integrates the merits of infilled frame, tuned mass control (MTMD) and the energy dissipation structures. The main influence factors of MVEW is analyzed in the paper, The results indicate that there are optimal value of the mass ratio, the stiffness ratio and the damping ratio of substructure in MVEW, and the damping effect become obvious as the increase of the number of substructure, it also shows that the best location ought to synthetically consider the number of substructure, the tuned frequency ratio and control mode instead of being fixed. In the end, the paper proposed the determine principle of the damping device's performance parameters as well as necessary optimization of MTMD parameters according to the specific case of actual structure1.

Introduction

As a new style of damping structure, the working mechanism of the enclosure wall multi-functional vibration-absorption structures MVEW) mentioned in previous work by Zhang and He [1] is to separate the main load-bearing structure and the functional enclosure wall, where the main structure endure the composite effects of vertical loads and horizontal earthquakes, and the enclosure wall works as a mass block of the technology of tuned mass control apart from surrounding function. It means that the enclosure wall as mass block cooperatively work with energy dissipated devices set below and the side of it to reduce the earthquake response, the elevation drawing of the control device in this structural system is shown in Fig.1. The features of this structure is that under small earthquakes, the main structure and enclosure wall will work as a whole to provide lateral stiffness to resist earthquakes, which similar to the infilled frame structure. Under the strong earthquakes, the enclosure wall with surrounding energy—dissipating device as Mass Damper Tuned (TMD) forms a multi-function tuned system combining with the main frame which can reduce earthquake damage effectively. Its principle is similar to multifunctional vibration-absorption and ant-vibration mega-frame structure [2,3]. The propose of MVEW, not only solve the insufficient of traditional tuned damping technology that it's need another special mass block which will increase the cost correspondingly, but also expand the theory space for the high-rise buildings and super-tall structures.

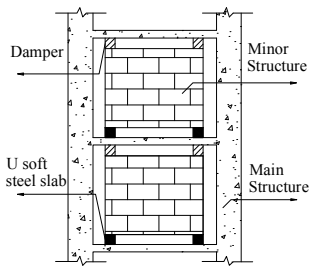


Fig.1 schematic diagram of multi-functional vibration-absorption enclosure wall structure

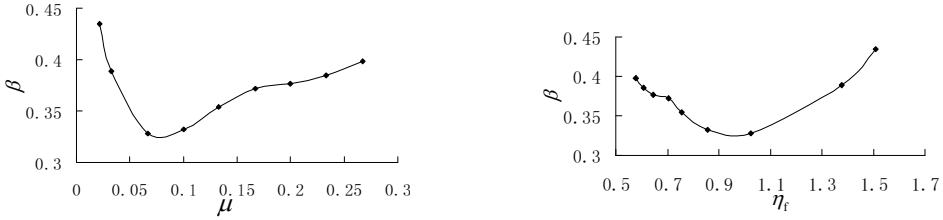
Based on the studied of the mechanism of the MVEW, the vibration response under the earthquake of the structure is analyzed deeply and summarizes the basic characteristics' influence on vibration reduction of this system.

Influence factors

There are many influence factors that affect the effectiveness of the multi-functional vibration-absorption enclosure wall structure. In this paper the earthquake response of a multi-functional vibration-absorption enclosure wall structure system whose main frame has six floors with each layer of 3 meters and lateral and longitudinal span of six meters is analyzed under EL-centro wave-for the peak acceleration 341.7 gal and set the enclosure wall substructure in its acting direction. We define damping coefficient β as the maximum displacement ratio between damping structure and seismic structures as the evaluation criterion for effectiveness of the system when analyze the influencing rule of every factors. Numerical simulations are conducted using SAP2000 and time-history method whose solution principle alike to Eq.1 is used to solve the earthquake response of the structure.

The influence factors of mass ratio between enclosure wall and the main frame on damping effect.

Mass ratio is an important factor that influences the effectiveness of the multi-functional vibration-absorption enclosure wall structure. We define mass ratio μ as the ratio between the substructure total mass and the mass of the main frame, also tuning frequency ratio η_f as the ratio between them. Fig3(a) shows the relation between mass ratio and damping coefficient of the structure vertex displacement, in unmerical examples, the value of stiffness of the U soft steel slab [4] takes 55 $kg \cdot rad^2/s^2$ and the value of damping ratio of every vibration mode take 0.05. Meanwhile, Fig.3(b) gives the relation between tuning frequency ratio corresponding to each mass ratio and damping coefficient of the structure vertex displacement.



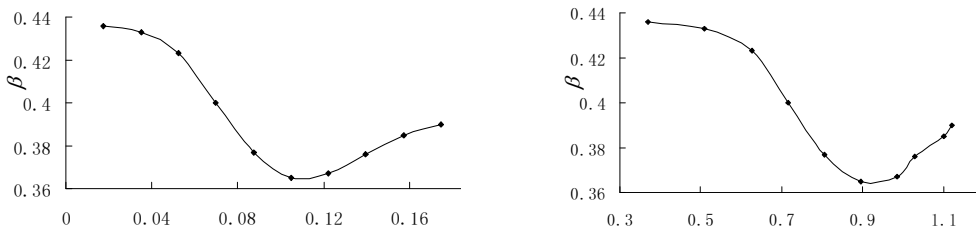
(a) The relation between μ and β b) The relation between η_f and β

Fig.2 Influencing curves of mass ratio on damping effect

It can be seen from Fig.2 that η_f decreases as the increase of μ, and β become smaller with the mass decrease of substructure under the case of certain number of enclosure wall substructure which shows that the damping effect is better. However, the increase of μ will aggravate the value of β when μ is greater than 0.1, which evince the worse of damping effect and the existence of optimal mass ratio. The proper range for the value of the tuning frequency ratio is between 0.8 and 1.0, so the optimal value range of μ in this example is about [0.05, 0.1] after comprehensive consideration of Fig.2(a) and Fig.2(b).

The influence of enclosure wall stiffness on damping effect.

In the numerical simulation with SAP2000, many parameters are interrelated and parameters will changed when one parameter is changed. So it need to undertake lots of simulations to get better effect. If the mass is certain, the vibration frequency of the structure under the earthquake will change with the change of stiffness as shown in Fig.3.



a) The relation stiffness ratio and β b) The relation η_f and β

Fig.3 Influencing curves of stiffness on damping effect

It can draw the conclusion after the analysis of Fig.3(a) and (b) that the corresponding frequency ratio will present the same trend when change the stiffness ratio, with the increase of the stiffness ratio, the damping coefficient β decrease gradually first and increase instead when stiffness ratio increases to a certain value. So damping coefficient has a minimum value which indicates that the stiffness ratio between the substructure and the main frame exist optimal value. The stiffness ratio takes 0.11 and the tuning frequency ratio for 0.9 in this example makes the best damping effect.

The influence of damping ratio of energy dissipation device on damping effect.

The inertia force of the enclosure wall plays the role of tuned damping with the main frame, besides, the energy dissipation device around the enclosure wall provides damping for the structure to dissipate energy. Consequently, the substructure' damping parameter also called damping ratio should be confirmed in structure design, namely $\xi = \frac{c}{2\omega m}$

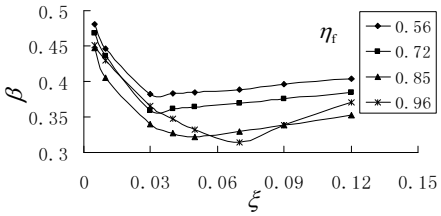


Fig.4 Influencing curves of damping ratio on damping effect

It can be observed that the variation trend of damping coefficient is similar when ξ increases from 0.005 to 0.120, and they all increases first and then decreases with an apparent minimal value in every influencing curves. It is thus clear that there is optimization problem of damping ratio as well. Here, when ξ=0.07, η_f=0.96, and the damping coefficient gets the minimal value which is 0.314, the structure achieves the best damping effect.

The influence of enclosure wall the number on damping effect.

The increase of number of substructure equivalent to broaden the scope of frequency distribution, but the structural response will be stable when the number increases to a certain amount. It is pointed out that the optimal number of TMD is 11 when the mass is certain as shown in the work of Li, Xiong and Wang [5].

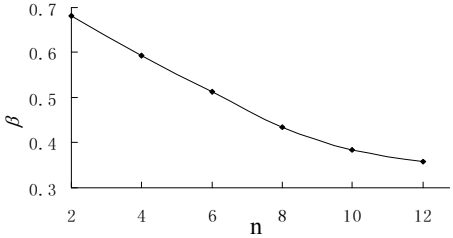


Fig.5 Relation the number of substructure and damping coefficient

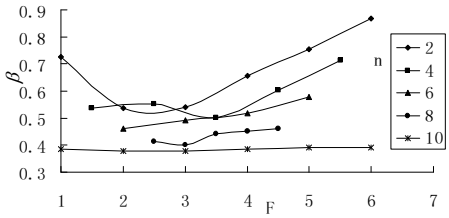


Fig.6 Relation curve between the locations of substructure and damping coefficient

Fig.6 shows the relation curve between the number of substructure and damping coefficient (an average value of the coefficient when the system has the same number of enclosure wall but different locations) under the condition that the total mass is fixed. It can be seen from the diagram that the damping coefficient decreases with the increasing of the number of enclosure wall which indicates that the more of the number of enclosure wall the less decrease degree of the vertex displacement of the structure, namely, the damping effect is more apparent. Meanwhile, the trend of the curve shows that the decrease degree of damping coefficient becomes smaller with he increase of the number of enclosure wall which indicates the existence of the optimal number of the enclosure in the structure. However, the number of enclosure wall in this example is 12 when all the wall are enclosure wall, but it also reach a conclusion from the diagram that the damping effect reach the best when n=12.

Furthermore, Fig.6 shows the influencing curve of the substructure location on the vertex displacement under the condition that the number of substructure is fixed. It can obtain a similar conclusion with Fig.5, which is to say the more of the number the better of the damping effect.

Conclusions

Under earthquake excitation, the lateral displacement cure of the main frame in MVEW is the same with traditional aseismic structures, but the MVEW has better effect on reducing the vertex displacement of the frame as the seismic response basically can decrease 30% to 60% in each plan.

It is found that the mass of enclosure wall substructures have optimal value which means that the thickness, material and the number of enclosure wall should be optimal designed.

The damping ratio and the stiffness ratio also exists optimal values, so parameter optimization should be considered in structural design to get better damping effect is suggested.

As the mass block in tuned mass control, enclosure wall not only reduce the cost of engineering construction, but also increase the convenience for construction due to its separation with the main frame. In addition, it is easier for seismic repair compared to traditional seismic structures.

References

[1] H.B.Zhang, M.S.Heng and C.G.Liu. Journal of Shihezi University, Vol.5 (2009), P637-641.
[2] Z.J.Lan, Y.J.Tian, S.Y.Cao and H.H.Wang. China Civil Engineering Journal, Vol.6 (2002), P1-5.
[3] G.P.Cai, F.Sun and J.Z.Huang. Engineering Mechanics, Vol.3 (2000), P55-59.
[4] Z.M.Xiong, X.P.Huo. Journal Of vibration And Shock Vol.27(2008), P124-129
[5] C.X.Li, X.Y.Xiong and Z.M.Wang. Journal of Vibration and Shock, Vol.4 (1999), P50-54.

doi:10.4028/www.scientific.net/AMM.50-51.323

Stocks Network Analysis Based on Visualization

Wangsen Lan [1,a], Guohao Zhao [2,b], Lijun Hou [1,c]

[1] Department of Mathematics, Xinzhou Teachers University, Shanxi, China

[2] School of Management Science and Engineering, Shanxi Finance and Economics University, Taiyuan, China.

[a]ws_lan@126.com, [b]gzhao1958@yahoo.com.cn, [c]houlyjun@126.com

Keywords: Stocks, Correlation, Complex network, Visualization.

Abstract. In order to explore interactions among stocks, a weighted complex network was modeled, whose nodes were stocks, connection between nodes was established by their correlation, and the weight on corresponding edge was their correlation coefficient. The article analyzed some influence behaves among stocks by using visualization method, and gave some algorithms to find out network maximum spanning tree and core network. The result shows that the network is scale-free, nodes coupling are of homology, and some nodes are of strong influence on others.

Introduction

Along with the widespread application of complex network in practice, stocks network research is more and more paid closely attention by people. By using complex network, researchers can find out the internal factors leading to stocks price fluctuations. The main significance of modeling stocks network is to explore interactions among stocks, estimate their cross-correlation level.

Representative stocks network models usually are some weighted indirect complex networks [1-5]. In the networks, nodes are stocks, the connection between nodes is usually established by the correlation between synchronizing time serial data of two stocks, edge weights are correlation coefficients. The article will model a stocks network for coal and power sectors in China stock market, and analyze the influence among the stocks by some visualization methods.

Model

In our research, the data set is composed by daily closing price of 98 stocks of coal and power sectors in China stocks markets. All data were collected from http://www.hexun.com with time period of most long span from the beginning of 1991 to the end of 2009.

Assume has n stocks with price $P_i(t)$ for stock i at time t. The logarithmic return of stock i is $s_i(t) = \ln P_i(t) - \ln P_{i-1}(t)$, which, for a certain consecutive sequence of trading days, forms the return vector s_i. In order to characterize the synchronous time evolution of stocks, we use same moment stocks return correlation coefficients between stocks i and j, and it is defined as

$$\gamma_{ij} = \frac{\langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle}{\sqrt{\left(\langle s_i^2 \rangle - \langle s_i \rangle^2\right)\left(\langle s_j^2 \rangle - \langle s_j \rangle^2\right)}} \tag{1}$$

Where $\langle \cdots \rangle$ denotes a time average over the trading days corresponding in the return vectors. All correlation coefficients, forming a $n \times n$ matrix with $-1 \le \gamma_{ij} \le 1$, are then transformed into a $n \times n$ influence matrix with elements $|\gamma_{ij}|$, such that $0 \le |\gamma_{ij}| \le 1$. In order to observe in detail mutual influence behavior of stocks, we prearrange different threshold value γ_0, and analyze the network nature for $|\gamma_{ij}| > \gamma_0$. Let

$$w_{ij} = \begin{cases} |\gamma_{ij}|, & |\gamma_{ij}| > \gamma_0 \\ 0, & |\gamma_{ij}| \le \gamma_0 \end{cases} \tag{2}$$

We model *Stocks Network for Coal and Power Sectors in China Stock Markets* (shorter form SNCPSCSM) denoted as $G(V, E, W)$, where $V = \{1, 2, \cdots, n\}$ is a set with n nodes to represent n stocks, E is edges set of the network, $W = (w_{ij})_{n \times n}$ is weight matrix satisfied $w_{ij} = w_{ji}$ $(i \neq j,\ i, j \in V)$ and $w_{ii} = 0 (i \in V)$ that are meaning the network is indirect and not self-connected for all nodes. The connection between nodes is dependent on their edge weight, in other words, node i and j are connected if and only if $w_{ij} > 0$. We are more interested in strong influence among stocks, and only take into account $\gamma_0 \ge 0.55$, and all analysis are took after having deleted isolated nodes.

Correlation Matrix Visualization

Weight matrix includes all information about network nodes connection. Weight matrix of SNCPSCSM we are researching is just stocks correlation matrix, whose visualization graph can give us the most intuitive survey for influence among stocks.

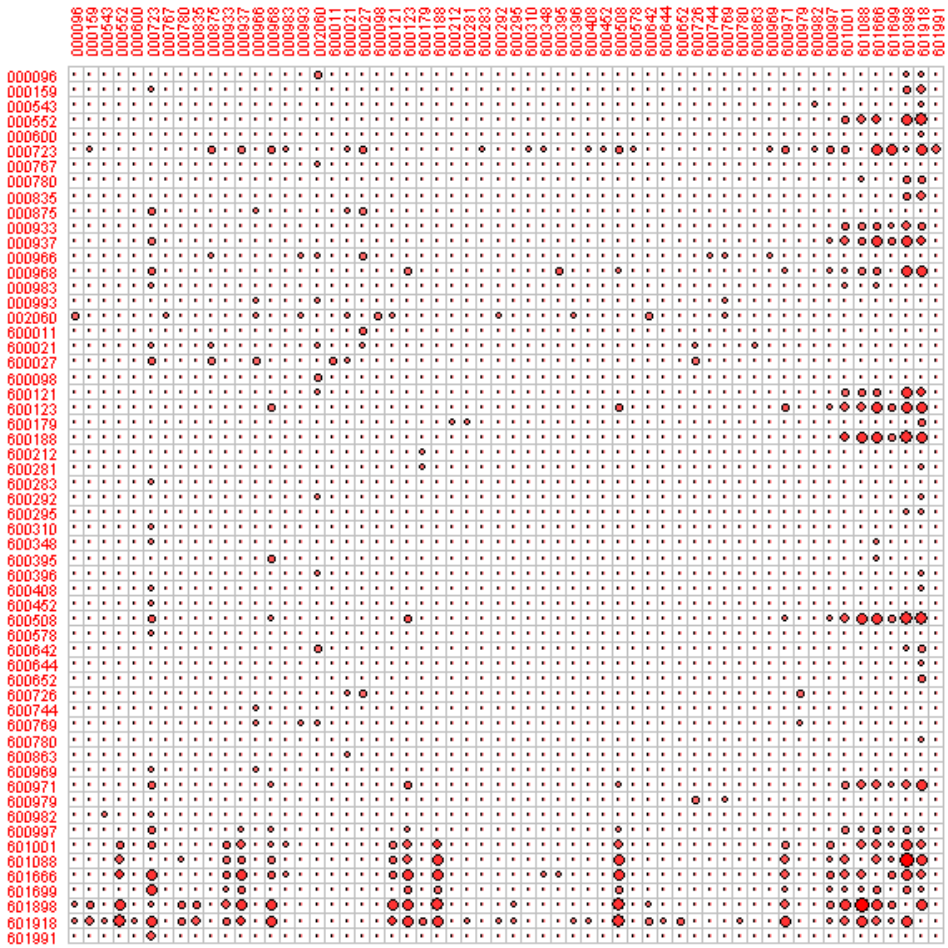


Fig. 1 SNCPSCSM nodes correlation matrix graph for $\gamma_0 = 0.65$. Bigger size plotted circles denotes stronger influences between stocks.

There are a lot of matrix visualization algorithm that can be searched in internet, which are all simple, and don't layout in the article. In practice operation, we must pay attention to using different size symbols to denote different value in order to plot correlation matrix.

As showed in Fig. 1, in SNCPSCSM, node 000723, 601001, 601088, 601666, 601699, 601898, 601918 produce great influence on other nodes.

Network Scale

Many researches showed that stocks networks were free-scale, whose node degree distributions were power law distribution $p(s) \sim s^{-\delta}$ [1-2, 5], such as S&P500 stocks network with $\delta = 1.8$ [1-2]. For the stocks network in real estate sector of china stock market, the node degree approximate power law distribution, whose δ shows a linear growth with threshold γ_0, $0.8 < \delta < 1.6$, and average 1.25 [5]. For SNCPSCSM that we are researching, the nodes degree distribution also approximate power law, $\delta = 0.41$ for $\gamma_0 = 0.55$, $\delta = 0.6$ for $\gamma_0 = 0.60$, $\delta = 0.87$ for $\gamma_0 = 0.65$, and $\delta = 0.94$ for $\gamma_0 = 0.70$, average 0.7, this shows indistinctive scale-free nature.

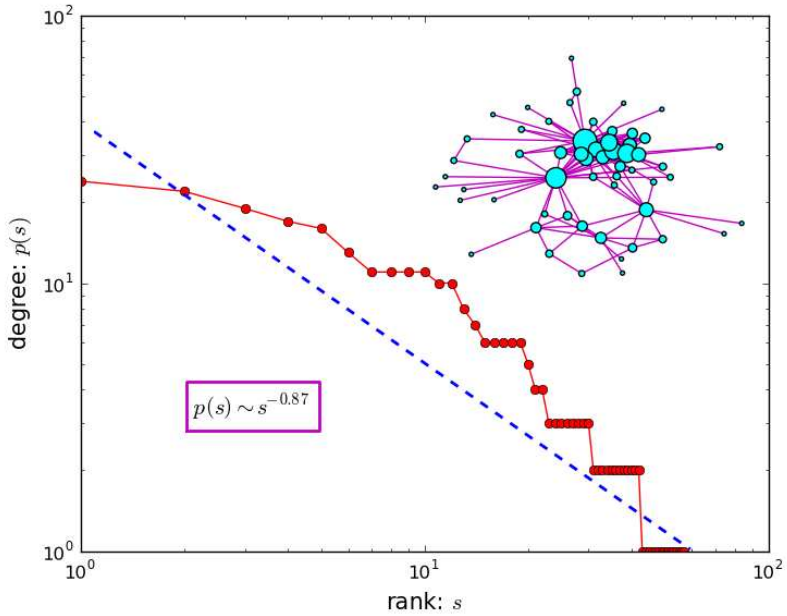


Fig. 2 SNCPSCSM degree distribution $p(s)$ plot for $\gamma_0 = 0.65$. The network layout is on the upper right corner, where bigger size nodes with bigger degree.

Maximum Spanning Tree and EGO Structure

In the practical application, many network optimization questions involve minimum spanning tree, whose some classical algorithm is good, such as Prim algorithm and Kruskal algorithm. In the stock network that we are researching, we emphasize strong influence among stocks, and discuss maximum spanning tree instead of minimum spanning tree.

Suppose the maximum spanning tree of $G(V, E, W)$ is T. Obviously, $T \subset G$. The maximum spanning tree problem can be summarized as following mathematical model:

$$\max Z = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}x_{ij}$$

$$s.t.\begin{cases} \sum_{i,j\in S} x_{ij} \le |U| - 1, \forall U \in V, U \neq \Phi \\ \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} = n - 1 \\ x_{ij} = 0 \text{ or } 1, \ i, j \in V \end{cases} \tag{3}$$

Where $|U|$ is nodes number in U. $x_{ij}=1$ if T includes the edge linked node i and j, otherwise $x_{ij}=0$. Obviously, maximum spanning tree algorithm can be implemented by either referring to or transforming into minimum spanning tree.

Stocks weighted network $G(V,E,W)$ is of some different characteristics from other network, such as $0<w_{ij}<1$ and $w_{ii}=0$. We can design a simple method to transform maximum spanning tree into minimum spanning tree.

Step 1. To make transformation $w'_{ij} \leftarrow 1-w_{ij}$ for $w_{ij} \in G$ and $w_{ij} \neq 0$, and obtain a new network $G'(V,E,W')$;

Step 2. To find out the minimum spanning tree T' of $G'(V,E,W')$;

Step 3. To make inverse transformation $w_{ij} \leftarrow 1-w'_{ij}$ for $w'_{ij} \in T'$, and obtain a new tree T, which is just the maximum spanning tree of $G(V,E,W)$.

We can see some things from visualization maximum spanning tree of SNCPSCSM (Fig. 3). There are in total 6 EGOnets, 4 of which are all composed of same sector stocks. 2 EGOnets belong to power sector, and 2 EGOnets belong to coal sector. Only 2 EGOnets are mixed by different sector stocks, whose center nodes are all from coal sector. All these results show that nodes couplings are of homology.

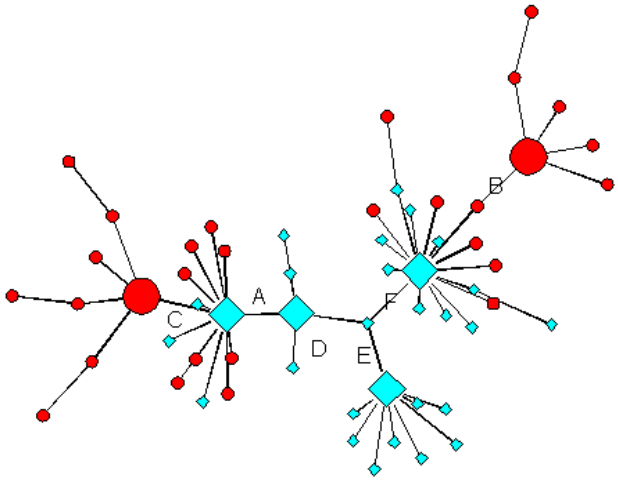


Fig. 3 SNCPSCSM maximum spanning tree. It be composed of 6 EGOnet, 6 bigger size nodes are respectively from left to right: C(600027), A(000723), D(601666), E(601898), F(601918), B(002060). Nodes with different colors and shapes denote stocks from different sectors, red circle is power sector, and cyan diamond is coal sector.

Core Network

In many literatures on complex network, communities are one of important research object. It is more meaningful to find a core network than communities for SNCPSCSM; moreover we have researched its maximum spanning tree and EGO structures, which can also indicate some clique natures though not exact communities.

Suppose weight matrix $W=(w_{ij})_{n\times n}$ eigenvalues are $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$, and corresponding unit eigenvector $v_1, v_2, \cdots, v_n$, satisfying $Wv_i=\lambda_i v_i$, where $1 \leq i \leq n$. From the opinion of geometry, v_i represents an n-dimensional vector, λ_i represents magnification scale of v_i after W transformation. We are only interested in $|\lambda_i|>1$, and might as well suppose $|\lambda_i|>1$ for $1 \leq i \leq k$. Further suppose $v_i=(a_{i1}, a_{i2}, \cdots, a_{in})^{\mathrm{T}}$, $\lambda_i a_{ij}$ $(1 \leq j \leq n)$ represent projection of v_i on dimension j after W transformation. We introduce a new variable:

$$p_j = \sum_{i=1}^{k} |\lambda_i a_{ij}| \quad (4)$$

Where $1 \le j \le n$, which represent the synthetically important degree of nodes in network, and use it to detect all core network nodes. We list the steps of their algorithm:

Step 1. Calculate eigenvalues and corresponding eigenvectors of weight matrix W;

Step 2. Find out the eigenvalues whose absolute values are bigger than 1 and eigenvectors corresponding with them;

Step 3. Calculate p_j for $1 \le j \le n$, sort p_j from big to small, and select some top nodes as core network nodes.

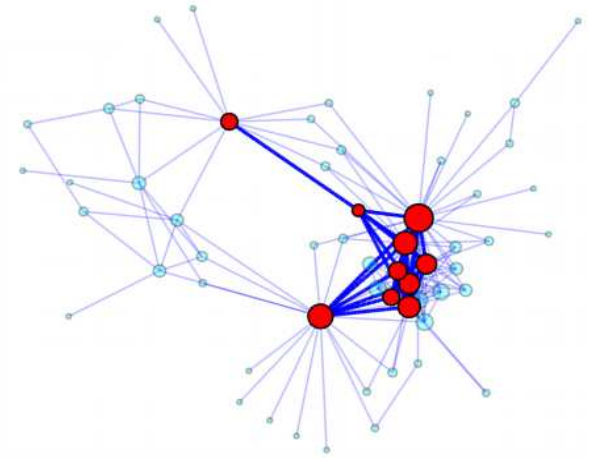


Fig. 4 SNCPSCSM core network. Translucent graph is mother network SNCPSCSM, and red nodes and blue edges are Core network plot. Bigger size nodes with bigger degree.

As showed in Fig. 4, SNCPSCSM core network with 10 nodes 000723, 002060, 600121, 600508, 601001, 601088, 601666, 601699, 601898, and 601918, which are either with bigger degree or with bigger betweenness.

Acknowledgments. Project supported by *National Natural Science Foundation in China*, No. 70873079 and No. 70941022, *Science & Technology Research & Development Projects in Higher Education Institution of Shanxi Province of China*, NO. 20091148.

References

[1] H. J. Kim, I. M. Kim: Scale-Free Network in Stock Markets. J. Kor. Phy. Soc. Vol. 40(2002), p. 1105-1108.

[2] H. J. Kim, Y. Lee, B. Hahng, I. M. Kim: Weighted Scale-Free Network in Financial Correlation. J. Phy. Soc. Jap. Vol. 71(2002), p. 2133-2136.

[3] K. E. Lee, J. W. Lee, B. H. Hong: Complex networks in a stock market. In: Conference on Computational Physics 2006. Com. Phy. Comm. Vol. 177(2007), p. 186.

[4] W.Q. Huang, X. T. Zhuang, S. Yao: A network analysis of the Chinese stock market. Phy. A. Vol. 388(2009), p. 2956-2964.

[5] W. S. Lan, S. D. Zhang. in: *Recent Advance in Statistics Application and Related Areas*, edited by K. L. Zhu. Vol. II, p. 2235-2239. AAPH Press, Sydney (2009).

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.328

Vibration of Double Cable Suspension Bridge under Vehicle Load

Nanhong Ding[1, a], Lixia Lin[2,b] and Weihua Liao[3,c]

[1]School of Civil Engineering Southwest Jiaotong University, Chengdu 610031, China

[2]School of Civil Engineering Lanzhou jiaotong University, Lanzhou 730070, china

[3]School of Civil Engineering Lanzhou jiaotong University, Lanzhou 730070, china

[a]nanhongd@126.com, [b]LinLixia@126.com, [c]whliao@126.com

Keywords: double cable suspension bridge, bridge carriageway pavement irregularity, vehicle-bridge coupling vibration, impact coefficient, cable shape coefficient

Abstract. A single moving mass-spring-damper model is adopted to simulate the vehicle model. The equation of vehicle-bridge coupling vibration is derived, using D'Alembert principle and the conditions of displacement compatibility. The dynamic responses of a double cable suspension bridge to the vehicle moving at different speeds are analyzed, considering the geometric nonlinearity and the bridge carriageway irregularity factor, under two types of vehicle loading conditions, namely moving along the center and the eccentric longitudinal axis. Then the influence of vehicle velocity and bridge carriageway irregularity on impact coefficient of double cable suspension bridge is discussed. Single cable suspension bridge can be taken as a special case of double cable suspension bridge, after the main cable shape coefficient is introduced. The dynamic responses of double cable suspension bridge and single cable suspension bridge are compared to reveal the character of vehicle vibration of double cable suspension bridge. The study of the dynamic responses character of double cable suspension bridge has a positive significance on structural form selection of such type bridge during designing, dynamic performance evaluation and vibration control.

Introduction

Double cable suspension bridge has two cables which intersect and link together in the middle of bridge span in one suspender plane as shown in figure 1. The double cable system has better stiffness than single cable system, and has the capability of adapting to unsymmetrical loads [1]. The double cable system has an obvious advantage in static property, but it has rare reports about the dynamic properties. Structural character of double cable suspension bridge determines its unique vibration characteristics. The natural vibration characteristics of the double cable suspension bridge has been studied by the paper author, and compared with the single main cable suspension bridges [1]. At present, scholars have done many studies on vehicle-bridge coupling vibration of long span suspension bridge[2~4], but almost all aimed at single cable suspension bridges (that is the ordinary suspension bridge), There have not researches on the vibration response of double cable suspension bridge under vehicle load.

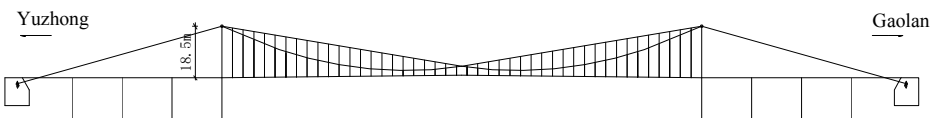


Fig.1 Double cable suspension bridge profile

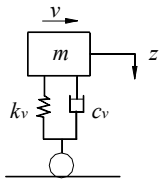


Fig.2 Vehicle model

Vehicle-bridge coupling vibration under 3 different vehicle models: the moving mass, 1/4 vehicle model and 1/2 vehicle model are studied in literature [5], results showed that: while considering the effects of springs, damping of vehicle system, it can more fully reflect the coupling effects of bridge, but bridge response become more complex. For small span bridge, the body nod stiffness has more

obvious influence on the bridge dynamic response[5], for large span bridge, such as the double cable suspension bridge whose span is far longer than the length of a vehicle, the vehicle dynamic characteristics has little effect on the dynamic response of bridge[6], in order to make analysis simple, take into account the impact of body springs and damping, and both consider the advantages of 3 models, a single moving mass-spring-damper model is applied to simulate vehicle in the paper. The equations of vehicle-bridge coupling vibration are derived and the dynamic responses of the bridge to moving vehicle are analyzed taking Qingcheng Yellow River Bridge as an example, considering the geometric nonlinearity, vehicle speed and the bridge carriageway irregularity factor. The dynamic responses of double cable suspension bridge and single cable suspension bridge are compared to reveal the character of vehicle vibration of double cable suspension bridge.

Theoretical Analysis

Motion Equation. Node displacement is measured from its static equilibrium position, signed y_j with down as positive. For each non-zero displacement of the nodes r (total number is N) the balance equations are listed as following:

$$m_r \ddot{y}_r + c_r \dot{y}_r + \sum_{j=1}^{N} k_{rj} y_j = F_r \ . \qquad (1)$$

In which, m_r , c_r are the concentration mass and viscous damping coefficient at node r respectively; k_{rj} is the stiffness coefficient; F_r is the bridge-vehicle interaction equivalent force at node r.

Using a single moving mass-spring-damper vehicle model shown in figure 2, we take the vehicle as a separate body, apply D'Alembert principle, motion equations of vehicle are listed as following:

$$m\ddot{z} + (F_r - F_{st,r}) = 0 \ . \qquad (2)$$

In which, m is the mass of vehicle; z is the displacement from the vehicle static equilibrium position; $F_{st,r}$ is the force of the vehicle in its static equilibrium position.

Equation (1) and (2) are coupled each other through its interaction force F_r, considering the vehicle spring-mass system, the relationship equations between force and displacement are added:

$$u = z - d_F - y_F \ . \qquad (3)$$

In which, u is the relative displacement for vehicle system, through $F_r \sim u$ relationship curve to determine F_r; d_F is the coordinate of bridge shape function $d(x)$ at the point F_r, bridge pavement irregularity can be considered by $d(x)$, y_F is vibration instantaneous displacement of the bridge at load point F_r.

The vehicle-bridge dynamics balance equation (1) ~ (3) can be solved by using Newmark-β step integration method, taking road irregularities as input excitation.

Simulation of Deck Roughness. Bridge deck roughness is the main factor for the vehicle vibration. Bridge deck roughness can be assumed to meet the zero-mean stationary Gaussian random process, which is obtained by numerical simulation of irregularity curves [7,8].

Coupling Vibration Response Analysis

Natural Vibration Characteristics. The dynamic characteristics are solved by using Newton-Raphson iteration to determine the structure tangent stiffness matrix under gravity [1], studies show that: mode shapes group obviously, the main beam vibration mode shapes appear first, then the cable and beam higher vibration modes, the vibration of the tower is more behind, which are consistent with trend of single cable suspension bridge. The first order mode of single cable suspension bridge is generally asymmetric vertical bending, while that of double cable suspension bridge is symmetric vertical bending, so in the same structural parameters, the double cable suspension bridge can effectively improve the frequency of vertical bending.

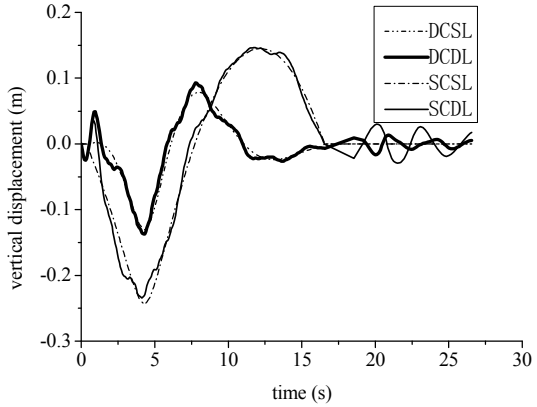


Fig. 3 Vertical displacement-time curve of ¼

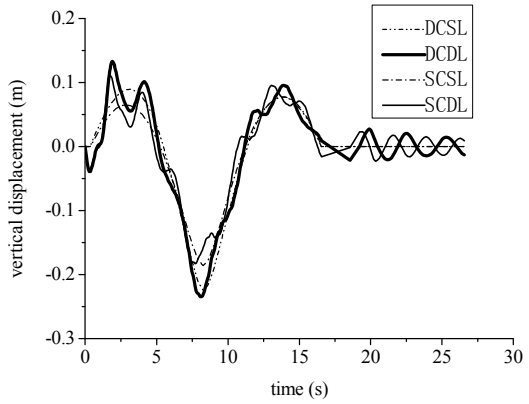


Fig.4 Vertical displacement-time curve of ½

Coupling Vibration Response Analysis. The Newmark-β numerical integration method is carried out, after taking the first 400 orders bridge vibration modes, damping ratio is 0.02, β is 0.25. The coupled vibration response of bridge is calculated considering the different vehicle speed from 20 km/h to 120 km/h and based on A grade road. Only displacement response curves under 40 km/h vehicle speed are drawn: The vertical displacement-time curve of ¼, ½ section when vehicle moving along the center longitudinal axis of stiffening beam is shown in figure 3 and 4. The transverse displacement-time curve of ¼, ½ section when vehicle moving along the eccentric 1.30m from longitudinal axis of stiffening beam is shown in figure 5 to 6. In figure 3 to 6, DCSL means double cable and static load, DCDL means double cable and dynamic load, SCSL means single cable and static load, SCDL means single cable and dynamic load.

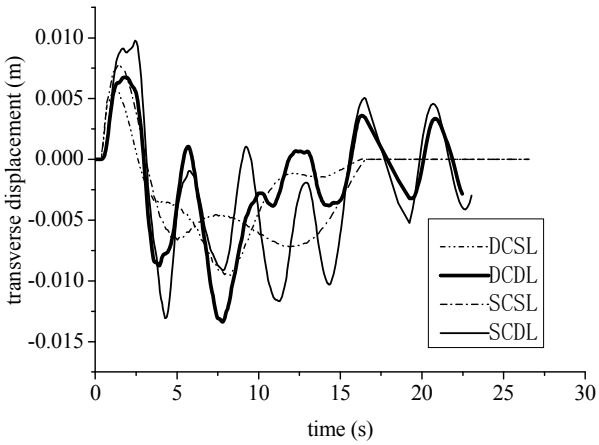


Fig.5 Transverse displacement-time curve of ¼

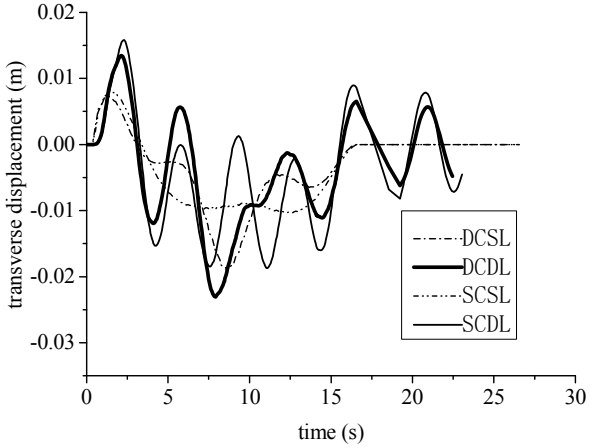


Fig.6 Transverse displacement-time curve of ½

Calculation shows that the fluctuations of vertical displacement slow down with increase of vehicle speed. When vehicle speed is 40 km/h, the vertical displacement-time curves of the stiffening girder ¼ and ½ section is close to static live load influence line. The maximum displacement of stiffening girder is -0.234 m occurred when the vehicle reaches to ½ section, while the maximum displacement of ¼ section is -0.138 m. The maximum value of bridge dynamic deflection reaches at

the middle span, mainly because the first order mode shape of double cable suspension bridge is positive symmetric vertical vibration.

Transverse displacement response fluctuation increase with increase of vehicle speed, such as the maximum transverse displacement reaches to 4.2 cm when vehicle speed is 120 km/h. Figure 6 and 7 show that the transverse displacement of the bridge changes gradually with the vehicle moving position, it can be concluded that the transverse displacement of the bridge is mainly determined by the bridge's first order transverse mode.

Vehicle Speed Effect. Introduce impact coefficient as following:

$$1+\mu = y_{d\max} / y_{s\max} \quad . \qquad (4)$$

In which, $y_{d\max}$ and $y_{s\max}$ means the maximum dynamic and static deflection of the bridge respectively. Three bridge road levels are discussed: no pavement roughness, A grade and B grade.

Impact coefficient does not increase linearly with the increase of vehicle speed as shown in figure 8. The impact coefficient is larger, when speed reaches to 40 km/h and 60 km/h. The impact coefficient increase with the increase of vehicle speed, when the speed exceeds 90 km/h. Impact coefficients under the case of no pavement roughness, A grade and B grade road are 1.064, 1.077, 1.142 respectively, when the speed is 40 km/h. The bridge pavement damage is more serious and the measured impact coefficient is 1.172 mentioned in the bridge inspection report, which shows that the worse the bridge deck roughness is, the greater the impact coefficient is.

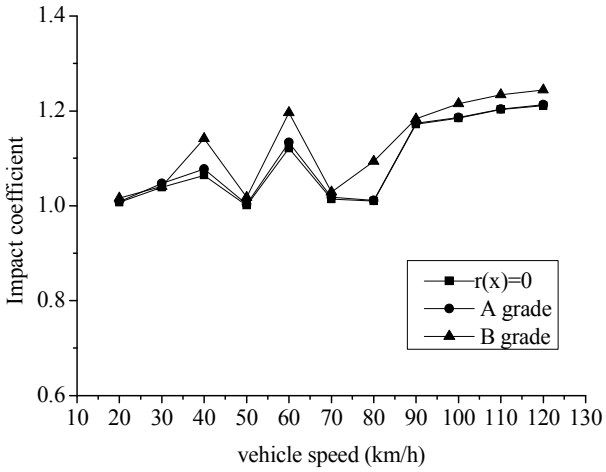


Fig.7 Impact coefficient under different vehicle speed

Comparison with Single Cable Suspension Bridge

Cable Shape Coefficient. The vertical coordinates of upper and lower cable of double cable suspension bridge is y_2^0, y_1^0, respectively, the vertical coordinates of upper and lower cable in mediate state is y_2, y_1, respectively, define the main cable shape coefficient:

$$\alpha = (y_1 - y_1^0)/(y_2^0 - y_1^0) = (y_2^0 - y_2)/(y_2^0 - y_1^0) . \qquad (5)$$

When $\alpha = 0$, it's double cable suspension bridge, when $\alpha = 0.5$, single cable suspension bridge, so single cable suspension bridge can be taken as a special case of double cable suspension bridge.

Comparison with Single Cable Suspension Bridge. The vibration response are compared between single cable and double cable suspension bridge at 40 km/h vehicle speed, the maximum dynamic displacements are shown in table 1. The vertical maximum displacements are close to when $\alpha = 0$ and $\alpha = 0.1$, but when $\alpha = 0.1$, the transverse maximum displacement is smaller. The maximum vertical displacement of the ¼ and ½ section are close to when $\alpha = 0.3$, which is most reasonable, and at this case the transverse maximum displacement is also smaller, which provides a dynamic basis to select reasonable structure form in design for this kind of suspension bridge.

Tab.1 Maximum dynamic displacement of stiffening girder (m)

Load condition	section	direction	$\alpha = 0$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.5$
center moving	*l*/4	vertical	-0.138	-0.14	-0.166	-0.234
	l/2	vertical	-0.234	-0.234	-0.176	-0.182
eccentric moving	*l*/4	transverse	-0.014~0.007	-0.010~0.006	-0.010~0.008	-0.014~0.010
	l/2	transverse	-0.023~0.014	-0.016~0.010	-0.016~0.012	-0.018~0.016

When vehicle speed is 40km/h, the comparison of vertical displacement-time curve under the center longitudinal axis moving load are shown in figure 3 and 4. The transverse displacement-time

curve under the eccentric longitudinal axis moving load are shown in figure 5 and 6.The maximum displacement of double cable suspension bridge is -0.234 m when vehicle moving to the ½ section, while the maximum displacement of ¼ section is -0.138 m; This is very different from the case that the maximum displacement of single cable suspension bridge is -0.234m occurred when vehicle moving to the ¼ section. All these show that the double-cabled system is superior to single-cable system in increasing vertical stiffness and resistance capacity to S-shaped deformation, but the vertical displacement of ½ section of double cable suspension bridge increases.

Summary

(1) The stiffening girder maximum vertical displacement of double cable suspension bridge occurs at vehicles moving to the middle section, which is very different from the maximum displacement of a single cable suspension bridge at vehicles moving to ¼ cross sections. Double-cabled system is superior to single-cabled system in increasing vertical stiffness and resistance to S-shaped deformation, which has a significant value on dynamic behavior assessment and strengthening design for this type of bridge.

(2) Stiffening girder vertical displacement fluctuations of double cable suspension bridge slow down with the increase of vehicle speed, while transverse displacement response fluctuations increase with increase of vehicle speed. The maximum deflection of span is not increases linearly with the increase of vehicle speed, the maximum deflection of span is relatively large when the speed is 40 km/h and 60 km/h; when the speed exceeds 90 km/h, the impact factor will increase with the speed increase.

(3) Surface roughness has an obvious influence on impact coefficient of double cable suspension bridge, the worse the bridge deck roughness is, the greater impact coefficient is. Therefore, the bridge maintenance work should be strengthened to reduce the impact on the bridge.

(4) Selecting a reasonable shape of the double cable will reduce the vertical and horizontal maximum displacement, which reflects the advantages of such bridge vibration characteristics, and it has guideline in bridge design and selection.

References

[1] LIN Li-xia, WU Ya-ping, DING Nan-hong: Influence of structure parameter on natural vibration characteristics of double cable suspension bridge. Journal of the china railway society Vol. 29(4) (2007), p. 91

[2] GUO Wei-wei, XIA He, XU You-lin: Dynamic response of long span suspension bridge and running safety of train under wind action. Engineering Mechanics Vol. 23(2) (2006), p. 103

[3] Ah med M Abded-Ghaffar: Vertical vibration analysis of suspension bridges. Journal of the Structural Division ASCE Vol. 106 (10) (1980), p. 2053

[4] Xia He, Xu Y L, Chan T H T: Dynamic interaction of long suspension bridges with running trains. Journal of Sound & Vibration Vol. 237(2) (2000) , p. 263

[5] XIAO Xin-biao, SHEN Huo-ming: Comparison of three models for vehicle-bridge coupled vibration analysis. Journal of southwest jiaotong university Vol. 39(2) (2004) , p. 172

[6] Tommy H.T. Chan, L.Guo, Z.X.Li: Finite element modeling for fatigue stress analysis of large suspension bridges. Journal of Sound and Vibration Vol. 261(3) (2003) , p. 443

[7] F.T.K Au, Y.S.Cheng, Y.K.Cheung: Effects of random road surface roughness and long-term deflection of prestressed concrete girder and cable-stayed bridges on impact due to moving vehicles. Comput. Struct Vol. 79(2001), p. 853

[8] Ding Nanhong, Lin Lixia, Sun Yingqiu: Theoretical and experimental study on highway arcade bridge's vibration under single vehicle load. Journal of Lanzhou J iaotong University Vol. 24(3) (2005), p. 28

©(2014) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.333

Optimize Fundamental Matrix Estimation Based on RANSAC

Jun Zhou

College of Computer and Information Science, Southwest University, Chongqing, P.R. China

zhouj@swu.edu.cn

Keywords: Fundamental matrix, robust estimation, RANSAC, matching

Abstract. Fundamental matrix estimation is a central problem in computer vision and forms the basis of tasks such as stereo imaging and structure from motion, and which is especially difficult since it is often based on correspondences that are spoilt by noise and outliers. The Random Sample Consensus (RANSAC) algorithm is a popular tool for robust estimation, primarily due to its ability to tolerate a tremendous fraction of outliers. In this article, we provide an approach for improve of RANSAC that has the benefit of offering fast and accurate RANSAC, and combine the M-estimation algorithm get the fundamental matrix. Experimental results are given that support the adopted approach and demonstrate the algorithm is a practical technique for fundamental matrix estimation.

Introduction

In computer vision, matching two relevant images photographed from the same scene is the basis of many applications including the recognition of objects and scenes, image-based 3D scene reconstruction, dynamic tracking, the estimation of the fundamental matrix is an important task in computer vision [1]. Generally image matching builds on the extraction of corresponding features from images, and epipolar geometric restriction relations between two images is the only information obtained from matched points, so epipolar geometric restriction can be used to match two images. Epipolar geometric depend on the internal parameters, the positions and orientations of cameras, it can be described by a singular fundamental matrix of which rank is two, and the problem of epipolar geometric could be transformed to the estimation problem of fundamental matrix [2], which could be helpful for matching images. The main difficulty stems from the unavoidable outliers inherent in the given matches. Most robust techniques require that the majority of matches be correct or else some form of outlier detection and removal is usually performed before actual parameter estimation.

The estimation of fundamental matrix has been a research hotspot, and its estimation accuracy will impact on the precision of further more works directly, which includes 3D reconstruction, motion estimation, camera calibration, matching, tracking, view synthesis, etc. Now the commonly used estimation algorithms mainly contain the linear algorithms, the iterative algorithms, and the robust algorithms. The 8-points algorithm proposed by Longuet-Higgins is a linear algorithm which is implemented fast and easily [3], but it is exceptionally sensitive to noise, and varies with the change of matched points. Hartley et al. proposed the improved 8-points algorithm that can enhance the stability of 8-points algorithm and reduce the disturbance of noise by normalization processing the selected matched points [4][5], but the calculated epipole position is unstable. The Above two algorithms neglect each couple of matched points has a different influence on fundamental matrix, and this assume results in a low estimation precision of fundamental matrix. The iterative algorithms mainly contains the algorithms calculating minimum geometric distances between matched points and epipolar lines and based-gradient methods [6][7]. They have an improvement in examination precision comparing to 8-points algorithm, but it is actually time-consuming, and can't eliminate the error matching.

The robust algorithms contains M-estimated algorithm [1][8], Least Median Squares algorithm (LMeds)[9], Random Sample Consensus (RANSAC)algorithm[10][11], etc. M-estimated algorithm solves fundamental matrix by the weighted iteration, but it really rests on the starting value and is influenced by the wrong matched points. LMeds algorithm estimates model parameters by

minimizing the square value of residual errors, which robustness is quite good. When the error rate of matched points will surpass 50%, M-estimated and LMeds algorithm aren't suitable. RANSAC algorithm is one of the most effective and robust estimation algorithms. Even if the error rate of matched points over 50%, it still can process effectively. RANSAC algorithm has a large range of applications in the estimation of fundamental matrix. In order to select the optimal relatively matched point set and reduce the influence of wrong matched points, we improve the RANSAC algorithm and combine the M-estimated algorithm, propose an algorithm of fundamental matrix estimation.

The paper is organized as follows: fundamental matrix is introduced in Section 2, Section 3 describes the improved RANSAC algorithm, our algorithm estimate the quality of the best found model and the experimental results is presented in Section 4. The paper concludes in Section 5.

Fundamental matrix

Consider two images I and I' of a common scene, Let C and C' represent the optical centre of the left and right cameras in the 3D coordinate, respectively. Points $\tilde{x}$ and $\tilde{x}_0$ (in homogeneous coordinates) are the projections of a 3D point x onto the left (I) and right (I') images, without loss of generality, we assume that the world coordinate system coincides with the left (first) camera coordinate system. It has been proven that

$$\tilde{x}'F\tilde{x} = 0 with F = A'^{-T}[t]_x RA^{-1} \quad (1)$$

where (R, t) is the rigid transformation (rotation and translation) which brings points expressed in the left camera coordinate system to the right one, and $[t]_x$ is the antisymmetric matrix defined by t such that $t * x = [t]_x x$ .The 3 * 3 matrix F is the fundamental matrix. Since $\det([t]_x) = 0$

$$\det(F) = 0 \quad (2)$$

Given a collection $\{x_1 x_2 .... x_n\}$ of image data and a meaningful cost function that characterizes the extent to which any particular F fails to satisfy the system of the copies of (1) associated with $x = x_i (i = 1, 2....n)$, find an estimate $\overline{F} \neq 0$ satisfying (2) for which the cost function attains its minimum. Since (1) and (2) do not change when F is multiplied by a nonzero scalar, $\overline{F}$ is to be found only up to scale. If the singularity constraint is set aside, then the estimate associated with a particular cost function $J = J(F; x_1, x_2 ..... x_n)$ is defined as the unconstrained minimizer $\overline{F}$ of J,

$$\overline{F} = \arg\min J(F; x_1 ..... x_n) \quad (3)$$

Improved RANSAC

RANSAC algorithm is an iterative method to estimate parameters of a mathematical model from a set of observed data which contains outliers. It is a non-deterministic algorithm in the sense that it produces a reasonable result only with a certain probability, with this probability increasing as more iteration are allowed. The RANSAC algorithm principle is that search M group samples in all data N, if the M big enough, we shall guarantee there is one group sample which only include inliers, and we use these M samples data set to evaluate model parameters, then get the best model parameters through a certainty criterion, on the other sides, the best model parameters can get from sample data set which not include outliers. Finally, we hypothesis the best model parameters as the model which is we look for, and use the model filter the other data in order to find all the inliers in the data N, in the same way, we can use these inliers evaluate the model parameters.

The speed of RANSAC depends on two factors. First, the percentage of outliers determines the number of random samples needed to guarantee a given confidence in the optimality of the solution. Second, the time needed to assess the quality of hypothesized model parameters is proportional to the number N of input data points. In order to enhance the efficiency of algorithm, we need to know the time of the algorithm consume. We hypothesis randomly sample a group data set from all data N need time t_s , and use this group sample data set compute the model parameters need the time t_c , use one

data which got from all data N verify model parameters need the time t, so use N number data verify model parameters need the time Nt. Form above, the time RANSAC algorithm are:

$$T = M\left(t_s + t_c\right) + MNt \tag{4}$$

In the formula (4), $M\left(t_s + t_c\right)$ are the sample M groups data samples set consume time, MNt are verify M model parameters consume time. From formula (4), we can get two factors determine the time of RANSAC algorithm consumed: one is the number of sample M, which values affected by the model complex, data error ratio and confidence probability. The other is time of verify model parameters, which are determined by the number of model parameters and the number of all data.In the RANSAC algorithm, t_s and t_c are commonly small, and t_s, t_c, Nt have relation with the model parameters, they have got different values if the model parameters change. In another way, we need to compute each data error to verify model parameters, and compute the square error. So, the Nt is a big time at the RANSAC algorithm consume, if we want to improve algorithm velocity, there are two choice realize it, one is reduce the number of sample M, the other is reduce the time verify model parameters need consume.

In optimized model verification has focused on minimizing the number of data points tested in the verification step of RANSAC. These techniques make use of the observation that most of the models hypothesized in RANSAC are likely to be contaminated, and are consistent with only a small fraction of the data points. Thus, it is often possible to discard bad hypotheses early on in the verification process. In particular, an optimal randomized model verification strategy is developed, based on Wald's theory of sequential decision making. The evaluation step is cast as an optimization problem which aims to decide whether a model is good (Hg) or bad (Hb), while simultaneously minimizing the number of verifications performed per model, as well as the probability that a good model is incorrectly rejected early on. Wald's equential Probability Ratio Test (SPRT) is based on the likelihood ratio

$$\lambda_j = \prod_{r=1}^{j} \frac{p(x_r \mid H_b)}{p(x_r \mid H_g)} \tag{5}$$

The core idea of the preview verify of RANSAC algorithm is that most evaluated model hypotheses are influenced by outliers. Many model parameters are affected by the outliers, the procedure of preview verify only need random select little datum, we can use this data verify the model parameters which is impossible became final result. From this method, we can filter a lot of wrong model parameters. Increase the number of samples M, insure at least one right model parameter can go through the preview verify, the datum which was used compute model parameters should be all the inliers. In order to reduce the number of verify model parameters, we randomly sample little data do preview verify, these data can take part in all data verify if data has pass the preview verify. In the procedure of algorithm, the number of inliers which are taken as quality criterion of verifying model parameters are independent of verify sample numbers, in other words, this criterion is correct under the condition of little samples.

Our Algorithm and Experimental Results

Our concrete step of algorithm as fellows:

Step1: use the improve 8-point algorithm compute fundamental matrix, and get the first values.

Step2: use the improve RANSAC get the right inliers, and rid of outliers.

Step3: based on the right inliers, use the M-estimation algorithm to compute the fundamental matrix.

The proposed algorithm was tested on many image pairs of indoor and outdoor scenes, several of which are presented here. The cases that are presented here are difficult cases in which the inliers rate is low and includes a dominant degeneracy. In our experimental, we contrast our algorithm and the

RANSAC algorithm, and the results indicate our algorithm is robust and accurate than the RANSAC algorithm.

For fundamental matrix estimation, the time needed to compute model parameters was set within the range observed in a large number of experiments (that is, in our implementation, checking whether a correspondence is consistent with a fundamental matrix is 200 times faster than estimating the matrix). The exact value depends on many factors including the CPU speed and type. The constant was set to the experimentally observed average of the number of models generated by the 8-point algorithm per sample. The initial values of η_0 and ε were set to η_0=0.05 and ε=0.2, respectively. Image pairs with experiments see the Fig.1. To demonstrate the performance of the improved algorithm, the algorithm and RANSAC was applied to the two pairs of images shown in Fig. 2.

(a) left (b) right

Fig.1 image pairs with experiments (a)left, (b)right

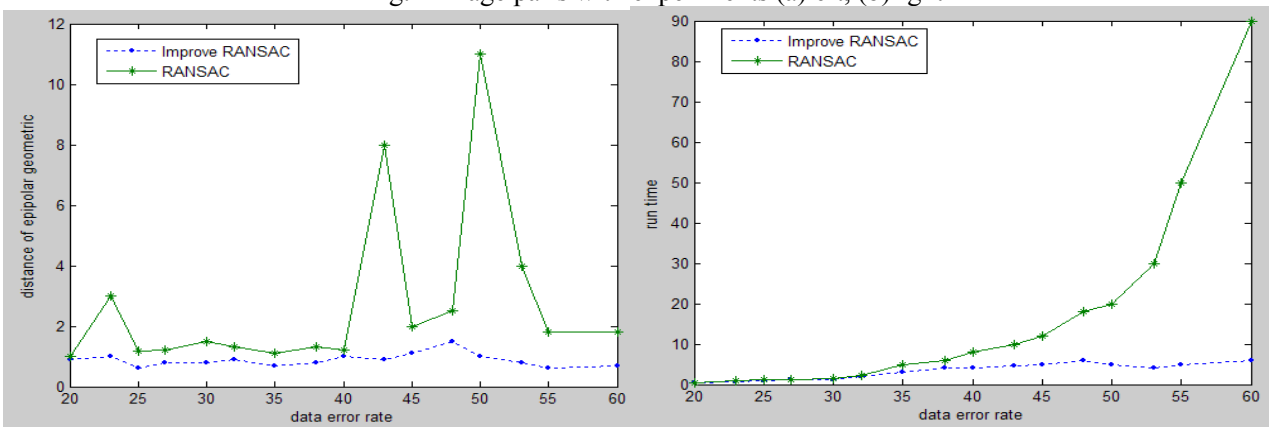


Fig2. Performance of improved algorithm and RANSAC

Conclusion and future Work

In this paper, we propose a improve RANSAC algorithm for robust estimation of fundamental matrix. The algorithm handles the above two difficult cases in a unified manner. The algorithm can handle not only the planar degeneracy but also scenes that include a variety of degeneracies or close to degenerate configurations. In order to enhance the accurate of RANSAC, we take the local optimize random sample process, and combine M-estimation algorithm to compute the fundamental matrix. In paper, we give the improve RANSAC algorithm basic theory and the concrete process.

In the future work, we shall do how to get the accurate, fast and robust fundamental matrix estimation algorithm. In recent years, numerous methods have been derived from RANSAC and form their family, but these algorithms keen on settle one aspect of above proposed, there is a lot of work to resolve this. And in this paper, the results show that our method can reduce the influence by noise and

error matching, and improve the precision of fundamental matrix, so it prepares well for matching other points in image and further works.

Acknowledgments

We would like to thank everyone who spent time on reading early versions of this paper, including the anonymous reviewers. The work is supported by "the Fundamental Research Funds for the Central Universities" XDJK2009C019.

References

[1] P.H.S.Torr and D.W.Murray : The development and comparison of robust methods for estimating the fundamental matrix, International Journal of Computer Vision ,vol. 24 (3)(1997) , p.271

[2]R. Hartley and A. Zisserman: Multiple View Geometry, Cambridge University Press(2000)

[3]H.C.Longuet-Higgins: A Computer Algorithm for Reconstructing a Scene from Two Projections, Nature, vol. 293(1981), no.10, p.133

[4] R.I. Hartley: Estimation of Relative Camera Positions for Uncalibrated Cameras, Proc. Second European Conf. Computer Vision, Lecture Notes in Computer Science, vol. 588(1992), p. 579

[5] R. Hartley: In Defense of the Eight-Point Algorithm, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19(1997), no. 6, p.580

[6] Z.Y. Zhang: Determining the epipolar geometry and its uncertainty: A review, Tech. Rep. 2927, Institut National de Recherche en Informatique et en Automatique(1996)

[7] Q.T. Luong and O.D. Faugeras: Determining the Fundamental matrix with planes: instability and new algorithms, Proceedings, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (1993), CVPR(93), p.489

[8] H. Chen and P.Meer: Robust Regression with Projection Based M-Estimators, Proc. IEEE Int'l Conf. Computer Vision (2003), p.878

[9] E. Malis and E. Marchand: Experiments with robust estimation techniques in real-time robot vision, IEEE/RSJ International Conference on Intelligent Robots and Systems (2006), p.223

[10] O. Chum, J. Matas, and J. Kittler: Locally Optimized RANSAC, Proc. Ann. Symp. German Assoc. for Pattern Recognition (2003)

[11] M. Fischler and R. Bolles: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, Comm. ACM, vol. 24(1981), no. 6, p. 381

doi:10.4028/www.scientific.net/AMM.50-51.338

Chasing Models under Different Conditions in Anti-Smuggling Problem

Jianli Cao and XiuLan Guo

College of Science, Henan University of Technology, Zhengzhou, Henan, 450001, China

cao_j_l@163.com

Keywords: Chasing Model; Anti-Smuggling Problem; Mathematica software

Abstract. Chasing problem is not only a common problem in reality including the classical Turtle Rabbit Race problem, but also a representative mathematical model which has widely application. In this paper, we study chasing models in some cases appearing in anti-smuggling problem. Different conditions are given to restrict the escaping and chasing way, thus more common situations are discussed and some new meaningful results are obtained.

Introduction

In the latest few years, mathematics has been applied more and more widely with the progress of science and technology, especially the improvement of computer. Practical problem in our life can be analyzed quantitatively and solved accurately by establishing appropriate mathematical model such as biology, population, radioactive, economics model [1-5].

This article focuses on chasing model in anti-smuggling problem which can be solved by differential equation method [6-10] or approximation method [11-15]. The anti-smuggling problem can be seen in many books in which the smugglers escaping along a fixed angle [1,6]. Therefore, the model has some limitations. In reality, we can not know in advance the direction that smugglers will escape. The angle is measured when capturers find smugglers and begin to chase them. Hence, this problem can be generalized by supposing the angle as a variable θ, thus a more general problem is discussed.

The Mathematical Model of Chasing Line

Model I. If the two directions of smugglers and capturers are vertical, suppose the initial position of smugglers and capturers are at $(0,0)$ and $(c,0)$, a, b are the velocities of smugglers and capturers. After establishing proper coordinate system, we denote x, y as the displacement of smugglers and capturers respectively. Through direct inference, the initial problem of chasing line is

$$\begin{cases} xy'' = k\sqrt{1+(y')^2} \\ y(c)=0, y'(c)=0 \end{cases} \tag{1}$$

where $k=\dfrac{a}{b}$.

Solving the problem resorting to basic calculus knowledge, we get:

(i) If $k<1$, the displacement of capturers satisfies

$$y=\frac{c}{2}\left(\frac{1}{1+k}\left(\frac{x}{c}\right)^{1+k}-\frac{1}{1-k}\left(\frac{x}{c}\right)^{1-k}\right)+\frac{ck}{1-k^2}. \tag{2}$$

Substituting x=0 into (2), we obtain the distance to catch up with smugglers is

$$y=\frac{abc}{b^2-a^2} \tag{3}$$

and the chasing time is

$$t=\frac{y}{a}=\frac{bc}{b^2-a^2}. \tag{4}$$

(ii) If $k=1$, the displacement of capturers satisfies

$$y=\frac{1}{2}\left(\frac{x^2-c^2}{2c}-c\ln\frac{x}{c}\right). \tag{5}$$

Since x can not equal 0 in (5), it means capturers can not catch up with smugglers.

(iii) If $k>1$, It is obviously that capturers can not catch up with smugglers.

Taking $c=200$, $k=0.78$ and using Mathematica software, we can get the figure of capturers displacement as below:

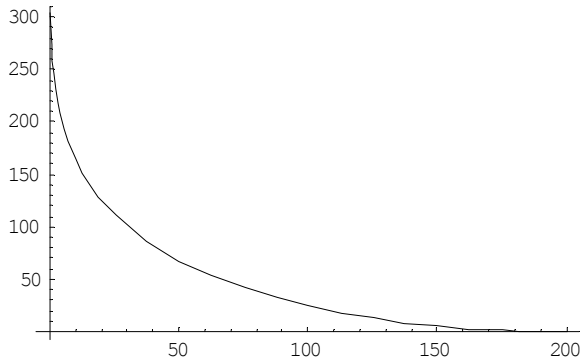


Figure1 The figure of capturers displacement

Model II. If smugglers escape along the direction which has an angle θ with the positive direction of x-axis, through direct inference, the differential equation of chasing line is

$$\frac{a}{b}\sqrt{1+(y')^2}=\frac{xtg\theta-y}{tg\theta-y'}y''. \tag{6}$$

Solving this equation with initial condition, we get:

(i) If $a\neq b$, the displacement of capturers satisfies

$$y=\frac{bc}{2(a+b)}\left(\frac{x}{c}\right)^{\frac{a+b}{b}}-\frac{bc}{2(b-a)}\left(\frac{x}{c}\right)^{\frac{b-a}{b}}-\frac{abc}{a^2-b^2}. \tag{7}$$

In this case, capturers can catch up with smugglers and the chasing time satisfies

$$at\sin\theta=\frac{bc}{2(a+b)}\left(\frac{at\cos\theta}{c}\right)^{\frac{a+b}{b}}-\frac{bc}{2(b-a)}\left(\frac{at\cos\theta}{c}\right)^{\frac{b-a}{b}}+\frac{abc}{a^2-b^2}. \tag{8}$$

(ii) If $a=b$, the displacement of capturers satisfies

$$y=\frac{x^2}{4c}-\frac{c}{2}\ln x+\frac{c}{2}\ln c-\frac{c^2}{4}. \tag{9}$$

Obviously capturers can not catch up with smugglers since x can not equal 0 in (9).

Model III. If the direction of capturers has an angle θ with the positive direction of x-axis, the initial problem of chasing line is

$$\begin{cases} xy'' = k\sqrt{1+(y')^2} \\ y(c\cos\theta) = c\sin\theta, \\ y'(c\cos\theta) = tg\theta \end{cases} \tag{10}$$

where $k = \frac{a}{b}$.

Solving the initial problem, we get:

(i) If $k = 1$, the derivative of capturers displacement is

$$y' = \frac{1}{2}\left[\frac{x}{c_1} - \frac{c_1}{x}\right], \tag{11}$$

where $c_1 = c\cos\theta\left(\frac{\cos\theta}{1+\sin\theta}\right)^{\frac{1}{k}}$. Obviously $x \neq 0$, so capturers can not catch up with smugglers.

(ii) If $k < 1$, the displacement of capturers satisfies

$$y = \frac{c_1}{2}\left[\frac{1}{k+1}\left(\frac{x}{c_1}\right)^{k+1} - \frac{1}{1-k}\left(\frac{x}{c_1}\right)^{1-k}\right] + c_2, \tag{12}$$

where $c_2 = \frac{ck(1-k\sin\theta)}{1-k^2}$. Obviously, capturers can catch up with smugglers. The distance to catch up with smugglers is

$$y = \frac{abc - a^2 c\sin\theta}{b^2 - a^2} \tag{13}$$

and it needs time

$$t = \frac{bc - ac\sin\theta}{b^2 - a^2}. \tag{14}$$

(iii) If $k > 1$, through calculation, we know that capturers can catch up with smugglers when $\theta \in (\arcsin\frac{b}{a}, \frac{a}{2})$ and the distance to catch up with smugglers is the same as (13), and the chasing time is the same as (14).

Model IV. If smugglers set out with an initial speed along the positive direction of y-axis while capturers are at a certain point $P_0(x_0, y_0)$, in the chasing process, the direction of capturers always point to smugglers. Through reasonable analysis, the initial problem of chasing line is as follows

$$\begin{cases} xy'' = k\sqrt{1+y'^2} \\ y(x_0) = y_0 \\ y'(x_0) = \dfrac{y_0}{x_0} \end{cases}, \tag{15}$$

where $k = \dfrac{a}{b}$.

Solving this initial problem, we can get:

(i) If $k < 1$, the displacement of capturers satisfies

$$y(x) = y_0 + \frac{C_1}{2(1+k)}(x^{1+k} - x_0^{1+k}) - \frac{C_1}{2(1-k)}(x^{1-k} - x_0^{1-k}), \tag{16}$$

where $C_1 = \dfrac{y_0 + \sqrt{x_0^2 + y_0^2}}{x_0^{1+k}}$.

In this case, capturers can catch up with smugglers. The distance to catch up with smugglers is

$$y = a(b\sqrt{x_0^2 + y_0^2} - ay_0)/(b^2 - a^2) \tag{17}$$

and the chasing time is

$$t = (b\sqrt{x_0^2 + y_0^2} - ay_0)/(b^2 - a^2). \tag{18}$$

(ii) If $k = 1$, the displacement of capturers satisfies

$$y(x) = y_0 + \frac{C_1}{4}(x^2 - x_0^2) - \frac{1}{2C_1}\ln\frac{x}{x_0}. \tag{19}$$

Since x can not equal 0 in (19), capturers can not catch up with smugglers.

(iii) If $k > 1$, the displacement of capturers satisfies

$$y(x) = y_0 + \frac{C_1}{2(1+k)}(x^{1+k} - x_0^{1+k}) - \frac{C_1}{2(1-k)}(x^{1-k} - x_0^{1-k}). \tag{20}$$

Consider the limitation when x approaches 0, we have

$$\lim_{x\to 0} y(x) = +\infty. \tag{21}$$

Obviously capturers can not catch up with smugglers.

Summary

Due to the complexity of the actual situation in the model, some reasonable assumptions are given in advance. Based on these assumptions, we give different solutions according to different problems. Meaningful results show that capturers can catch up with smugglers if the velocities of them satisfy certain expression.

Acknowledgment. This project is supported by the Funds for Basic Research Project under Grant 10971200, 102300410118, 2008B110003, 2008GJYJ-C42, 2009GJYJ-A19 and 2009GGJS-060.

References

[1] Qiyuan Jiang: Mathematical Model (Mechanical Industry Press, Beijing 2003)

[2] Zhonggeng Han: Mathematical Modeling Method and its Application (Higher Education Press, Beijing 2005)

[3] Zhenhang Liu: Mathematical Modeling (Chinese People's University Press, Beijing 2004)

[4] Jianli Cao, Wantang Jiao and Landong Hou: A Nonlinear Regression Model of Rearing Cattle in Biostatistics Proceedings of the 6th Conference of Biomathematics, Vol.1(2008), p. 241-244

[5] Jianli Cao: Research on China's Population Trend Based on UN Population Data, Conference Proceedings of 2009 International Institute of Applied Statistics, Vol.2(2009), p. 1660-1666

[6] Yunfeng Gao: Small Problems, Mechanics in Engineering, Vol.20(1998), p.78-80

[7] Zhengguo Zhao: More Intuitive Solution for Chasing Problem. Mechanics in Engineering (2000), p.68~69

[8] Chengyin Mao: The Proof of Oval Critical Range in Chasing Problem. Mechanics in Engineering (2000), p.77~80

[9] Fei Tan: Application of Differential Equation in Mathematics Modeling Documentation Analysis, Journal of Lianyungang College of Chemical Technology, Vol.15, no.2(2002)

[10] Linchuan Xiang: The Numerical Solution of a Pursuant Trajectory Problem with Generalization, College Physics,Vol.23,no.8(2004), p.24-26

[11] Xiaoping Li: A Simple Proof and Annotations About Critical Region of Pursue Problem, Mathematics in Practice and Theory,Vol.35,no.1(2005), p.222-223

[12] Weigang Qiu, Mingming Jiang, Hong Yang: On Critical Range in Chasing Problem, Mechanics in Engineering (2007), p.73

[13] Hua Zhou: Research on Complex Technology of Chasing Path in Large Aircraft Project. Defense Science and Technology (2007), p.36~38

[14] Qingxin Yuan, Yongwei Zhou: The Mechanical Solution to Chasing Problems of Many Dogs, Mechanics in Engineering (2010)

© (2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.343

Coupling Acoustic Analysis of Enclosure Consisting of Multi-Flexible Plates and Its Active Structural Acoustic Control

Gu Linning[1, a], Chen Nan[1, b]

[1]School of Mechanical Engineering, Southeast University, Nanjing, China

[a]aku_linney@163.com, [b]nchen@seu.edu.cn

Keywords: coupled system, acoustic mode, structural mode, ASAC, noise reduction.

Abstract. This paper develops a model of structural-acoustic coupled system of a rectangular enclosure involving two simply supported flexible plates. The acoustical potential energy and the optimal secondary control forces are derived theoretically. An example is presented to analyze the characteristics of the coupled system, the dominant factors that impact the coupling effects and the noise reduction with ASAC method. The paper shows the possibility of applying distributed cooperative control method to the structural-acoustic coupled system.

Introduction

Enclosed structure with thin walls is typically used in many engineering areas, such as submarine hull, various carrying tools and shells of aircraft cabin, etc. Both theorists and engineers [1~2] are paying more attentions to the solution of the structural-acoustic coupled system. Wu[3] has solved the interior acoustic field with the concept of covering domain. Xie[4] has researched on the active structural acoustic control to a simple enclosure. And Yao[5] has analyzed the impact of different boundary conditions to the interior acoustic field.

In this paper, a model of structural-acoustic coupled system of a rectangular enclosure consisting of two simply supported flexible plates and four rigid plates is established. And the acoustical potential energy and the optimal secondary control forces are derived. Then an example is presented to analyze the characteristics of the coupled system, and the noise reduction of the Active Structural Acoustic Control, which is to lay foundation for deep analysis of dominant factors that impact it and to compare the control effects to dominant factors and noise reduction with different secondary forces applied on the flexible plates. At last, the possibility of applying distributed cooperative control method to the structural-acoustic coupled system is studied.

Structural-acoustic Coupling Theory Derivation

Theoretical Analysis Model. Consider a rectangular enclosure consisting of two simply supported flexible plates and four rigid plates, as shown in Fig. 1. The dimensions of the cavity are $L_x \times L_y \times L_z$, with the volume V, and the areas of two flexible plates respectively are A_a and A_b, with the thickness h. The density of the plate and air respectively are ρ_1 and ρ. It is assumed that the noise of the structural-acoustic system is induced by a point force excitation $f(\mathbf{r}, t)$ on Plate a or b.

Acoustroelasticity Theory. According to Modal Synthesis Method**,** the interior acoustic pressure and the normal vibrating displacement of Plate a and b could be written as [6]

$$p=\sum_{n=1}^{N}\psi_n p_n\,,\quad w_a=\sum_{r=1}^{R}\phi_r^a q_r^a\,,\quad w_b=\sum_{s=1}^{S}\phi_s^b q_s^b\,, \tag{1}$$

where $\psi_n(\vec{r})$, $\phi_r^a(\vec{r})$ and $\phi_s^b(\vec{r})$ are mode shape functions of a rigid-welled rectangular enclosure, structural mode shape functions of Plate *a* and *b*, respectively. Due to the orthogonality, one has

$$\frac{1}{V}\int_V \psi_n\psi_m dv=M_{1n}\delta(n-m)\,,$$

$$\int_A \rho_1 h\phi_m\phi_s dA=M_{2m}\delta(m-s)\,. \tag{2}$$

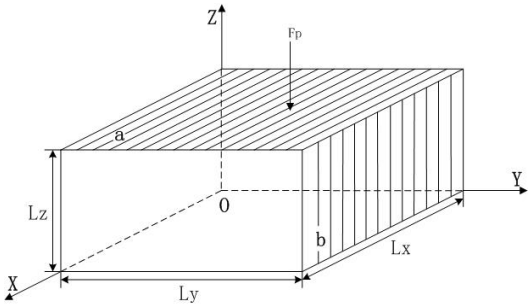


Fig. 1. A structural-acoustic coupled system with structural excitation *f(**r**, t)*

According to the wave equation and boundary conditions, and ψ_n meets Green's formula, thus

$$\ddot{p}_n+\omega_n^2 p_n=-\frac{\rho c^2}{VM_{1n}}\left(A_a\sum_r C_{nr}^a\ddot{q}_r^a+A_b\sum_s C_{ns}^b\ddot{q}_s^b\right). \tag{3}$$

Structural Mode Theory. According to structural mode theory, one has

$$\ddot{q}_m(t)+\omega_m^2 q_m(t)=\frac{A}{M_{2m}}\sum_n p_n C_{nm}-\frac{\int_A f(\vec{r},t)\phi_m dA}{M_{2m}}. \tag{4}$$

In this paper, it is assumed that there is only one primary force excitation located on Plate *a* or *b*. In this section it is assumed that there are *f(**r**$_a$, t)* and *f(**r**$_b$, t)* exciting Plate *a* and *b* respectively just for the derivation. The vibration equations of Plate *a* and *b* can be expressed as

$$\ddot{q}_r^a+(\omega_r^a)^2 q_r^a=\frac{A_a}{M_{2r}^a}\sum_n p_n C_{nr}^a-\frac{Q_{ar}}{M_{2r}^a}\,,\quad \ddot{q}_s^b+(\omega_s^b)^2 q_s^b=\frac{A_b}{M_{2s}^b}\sum_n p_n C_{ns}^b-\frac{Q_{bs}}{M_{2s}^b}\,, \tag{5}$$

where $Q_{ar}=\int_{A_a} f_a(\vec{r}_a,t)\phi_r^a dA_a$, $Q_{bs}=\int_{A_b} f_b(\vec{r}_b,t)\phi_s^b dA_b$ are defined as generalized modal forces.

Primary Acoustic Field. When the excitation and responses are both simple-harmonic, considering the damping, it is available to utilize Fourier Transform Algorithm to convert Eq. (3) and (5) into

$$p_n(\omega)=\frac{h_n\rho c^2\omega^2}{VM_{1n}}\left[A_a\sum_r C_{nr}^a q_r^a(\omega)+A_b\sum_s C_{ns}^b q_s^b(\omega)\right], \tag{6}$$

$$q_r^a(\omega)=\frac{h_r^a A_a}{M_{2r}^a}\sum_n p_n(\omega)C_{nr}^a-\frac{Q_{ar}(\omega)}{M_{2r}^a}\,,\quad q_s^b(\omega)=\frac{h_s^b A_b}{M_{2s}^b}\sum_n p_n(\omega)C_{ns}^b-\frac{Q_{bs}(\omega)}{M_{2s}^b}\,, \tag{7}$$

where $h_n=\dfrac{1}{\omega_n^2+j2\xi_n\omega_n\omega-\omega^2}$, $h_r^a=\dfrac{1}{(\omega_r^a)^2+j2\xi_r^a\omega_r^a\omega-\omega^2}$, $h_s^b=\dfrac{1}{(\omega_s^b)^2+j2\xi_s^b\omega_s^b\omega-\omega^2}$.

Then Eq. (6) and (7) can be written in the matrix form as

$$\{P\}=[Z^a]\{Q^a\}+[Z^b]\{Q^b\}\,,\quad \{Q^a\}=[H^a]\{P\}-\{QF^a\}\,,\quad \{Q^b\}=[H^b]\{P\}-\{QF^b\}\,. \tag{8}$$

Thus, it is available to obtain

$$\{P\}=\left[\left[Z^a\right]\left[H^a\right]+\left[Z^b\right]\left[H^b\right]-I\right]^{-1}\left(\left[Z^a\right]\left\{QF^a\right\}+\left[Z^b\right]\left\{QF^b\right\}\right). \tag{9}$$

It is assumed that there is only one primary force excitation (assuming on Plate *a*), thus

$$\{P_p\} = \left[\left[Z^a\right]\left[H^a\right] + \left[Z^b\right]\left[H^b\right] - I\right]^{-1}\left[Z^a\right]\{QF^a\}. \tag{10}$$

Active Structure Acoustic Control. In order to reduce noises based on active structural acoustic control method, there must be several secondary forces $fs_i(\boldsymbol{r}_i,t)$ (i=1, 2, …, L) exciting Plate *a* to produce secondary acoustic field, and its acoustic modal coefficient vector can be given by

$$\{P_s\} = \left[\left[Z^a\right]\left[H^a\right] + \left[Z^b\right]\left[H^b\right] - I\right]^{-1}\left[Z^a\right][QL]\{fs\} = [A]\{fs\}, \tag{11}$$

where $QL_{ij} = \dfrac{h_i^a \int_{A_a} \delta(\vec{r} - \vec{r}_j)\phi_i^a dA_a}{M_{2i}^a}$, $\{fs\} = \{fs_1 \quad fs_2 \quad \cdots \quad fs_L\}$. Thus, the total acoustic pressure is

$$p_{总}(\vec{r},\omega) = \left(\{P_p\} + \{P_s\}\right)^T \{\psi\}. \tag{12}$$

Then the total time averaged acoustic potential energy can be given by

$$E_p = \frac{1}{4\rho c^2}\int_V \left|p_{总}(\vec{r},\omega)\right|^2 dv = \frac{V}{4\rho c^2}\left(\{B\} + [D]\{fs\}\right)^H \left(\{B\} + [D]\{fs\}\right), \tag{13}$$

where $\{B\} = diag\left\{\sqrt{M_{11}} \quad \sqrt{M_{12}} \quad \cdots \quad \sqrt{M_{1N}}\right\}\{P_p\}$, $[D] = diag\left\{\sqrt{M_{11}} \quad \sqrt{M_{12}} \quad \cdots \quad \sqrt{M_{1N}}\right\}[A]$. Thus, it is easy to get an optimal $\{f_s\}$ that produces the minimal acoustical potential energy

$$\{fs\}_{opt} = -[D^H D]^{-1}[D]^H \{B\}. \tag{14}$$

Analysis Cases

The model has the same parameters as model of Chen [6], listed in Table 1.

Table 1. Material properties of the model.

c=344m/s	E=68.5GPa	υ=0.33	$\eta_1=\eta_0$=0.01	ρ_1 =2700kg/m^3	ρ=1.21kg/m^3
L_x=0.868m	L_y=1.150m	L_z=1.0m	h=0.006m	F_p=10N	

To check the validity of the previous derivation, compare it with the results of Chen [6]. Calculate the frequency domain responses of acoustic potential energy (APE), shown in Fig. 2. It can be seen that there is good agreement with the results of Chen [6], and it reflects the frequency-domain characteristics of the coupled acoustic field, except for little differences at some frequencies due to damping assumption and ignorance.

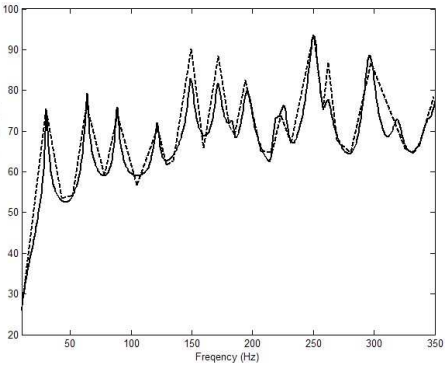
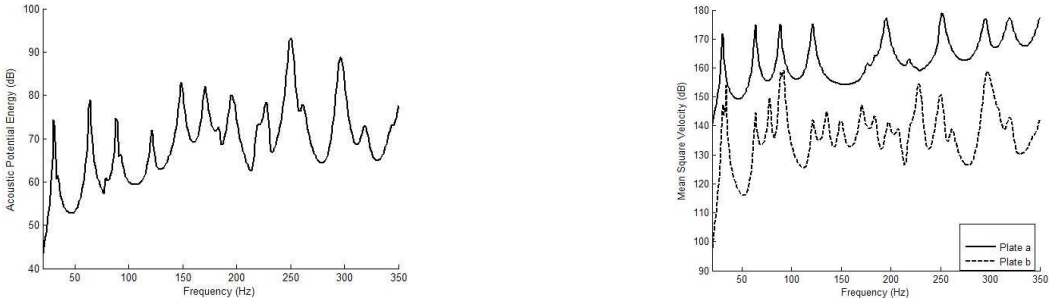


Fig. 2. APE responses (—) and the result of Chen [6] (---)

The coupled system shown in Fig. 1 is excited by a primary force located on (0.3, 0.4, 1.0) on Plate *a*. The calculated frequency domain responses of APE and the surface mean square velocity (MSV) of Plate *a* and *b* are shown in Fig. 3, respectively.

In Fig. 3 (a), it can be seen that the most APE peaks are related to the acoustic natural frequencies and the structural natural frequencies of Plate *a*, except for 34Hz, 78Hz and 92Hz where are decided by the natural frequencies of Plate *b*; however their peak-amplitudes is 10 dB less than those at other peak frequencies. In Fig. 3 (b), the MSV of Plate *a* which is excited directly is 20 dB higher than that of Plate *b*. The most MSV peaks of Plate *a* are decided by its natural frequencies, with a few peaks impacted by acoustic natural frequencies, but none of Plate *b*.



(a) Acoustic potential energy responses (b) Mean square velocity responses

Fig. 3. The coupled system responses to a point force excitation located on Plate a

When the coupled system is excited by the primary force located on (0.492, 1.150, 0.433) on Plate *b*, the frequency domain responses of APE and the MSV of Plate *a* and *b* are shown in Fig. 4.

(a) Acoustic potential energy responses (b) Mean square velocity responses

Fig. 4. The coupled system responses to a point force excitation located on Plate *b*

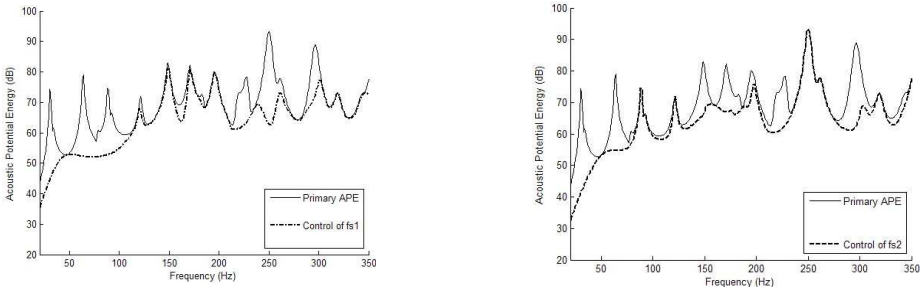
Similarly, it can be found that the energy transfer path is: Plate *b* to acoustic cavity, acoustic cavity to Plate *a*, and feedback to acoustic cavity from Plate *a* at some lower frequencies, such as 64Hz, 119Hz and 184Hz, obviously it is caused by the coupling between acoustic modes and structural modes. But the feedback energy is not transferred to Plate *b*, and this weak coupling effect is not playing a leading role compared to Plate b which is excited directly.

Generally, both natural frequencies of Plate *a* and *b* influence the acoustic responses, but the plate excited directly and acoustic modes play a leading role. Especially when acoustic frequency is close to structural natural frequency, it will lead to strong coupling effect, such as 249Hz and 295Hz at Fig. 3(a), as well as 229Hz and 303Hz at Fig. 4(a), where peak amplitudes are 10 dB higher than others. It can be concluded that the 10^{th} and 12^{th} structural mode of Plate *a* play the most leading role when Plate *a* is excited directly, and the 8^{th} and 10^{th} structural mode of Plate *b* play the most leading role when Plate *b* is excited directly. Thus, ASAC method can be applied to control the most leading structural modes according to different excitations, to restrain the plate vibration, to lower the acoustic radiation of flexible plates, so as to achieve the ultimate goal of noise reduction.

When apply a secondary force f_{s1} located on (0.651, 1.006, 1.0) where the 10^{th} structural mode vibrates within maximum amplitude when Plate *a* is excited directly, and Fig. 5(a) shows the control effect. It is seen that the interior noise is reduced at majority of frequencies, especially at 249 Hz the noise reduction reaches almost 30 dB, which exactly indicates that f_{s1} controls the 10^{th} structural mode of Plate *a* efficiently. Besides, the noise reduction at some other peak-frequencies reaches 20 dB, such as 64Hz and 228Hz. When apply a secondary force f_{s2} located on (0.434, 0.805, 1.0) where the 12^{th} mode vibrates within maximum amplitude, shown at Fig. 5(b).

Comparing Fig. 5(a) and (b), it can be found that f_{s2} can increase the noise reduction at 120 – 180Hz which is 'control blind area' of f_{s1}, and f_{s1} also can cause obvious noise reduction at 70-110 Hz and 230 -260 Hz where f_{s2} cannot. Thus, it can be included that f_{s1} and f_{s2} have

complementary effect on the noise reduction. Fig. 6 shows the simultaneous control effect of two secondary forces.



(a) The control effect of f_{s1} (b) The control effect of f_{s2}

Fig. 5. The control effect of ASAC when Plate a is excited directly

Generally, the simultaneous control effect is better than that of one, especially at the natural frequencies out of control of f_{s1} and $f_{s2,}$ but the increase of noise reduction can reach 10 dB merely at some frequency range, and at other frequencies there is nearly no increase, especially the natural frequencies at which the modes are under the control of f_{s1} or $f_{s2.}$

Conclusions

It is concluded that:

1) The flexible plate excited directly plays a leading role in the coupled system, and there is only need to control its most leading modes to reduce the noise with ASAC method.
2) When a secondary force is located where the N^{th} structural mode vibrates within maximum amplitude, then the noise dominated by the N^{th} mode will be reduced greatly. Moreover, it can reduce noises induced mainly by other structural modes in different degrees.
3) When noises cannot be reduced by single secondary forces, simultaneous control of several secondary forces will make the situation better. But it cannot reduce the noise components at every frequency, especially components which are at the modes controlled by single secondary forces. It is to say, there is no need to adopt the simultaneous control all the time.

The above conclusions lay foundation for applying distributed cooperative control method to the structural-acoustic coupled system containing several flexible plates. Deeper research is under way.

Acknowledgement

This work is supported by China National Natural Science Fund (No. 50975047).

References

[1] Y. Y. Li, L. Cheng. Active noise control of a mechanically linked double panel system coupled with an acoustic enclosure [J]. Journal of Sound & Vibration, 2006, 297(3-5): 1068-1074

[2] T. Habib, M. Kepesi. Open issues of active noise control application[C]. Radioelektronika 2007, 17th Inter. Conf, 24-25 April 2007: 1-4

[3] J. H. Wu, H. L. Chen, X. L. Hu. A new method to calculate interior sound field of arbitrary-shaped closed thin shell [J]. Chinese Journal of Acoustics, 2000, (5): 29-34

[4] J. L. Xie, N. Chen, B. L. Zhong. Study on forces control of structural-borne sound in enclosed space [J]. Journal of Vibration Engineering, 1999, 12(1): 15-20

[5] H. P. Yao, J. R. Zhang. Analysis of structural-acoustic coupling of elastic rectangular enclosure with arbitrary boundary conditions [J]. Chinese Journal of Acoustics, 2007, 32(6): 497-501

[6] K. A. Chen, L. Y. Ma. Active Noise Control- Principles, algorithms and implementation [M]. Xi'an: Northwestern Polytechnical University Press, 1993: 101- 111

Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.348

Study of Dry Sliding Friction and Wear Characteristic of High-speed Steel/ Ti_6Al_4V Alloy under Direct Steady Magnetic Field

WEI Yong-hui[1,2,3,a], ZHANG Yong-zhen[1,2,3,b], CHEN Yue[2,3,c]

[1.]School of Mechatronics，Northwestern Polytechnical University; Xi'an 710072, China

[2.] School of Material Science and Engineering, Henan University of Science and Technology, LuoYang 471003,China

[3.]Henan Laboratory for Material damage and Tribology, LuoYang 471003,China

[a]yangguangyingyu@163.com, [b]yzzhang@mail.haust.edu.cn, [c]chenyue@mail.haust.edu.cn

Key words: DC magnetic field，magnetic field intensity，super-paramagnetic

Abstract: The influence of DC steady magnetic field on dry-sliding friction and wear characteristics of friction pair of high-speed steel (HSS) ring / Ti_6Al_4V pin is studied by use of improved MPV-1500 type under normal temperature. It is shown that the wear loss of ferromagnetic materials can be decreased by applying a steady magnetic field. Under the load of 250N and the friction velocity of 0.6m/s, the wear loss of the sample HSS ring keeps decreasing along with the enhancement of magnetic field intensity, and in the scope of 64mT-160mT, the average friction coefficient becomes smaller and smaller. Magnetic field can accelerate debris refinement. The wear mechanism covers oxidative wear and adhesive wear mainly.

Introduction

Along with the development of science and technology, especially the extension of electromagnetic technology application, more and more friction pairs of electrical equipments and appliances have been used under the electromagnetic condition, the study on the influential factors and regularity of friction and wear characteristics of materials of ferromagnetic and non-ferromagnetic pairs under a magnetic field condition and the establishment of theoretical basis and guideline for concurrent tribology have become critical topics for tribological design of parallel in mechanical designs[1-4]. We have reported that wear characteristic of the ferromagnetic materials can be decreased if magnetic field intensity can be controlled properly; Therefore, the effect of magnetic field on the wear characteristics of ferromagnetic and non-ferromagnetic pairs also should be investigated further. Then Ti6Al4V alloy with high specific strength, good hot strength, excellent corrosion resistance and other advantages, has become the aviation, space rockets, missiles, ships, chemical, metallurgy, medical care and energy and chemical industrial sectors important structural materials [5,6].

This text, by taking the friction pair of HSS ring / Ti6Al4V alloy pin as the subject investigated, introduces the influence of DC steady magnetic field on friction and wear characteristics of friction pair of HSS ring / Ti6Al4V alloy pin and describes its wear mechanism.

Corresponding author: Tel: +86-379-64231723, E-mail: yzzhang@mail.haust.edu.cn.

Test equipment and method

This test is performed on an improved MPV-1500 type friction tester, where the pin-ring contact mode is adopted. During the test, the friction test will be operated for 20min under the relative friction velocity of 0.6m/s and the load of 250N. A pressure sensor is used for calibration of actual load before each test for fear of extra pressure for the magnetic field force.

The hardness of the ¢7mm×38mm sample ring, being made of W18Cr4V HSS, is HRC60. The hardness of the ¢ 35mm×20mm sample pin, being made of Ti6Al4V alloy, is HRC32. As for the chemical compositions of the sample pin-ring, see the following Table 1.

Table 1 The chemical compositions of high-speed steel and Ti_6Al_4V alloy[%]

Pin/Ring	C	Si	Mn	Cr	Al	V	W	H	Zr	Ti	Fe
high-speed steel	0.79	0.21	0.15	4.20	-	1.16	18.65	-	-	-	others
Ti6Al4V alloy	0.30	0.27	0.64	0.25	5.5-7.0	4.2-6.0	-	0.30	0.20	others	0.10

The magnetic field can be obtained by using an energized coil and the magnetic field intensity can be controlled by adjusting the voltage or current. The magnetic flux flowing through the sample pin section will be measured by use of HT701 flux-meter, and through formula (1) calculate the cross section of the magnetic field intensity.

$$B = \frac{\phi}{nS} = \frac{4\phi}{n\pi D^2} \quad (1)$$

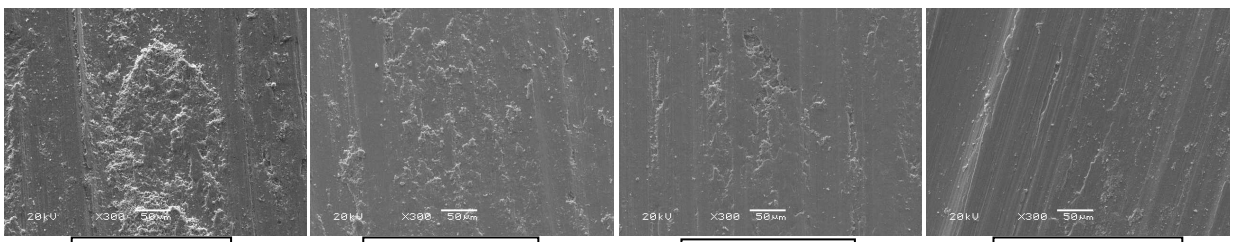
Where, B is the magnetic field intensity; Φ is the flux; S is the cross-sectional area; D is the sample pin section diameter;

The sample pin and ring shall be subject to demagnetization (TC-2 demagnetizer) and cleaning respectively after each test and weighed by a BS210S electronic balance with an accuracy of 0.1mg for calculation of wear loss before and after tests. The friction surface appearance of the worn pin and the abrasive dust compositions shall be analyzed with a spectrum analyzer JSM5600LV scanning electron microscopy and X-ray diffraction D/max-3C automatically.

Result and analysis of test

Influence of magnetic field intensity on wear loss

Fig. 1 shows the friction surface appearances (x300) of the Ti_6Al_4V pin under the load of 250N, the friction velocity of 0.6m/s (20min) and different magnetic field intensities. According to Fig. 1, when no magnetic field or magnetic field intensity is low, much debris adhere to the friction surface adhesion. Along with the enhancement of the magnetic field intensity, the adhesion debris on friction surface becomes less and less. When magnetic field intensity is 160mT, the friction surface keeps fairly clean, micro-cutting wear mainly. Possible reason is that, debris will be caused under a fluctuant positive pressure and friction force of the friction surfaces during the test of friction and wear; As the Ti6Al4V ring / HSS pin friction pair of contact surface parallel to the ground plane,

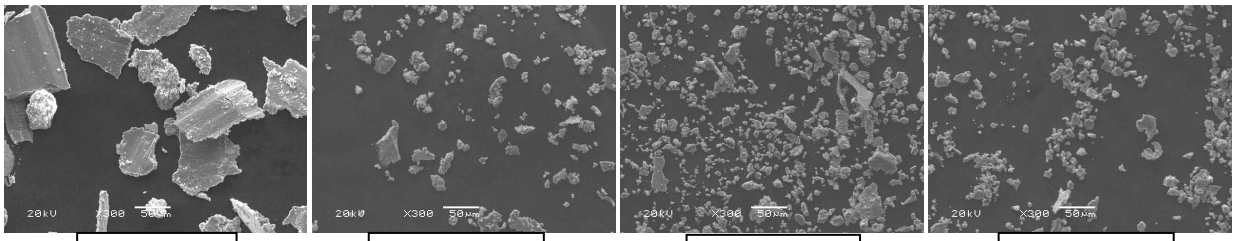


（a）B=0mT （b）B=64mT （c）B=96mT （d）B=160mT

Fig.1 Under different magnetic field intensity, Ti_6Al_4V pin surface morphology(×300).

HSS ring at the above, Ti6Al4V pin at the below. When no magnetic field or magnetic field is small, due to the role of gravity and the molecular force, debris may easily adhere to the Ti6Al4V pin friction surface. Also, because HSS is a ferromagnetic material, while the Ti6Al4V alloy non-ferromagnetic material, when the magnetic field is large, the produced debris are adsorbed easily and remaining in the high-speed steel ring contact surface, and stopped on the direct contact surfaces between the HSS and Ti6Al4V matrix without shed-off. So when magnetic field intensity higher, friction surface adhesion less debris, the more friction surface cleaner.

Fig. 2 shows the debris morphology (x300) with different magnetic field intensity. According to Fig. 2, when no magnetic field, sharp flakes and larger debris exit; along with the enhancement of the magnetic field intensity, the debris becomes finer and rounder. The reason for it is that, as the magnetic field intensity increases, the debris more easily absorbed on friction surface. The greater magnetic field, the bigger adsorption force, the longer participation time of debris in test, smaller debris particles, the more spherical the particle diameter less than 30nm for more debris, or even much less than 20nm. The reference [7-8] reported that, when the particle diameter less than 30nm (Fig.2), the non-magnetic alloy particles will also occur super-paramagnetic phenomenon in the magnetic field, showing a strong magnetic, resulting in significantly increased adhesion strength.

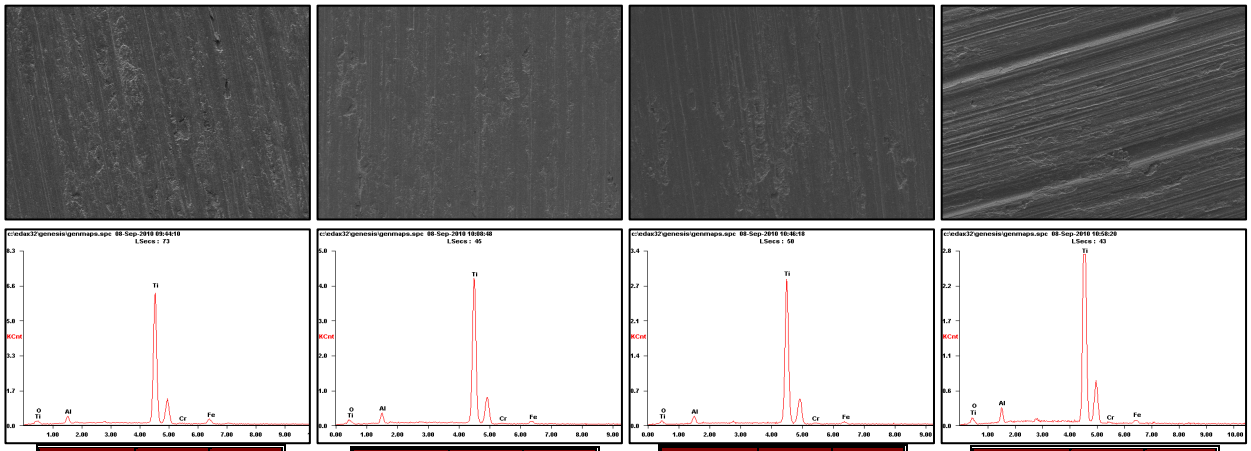


（a）B=0mT　（b）B=32mT　（c）B=64mT　（d）B=96mT

Fig.2 Under different magnetic field intensity, debris morphology (× 300).

Fig.3 shown overall friction surface of Ti_6Al_4V pin EDS spectrum analysis (× 100) under different magnetic field intensity. According to Fig.3, friction surface oxidation did take place. As the magnetic field strength increases, the pin friction surfaces remains less Fe element content, O is also decreased, but the content of Ti element is more and more. Possible reason is that, the larger the magnetic field, debris produced more easily attracted to the HSS friction surface, and less debris stuck in the Ti6Al4V friction surface, so Ti6Al4V substrate surface more easily exposed. Therefore, the greater the magnetic field, and marketing elements of the friction surface of the Fe content of less, Ti element content increases. This is also confirmed under the magnetic field environment, the ferromagnetic material can really absorb small non-iron magnetic alloy particles.

Fig.4 shows the XRD analysis of debris under different magnetic field intensity. According to Fig. 4, when no magnetic field or magnetic field is low, the main constituents of the debris are β-Ti and Ti_3O and a small amount of Fe_3O_4 and Fe_2O_3. As the magnetic field increases, β-Ti content was reduced and then increased. Fig.5 show the influence of magnetic field intensity on friction pairs of HSS / Ti_6Al_4V alloy wear loss. According to Fig.5, along with the enhancement of the magnetic field intensity, the wear loss of the sample ring decreases, while the sample pin increases gradually. Applied magnetic field reduces the wear of the ferromagnetic material HSS. The reason is that, because of higher hardness of HSS, while the hardness of Ti6Al4V alloy smaller, so Ti6Al4V worn more serious during the test process, the initial debris mainly from Ti6Al4V. When no magnetic field, due to the effects of gravity and molecular force, a little of debris remain on the friction surface, oxidation partly, most of β-Ti was too late to oxidation and break away the friction surface; When the magnetic field is large, a lot of debris which protect the HSS friction surface, was adsorbed on the HSS ring friction surface. Ti6Al4V alloy substrates frequently involved in wear. When the wear



Element	Wt%	At%
O K	05.16	17.53
TiK	81.03	67.84
FeK	03.81	04.26

（a）B=0mT

Element	Wt%	At%
O K	03.83	10.16
TiK	85.48	75.67
FeK	02.96	02.25

（b）B=96mT

Element	Wt%	At%
O K	04.32	11.45
TiK	86.15	76.32
FeK	02.77	02.10

（b）B=128mT

Element	Wt%	At%
O K	01.92	05.29
TiK	88.46	81.59
FeK	02.76	02.18

（b）B=192mT

Fig.3 Overall friction surface of Ti_6Al_4V pin EDS spectrum analysis (× 100).

debris adsorbed by HSS ring friction surface reaches a certain level, the excess debris break away the friction surface for too late to complete oxidation. Therefore, along with the enhancement of the magnetic field intensity, the wear loss of the sample ring decreases, while that of sample pin increases gradually, the friction surface becomes cleaner(Fig.1). When no magnetic field or magnetic field intensity is very high, the β-Ti phase of both of its wear debris was higher (Fig.4).

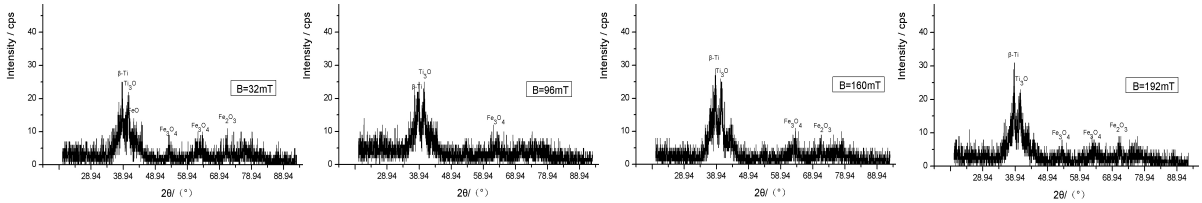


Fig.4 Under different magnetic field intensity, XRD analysis of the debris.

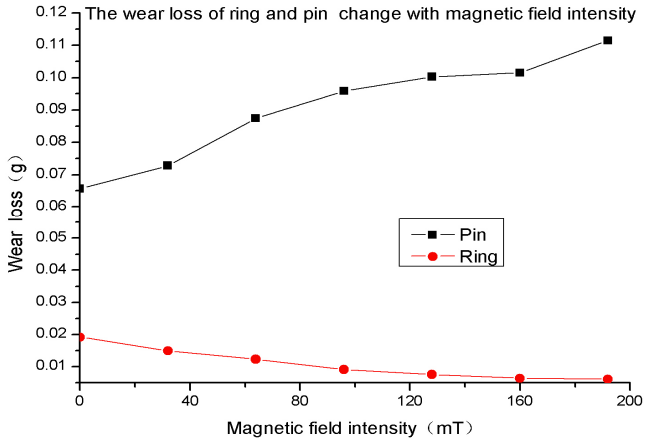


Fig.5 Influence of Magnetic field intensity on friction pairs of HSS / Ti_6Al_4V alloy wear loss.

Influence of magnetic field intensity on friction coefficients

Fig. 6 shows the diversification of average coefficients with the magnetic field intensity. As shown in Fig. 6, along with the enhancement of the magnetic field intensity, the trend of the average coefficients decreases and then increases. The reason for it is that, as the analysis shows in Fig.1-4, during the experiment friction surface will produce oxide, while the oxide layer could play a

lubricant. When no magnetic field, the average coefficient is high due to major debris adsorbed on the Ti_6Al_4V pin friction surface; When magnetic field intensity is high at 192mT, the average coefficient is high due to major debris which were adsorbed on the HSS ring friction surface; while the magnetic field intensity in 32mT-160mT range, the friction surface of both HSS ring and Ti6Al4V pin have thick oxide layer, so the friction coefficient smaller.

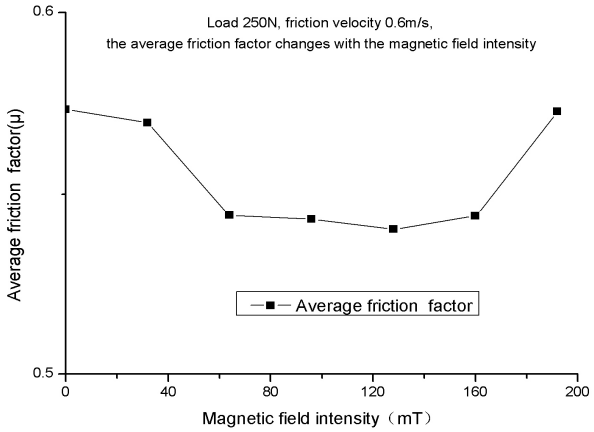


Fig.6 Average coefficient of friction changes with Magnetic field intensity increasing.

Conclusion

a. The wear of ferromagnetic materials can be decreased in the steady magnetic field; The wear loss of the sample HSS ring keeps decreasing comparing with that of the sample Ti_6Al_4V pin which increases along with the enhancement of magnetic field intensity under the load of 250N and the friction velocity of 0.6m/s;

b. Along with enhancement of magnetic field intensity within the scope of 0mT-192mT, the trend of the average coefficients decreases and then increases;

c. Magnetic field can accelerate debris refinement;

d. The wear mechanism covers oxidative wear and adhesive wear mainly.

Acknowledgment

This research is supported by the National 973 Project (2007CB607603) grants in China.

References

[1] Y.H. WEI, Y.Z. Zhang, Y.Chen: Lubrication Engineering. Vol.35-5(2010), p.94-98
[2] Y.H. WEI, Y.Z. Zhang, Y.Chen: Tribology. Vol.30-4 (2010), p.399-403
[3] M. Amirat，H. Zaidi，J. Frene: Engineering Tribology. 2008.222:279-290.
[4] A. Senouci, H. Zaidi, J. Frene: Applied Surface Science. Vol.144-C145(1999), p.287-291
[5] G. Straffelini, A. Molinari: Wear. Vol. 236(1999), p.328-338
[6] A.M.Ohidul, A.S.M.A. Haseeb: Tribology International. Vol. 35(2002),p.357-362
[7] A. Molinari, G. Straffelini, B. Tesi: Wear. Vol. 208(1997), p.105-112
[8]A. Zhe, L. ZHU, F. LIN: Journal of Harbin Medical University. Vol. 38-5(2004), p.424-425
[9] G..D. ZHANG, Y.P. GUAN, G.B. SHAN: The Chinese Journal of Process Engineering. Vol. 2-4(2002), p.319-324

doi:10.4028/www.scientific.net/AMM.50-51.353

A Maturity-based Adaptive Ant Colony Optimization Algorithm

Haining Wang[1, a], Shouqian Sun[1, b], Bo Liu[2, c]

[1] Institute of Modern Industrial Design, Zhejiang University, Hangzhou 310027

[2]Insisitute of Information Science and Engineering, Central South University, Changsha 410083

[a]wanghn@zju.edu.cn, [b]ssq@zju.edu.cn, [c]cocoaliu07@gmail.com

Keywords: Ant colony optimization, Average path similarity, Adaptive parameter control

Abstract. In this paper, for the problems of low convergence rate and getting trapped in local optima easily, the average path similarity (APS) was proposed to present the optimization maturity by analyzing the relationship between parameters of local pheromone updating and global pheromone updating, as well as the optimizing capacity and convergence rate. Furthermore, the coefficients of pheromone updating adaptively were adjusted to improve the convergence rate and prevent the algorithm from getting stuck in local optima. The adaptive ACS has been applied to optimize several benchmark TSP instances. The solution quality and convergence rate of the algorithm were compared comprehensively with conventional ACS to verify the validity and the effectiveness.

Introduction

Ant colony optimization (ACO) is a novel intelligent optimization algorithm proposed by M. Dorigo from observing the ant foraging behavior [1-2]. This algorithm has received good results in traveling salesman problem (TSP), the well-known NP problem, and some of complicated optimal combination problems. It has drawn more and more attention by researchers and has been modified by a series of strategies to improve the algorithm performance. Some famous modifications include rank based ant system (RAS) [3] proposed by Bullnheimer et al, Max-Min ant system (MMAS) [4] proposed by Stutzle etal, Ant-Q algorithm presented by Dorigo and Gambardella [5], as well as ant colony system (ACS) , the improved version of Ant-Q. Among them, ACS is the best ACO algorithm since it has two rules of global updating and local updating for pheromone [6].

ACS algorithm has problems of low convergence rate and getting trapped in local optima easily as other intelligent optimization algorithms when addressing large scale issues [7]. The contradiction between convergence rate and search stagnation can be alleviated by appropriately setting a series of control parameters in the ACS algorithm. However, there is no strict theoretical basis for parameter selection and the universality is questionable since the optimal value taken by experience is usually associated to specific problems [8]. Therefore, the design of one strategy which can adaptively adjust parameters would be of great significance for improving the convergence rate, solution precision and the universality of the algorithm. This paper analyzed the relation between parameters of pheromone updating, as well as the optimizing capacity and convergence rate, then proposed ATS to present the optimization maturity based on which we adjusted parameters of pheromone updating and proposed an improved adaptive ACS algorithm. We demonstrated that this new algorithm has faster convergence rate and better global optimizing capacity by comparing with classic ACS on several TSP instances.

ACS Algorithm Parameter Analysis

The goal of pheromone updating is to enhance superior paths and weaken worse paths aiming at lead the search algorithm to converge to superior paths rapidly. However, after server numbers of iterations, pheromones would concentrated distribution in few superior paths and reduce the probability of other selection, on which condition the algorithm would gets stuck in local optima

easily. Utilize the pheromone updating rule appropriately would be the key to achieve fast convergence rate and let the algorithm leave local optima.

For the ACS algorithm, the bigger the control perimeter α of global pheromone updating is, the faster pheromones accumulate on superior paths. The convergence rate will be accelerated but with the bigger probability to be trapped in local optima. When the control parameter ρ is relative big, the search randomness would be increased but with low guidance quality and slow convergence rate.

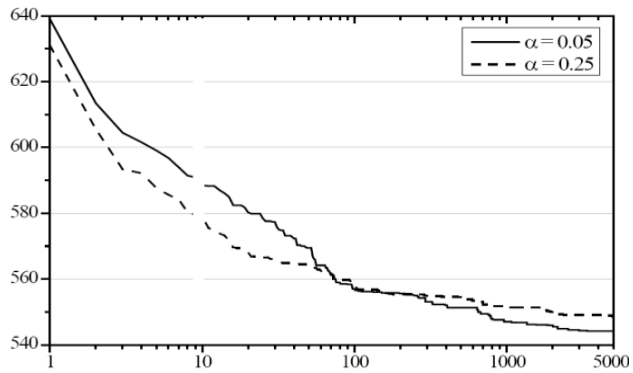


Fig. 1. Optimal path length in each iteration

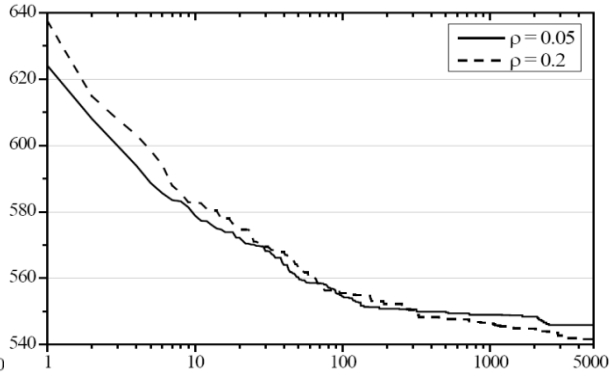


Fig. 2. Optimal path length in each iteration

Fig.1 shows the change of optimal path length on TSP problems when α and ρ are assigned with bigger or smaller values. The lateral axis represents the iteration times as well as the longitudinal axis means the optimal path length in each iteration. We can see that in the first 100 iterations, the convergence rate when α is assigned with bigger value 0.25 is faster than the situation when α is 0.05, which is a smaller value. However, with the increase of iteration times, smaller α would lead to better global search ability, which manifests as stronger convergence ability in late stage and superior solutions.

Fig.2 shows the slow volatilization rate of pheromones is beneficial to the accumulation of pheromones on shorter paths and also the fast convergence rate at early stage when ρ is relative small. However, ρ should be assigned with bigger value to guarantee good global search ability at the mature stage since fast rate is needed to volatilize pheromones on shorter paths at this time and diversity of paths is required to prevent the search process from getting stuck in local optima.

At the early stage of ACS algorithm, to improve the convergence rate of the algorithm, more pheromones must be released on shorter paths and bigger α and smaller ρ should be assigned at this period. When the optimization process enters the mature stage, plenty of pheromones concentrated in several shorter paths where the volatilization should be reduced to avoid trapping in local optima. Therefore, smaller α and bigger ρ should be assigned.

Maturity Strategy of Optimization

The algorithm needs to judge whether current optimization state is mature to achieve the adaptively adjustment of α and ρ. Eq. 1 and Eq. 2 represent the adjustment process of these two parameters.

$$\alpha = F_{\alpha}(\varphi) \tag{1}$$

$$\rho = F_{\rho}(\varphi) \tag{2}$$

φ is the indicator of the optimization process.

Judgment of Optimization Maturity based on Path Similarity. At the early stage of the algorithm, the difference of paths constructed by ants is relatively big. With the updating of global pheromones, the optimization process enters the mature stage when distribution of pheromones concentrates on shorter paths where ants move toward. The similarity of paths constructed by ants in this period would be very high; hence we can judge the optimization maturity of the algorithm according to the similarity of all paths and the shortest path. The computing steps of the method which we call it average path similarity (APS) is as follows.

Step1. Every ant k computes the similarity $s(T_k, T_{best})$ between the path T_k constructed by itself and current optimal path T_{best} respectively. The meaning of Eq. 3 is the number of common edges between T_k and T_{best}.

$$s(T_k, T_{best}) = \left| T_k \bigcap T_{best} \right| \tag{3}$$

Step2. Compute the APS according to Eq. 4. (m and n are the number of ants and cities respectively)

$$APS = \frac{1}{m * n} \sum_{k=1}^{m} s(T_k, T_{best}) \tag{4}$$

Realization of Adaptive Adjustment. We use APS to quantify the optimization maturity of the algorithm and make adjustment to values of α and ρ. Specific implementation is listed as follows.

$$\alpha = \begin{cases} \alpha_{\min}, & if\ a_\alpha * APS + b_\alpha < \alpha_{\min} \\ a_\alpha * APS + b_\alpha, & if\ \alpha_{\min} \le a_\alpha * APS + b_\alpha \le \alpha_{\max} \\ \alpha_{\max}, & if\ a_\alpha * APS + b_\alpha \ge \alpha_{\max} \end{cases} \tag{5}$$

$$\rho = \begin{cases} \rho_{\min}, & if\ a_\rho * APS + b_\rho < \rho_{\min} \\ a_\rho * APS + b_\rho, & if\ \rho_{\min} \le a_\rho * APS + b_\rho \le \rho_{\max} \\ \rho_{\max}, & if\ a_\rho * APS + b_\rho \ge \rho_{\max} \end{cases} \tag{6}$$

At the early stage of the optimization, small APS value requires bigger α and smaller ρ to accelerate the convergence; with the APS grow bigger, the optimization enters the mature stage and require smaller α and bigger ρ to weaken the attraction of shorter paths. Therefore, the value of a_α should be negative and value of a_ρ be positive. These two parameters also determine the slope value of the linear transformation of α and ρ.

In sum, we proposed the improved adaptive ACS algorithm based on optimization maturity, the process is shown in Fig.3.

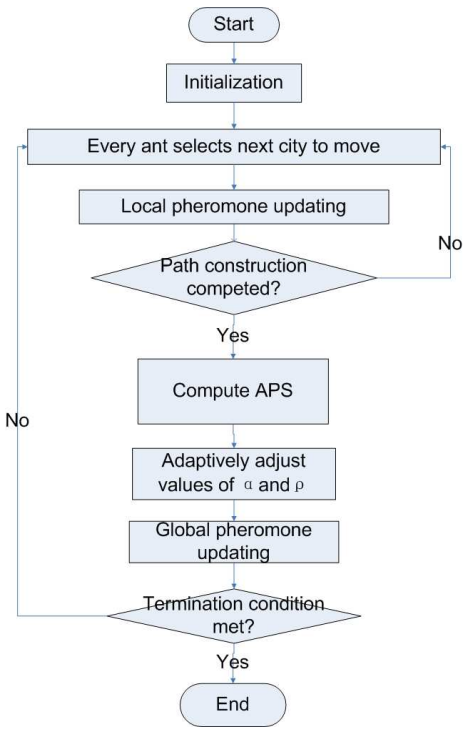


Fig. 3. Algorithm procedure of AACS

Comparison and Analysis of simulation experiments

To verify the universality and effectiveness of the AACS algorithm, we use the traditional ACS and four benchmark examples in TSPLIB to make contrast test. The algorithm runs 100 times on every test instance and every time ends after 5000 iterations. Parameters are taken as shown in Table 1.

Table 1. Parameters of benchmark TSP instances.

TSP instances	TSP City numbers	Shortest path
oliver30	30	421
eil51	51	426
eil76	76	538
eil101	101	629

We can draw the conclusion from Table 2 that the solutions of AACS are better than ACS. For the instance of oliver30, which has the smallest scale, AACS finds the global optimal solution after 100 operations, so does ACS. However, with the increase of the problem scale, the advantage of AACS is more and more obvious. AACS finds global optimal solutions in eil51 and eil76 far more than ACS, and the standard variance is smaller with high stability. It is noteworthy that in the instance of eil101, ACS cannot obtain the optimal solution known while AACS still can keep satisfactory performance, which is benefit from the adaptively adjustment to parameters during the mature stage of the algorithm to make search process leave local optima timely and keep the ability to exploit new paths.

Table 2. Quality comparison of solution paths computed by ACS and AACS.

TSP instances	Algorithms	Average value	Worst Value	Optimal Value	Variance	Optimal Value Times
oliver30	ACS	421.17	421	438	1.70	999
	AACS	421	421	421	0	100
eil51	ACS	428.21	426	435	2.05	18
	AACS	426.51	426	430	0.89	73
eil76	ACS	541.27	538	551	2.77	15
	AACS	539.10	538	542	1.27	43
eil101	ACS	640.67	630	655	5.86	0
	AACS	635.89	629	647	4.75	8

Conclusions

This paper established the adaptive parameter adjustment strategy by analyzing volatile coefficients of pheromones and its influence on algorithm performance during various optimization stages, utilized the average path similarity to represent optimization maturity and proposed an adaptive ACS algorithm to address the contradiction between convergence rate and the stagnation behavior in traditional ACS, as well as improving the algorithm universality. The AACS algorithm has better performance in solution quality and precision, convergence rate and stability compared with traditional ACS algorithm. However, the parameter selection research and the adaptive algorithm in this paper are all based on TSP applications, which are representative optimal combination problems.

In further research, we would extend the adaptive algorithm and parameter analysis to more optimal combination applications.

Acknowledgments

This research was supported jointly by Science and Technology Project of Zhejiang Province (Grant No. 2008C11111), the Universities and Research Institutes of Ministry of Education and Guangdong Province (Grant No. 2009B090300246 and 2009B090300041).

References

[1] A.Colorini, M. Dorigo, V. Maniezzo. Distributed optimization by ant colonies. Proceedings of the First European Conference on Artificial Life, (1991) December 11-13; Paris, France

[2] M. Dorigo, V. Maniezzo, A. Colorni. IEEE Transaction on Systems, Man, and Cybernetics – Part B. 1, 26 (1996).

[3] B. Bullnheimer, R. F. Harl, C. Strauss. Central European Journal of operations research and economic. 7, 1 (1999).

[4] T. Stützle, H. Hoos. MAX-MIN Ant System and Local Search for the Traveling Salesman Problem. Proceedings of the IEEE International Conference on Evolutionary Computation, (1997) April 13-16; Indianapolis, USA.

[5] L. M. Gambardella, M. Dorigo. Ant-Q: A Reinforcement Learning Approach to the Traveling Salesman Problem. Proceedings of ML-95, Twelfth International Conference on Machine Learning, (1995) July 9-12; CA, USA.

[6] M. Dorigo, L. M. Gambardella. IEEE Transaction on Evolutionary Computation, 1, 1 (1997).

[7] L. Chen, J. SHEN, L.QIN, H.J. CHEN. Journal of Software, 14, 8 (2003).

[8] L.Y.JIANG, J.ZHANG, S.H. ZHONG. Computer Engineering and Applications, 43, 20, (2007).

doi:10.4028/www.scientific.net/AMM.50-51.358

An Effective Grid Resource Management Model Based on Cloud Computing Platform

Yue ZHANG[1]，Yunxia PEI[2]

[1]Center of network, Henan college of police, Zhengzhou, China
[2]Department of Math and Computer Science, Zhengzhou University of light Industry, Zhengzhou, China
Email: peiyunxia2010@163.com

Keywords: Grid, resource management, cloud computing, Elastic Compute Cloud

Abstract: Resource management model was presented based on cloud computing using virtualization technologies. According to this strategy, a resource scheduling model was designed. This strategy can really effectively improve resource utilization rate. This algorithm met the needs of cloud computing more than others for grid environment with shorter response time and better performance, which were proved by the simulation results in the Gridsim environment.

Introduction

Cloud computing are a distributed processing, parallel processing and grid computing development and extension of the concept of computer science is the realization of a business [1]. Cloud computing is a new method of shared infrastructure, the basic principle is through a large network storage and computing programs distributed to a large number of distributed computers, and provide application services, allowing companies to be able to switch resources the need for the application, on demand access to the computer and storage systems[2]. Cloud computing is Internet-based super-computing model, through the framework of a distributed, global access to the resource structure to make Internet data center environment to run a similar calculation, that is stored in personal computers, mobile phones and other devices on a large number of Information and processor resources together, work together.

Cloud computing is a Producer-Consumer model, cloud computing system uses fast Ethernet network connection with a number of clusters, the user computing system through the Internet to provide access to a variety of cloud data processing services. Grid system is a resource sharing model, resource providers are also can be resources for consumers, and the study focused on how the grid will be a dynamic combination of resources distributed virtual organizations [3].

Cloud computing and grid computing [4] an important difference is that resource scheduling model. Cloud computing uses clusters to store and manage data resources, the task of running the data center. The scheduling computing tasks to the data storage node running. The grid, places the center calculated. Computing resources and storage resources distributed in every corner of the Internet do not emphasize the task of computing and storage resources required for the same at one place. Due to network bandwidth limitations, the grid computing in the data transmission time a large part of the total running time.

Grid is the data and computing resource virtualization and cloud computing is further virtualization of hardware resources, utilize virtual machine technology, the failure to re-execute the task, and without restarting the task [5]. At the same time, a uniform grid nodes within the operating system (mostly UNIX), and the cloud to relax the conditions for a variety of operating systems in virtual Provide various services to be machine. And the complexity of managing the grid in

different ways provides a simple-to-use cloud computing management environment. In addition, grid and cloud on the payment method has a significant difference. Grid tariff in accordance with the standards set by a number of charges or the sharing of idle resources between the organizations [6]. When the cloud is paid by using the model and service level agreement charges.

The Development of Cloud computing

Currently, Sun, IBM, Microsoft, Google, Amazon and other information giants have been involved in cloud computing research and development [7]. End of 2006, the pioneer of Internet companies Google launched "Google 10l plans," and formally put forward the "cloud" of the concepts and theories, will cloud the thinking of this new shared infrastructure approach, reflected needs and applications. Sun, launched in 2006, the cloud-based theory of the "black box" program, available now in phase; named Big Blue IBM "Blue Cloud" plan; Microsoft hundreds of millions of Windows users around the world by Windows Live provide cloud computing services to achieve the transfer to the general device memory can store any time pattern; the Internet to provide the largest online retailer Amazon Elastic Compute Cloud, an independent developer and the developer provide cloud computing services platform.

Amazon has its own transformation and optimization platform developed on the basis of being known as the Elastic Compute Cloud (EC2) cloud computing platform, the use of surplus enterprise IT infrastructure resources to provide remote cloud outside services. Based on the powerful virtualization technology and network security protocols, Amazon provides a simple way to provide users with flexible billing computing resources rental service. As a leader in the field of cloud computing, Amazon EC2 as the core of the company to build AWS (Amazon Web Services) to create the now widely applied IaaS (Infrastructure as a Service, Infrastructure as a Service) cloud computing model. Amazon cloud computing program called AWS (Amazon Web Service). A WS is a set of services that allow applications to access Amazon computing infrastructure. These services include storage, computing, messaging, and data sets.

Google is a pioneer in cloud computing, cloud computing is the core of its technology. Cheap, efficient cloud computing platform, Google also proud of a great invention, its glory in no way inferior to Web ranking. Google succeeded in making many of the ordinary low-cost PC connected into a "cloud" to provide reliable and efficient computing services, powerful search service is only one of them. Google will be entirely in the framework of its own cloud computing platform, they use the production server all by himself. As the largest manufacturer of servers, Google cans other companies 1/5 or even 1/20 the cost of manufacturing the server, it is cost-effective server based on these, Google was able to maintain so many applications, and innovation. Google recently introduced a system based on Linux open-source Chrome OS, integrated Google's services, aimed at creating a "network-centric," the new computer operating experience, to become the world's largest cloud computing services provider. Google's cloud computing infrastructure consists of 4 independent and closely linked to the system components, this four cluster systems were built on top of the file system (Google file system), the characteristics of the application for Google's Map/Reduce distributed computing system, distributed lock service system (Chubby) and large-scale distributed database system (Big Table). Google's cloud can be seen as using virtualization cloud computing infrastructure (hardware architecture), plus the cloud-based file systems and databases and the corresponding development and application environment, users can use the browser to the distribution in the cloud Google DOCS applications.

Hadoop open source cloud computing system is developed by the Apache Software Foundation open source programming tools for parallel computing and distributed file systems, and the Map

Reduce and file system similar to the Google. Is used in a large cluster of low-cost hardware devices run applications on the framework, to provide efficient, high fault tolerance, stable operation interface and distributed storage. Hadoop cloud-based computing environment, providing cloud computing and cloud storage capacity of the online service, end users can use these services through a browser.

Grid Resource Management Model

Cloud computing service model, including "end", "pipe", "cloud" three levels, shown in Figure 1. "End" refers to the user access to the "cloud" of the terminal equipment, which can be desktop computers, notebook computers, mobile phones or any other terminal to complete the information exchange; "control" refers to the network information transmission channels, for public cloud is present, telecom operators to provide communications network for the private cloud refers to the internal communication network or virtual private network; "cloud" refers to the provision of ICT services, resources or information center of the infrastructure, platform and application servers, etc. types of services provided, including infrastructure, platforms and applications. Figure 2 shows the "cloud" contains three levels: the infrastructure level, such as a variety of servers, databases, storage devices, parallel distributed computing system, etc.; platform level, by the operation, support and development of three platforms composition; application level, providing software, data and information, and other applications.

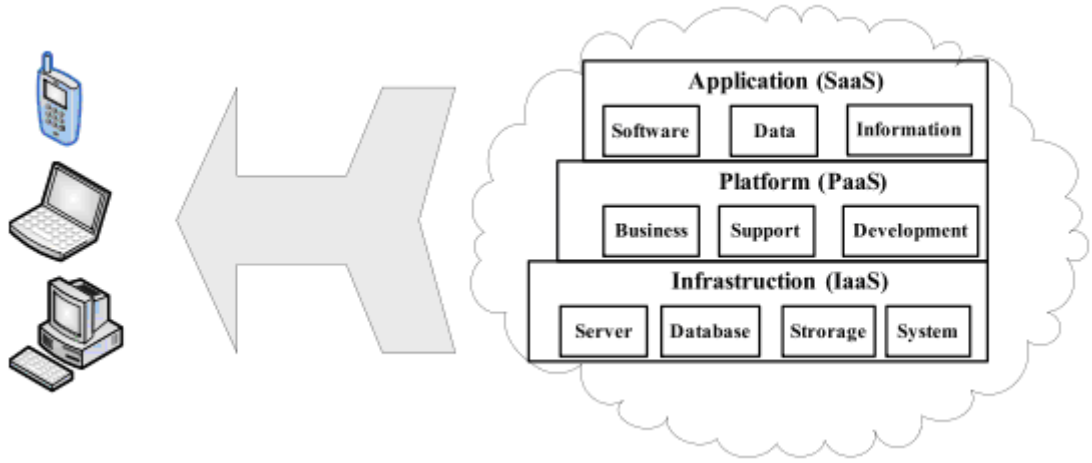


Fig.1. Cloud computing service model

Analyze in the cloud, the cloud architecture can be is divided into five levels: physical layer, the core layer, resource layer architecture, development platform layer, application layer. In accordance with the following hierarchy levels, respectively, the role of the introduction:

1) Physical Layer: refers to the geographic distribution of different parts of the local resources to provide local resources to support. Local resources can be computing resources, storage resources, sensors, servers, networks and other local resources. Physical layer is the bottom of the cloud infrastructure. Providers responsible for physical layer operation, administration, maintenance and upgrade of physical resources to provide HaaS IT services to a huge demand for large enterprises.

2) Core layer: refers to the basic management of distributed resources, distributed applications by abstracting services deployment environment. Core functions can be OS kernel, hyper visor, virtual machine monitor or cluster middleware abstraction services. Providing KaaS to the deployment of distributed applications are.

3) Resource layer structure: the core layer is above the deployment of distributed applications,

distributed resources to provide basic services. This layer provides the basic services of distributed resources, including IaaS, DaaS, CaaS. IaaS is distributed computing services, providing flexible, efficient, high-intensity computing services, IaaS primarily through virtual technology. DaaS is a distributed storage service to provide reliable, safe, large capacity, easy data storage services.

CaaS is a network communications services, providing reliable, secure network communications services.

4) Development platform layer: is the API for application developers by providing a variety of cloud computing programming environment, but also for application developers to provide extension, load balancing, authorization, email, user interface and other various support services. P aa S accelerated the deployment of application services to support the application service extensions.

5) Application layer: is provided through the development of platform development environment and market demand, developed a variety of applications. Application provider is responsible for software development, testing, operation, maintenance, upgrades, provision of a safe, reliable service.

Cloud computing will be the main process on the "cloud" to complete, "end" function can be simplified or become "thin", but this does not mean that the existing types of intelligent terminals will be abandoned. Precisely because the terminal .The improvement of intelligent data processing needs increase, the relative lack of terminal computing resources to promote the emergence of cloud computing and development of various types of "clouds" in turn, promote the emergence of intelligent terminals and integration to enhance user IT resources and information services. Depends on the evolution of cloud computing, "client" - "pipe" - "cloud" of coordinated development.

Grid scheduling is an optimization problem. Weights can be used with every objective to change the behavior of the scheduler on-the-fly, allowing the scheduler to prioritize on multiple objectives [8]. A cloud computing based scheduler can adapt easily to minor changes in the problem space. Hence, a GA based scheduler for grid applications has been chosen for the proposed resource allocation system. The architecture of a service grid is shown in Fig.2.

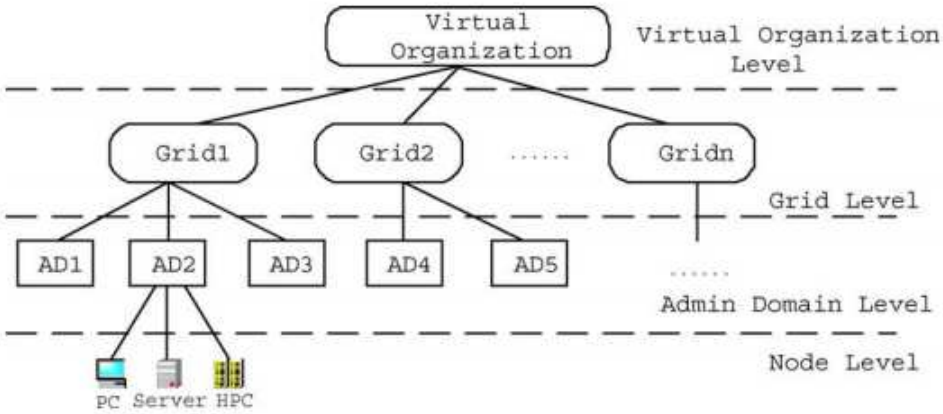


Fig.2.The architecture of a service grid

Conclusions

This work addresses grid resource management issues using a combination of intelligent agents and cloud computing approaches. For local grid resource management, the iterative heuristic algorithm is more efficient than the genetic algorithm. For global grid resource management, a cloud computing service advertisement and discovery technique is shown to be effective. We found several sub-optimal solutions that may be acceptable ones producing a limited waste of the machine

available power. To reach such suboptimal solutions a parallel run of the solver should limit the search time to less than a minute. The average quality of the solutions is quite good, and because of this, application to a real world environment would be interesting. The next generation grid computing environment must be intelligent and autonomous to meet requirements of self management. Related research topics include semantic grids and knowledge grids.

References

[1] Stephen B.Google and the Wisdom of Clouds[Z]. *http：//www．businessweek．com／magazine／content／07_52／b4064048925836*．Htm,2007-12-13.

[2]Guo Jun, Peng, Chen Gaoyun, Huang Jian. MPI-based cloud computing model [J]. *Computer Engineering,* 2009, 35 (24):84-86

[3] CHEN Quan, Deng Qian Ni, cloud computing and its key technology [J], *Computer Applications, 2009, 29 (9):2562 -2567*

[4]BERMAN F，FOX G，HEY T. The grid：Past, present，and future[M]//*Grid computing：Making the global infrastructure a reality．England：John Wiley and Sons*．2003：9-50.

[5] J. Cao, D.J. Kerbyson, G.R. Nudd, Use of agent-based service discovery for resource management in metacomputing environment, in: *Proceedings of the Euro-Par'01, Lecture Notes on Computer Science*, vol. 2150, Springer, Berlin, 2001, pp.882–886.

[6] I. Foster, C. Kesselman, J.M. Nick, S. Tuecke, Grid services for distributed system integration, *IEEE Comput*. 2002, 35 (6) :37–46.

[7] S. Cahon, E.G. Talbi, PARADISEO: A Parallel and distributed library for metaheuristics, in: PAREO'2002, *Third meeting of the PAREO Euro Working group on Parallel Processing in Operations Research, Guadeloupe*, France, May 20–24 2002.

[8] V. Hamscher, U. Schwiegelshohn, A. Streit, R. Yahyapour, Evaluation of Job-Scheduling Strategies for Grid Computing. Grid Computing—GRID 2000, *in First IEEE/ACM International Workshop, Proceedings*, Springer, 2000.

doi:10.4028/www.scientific.net/AMM.50-51.363

Design for automatic cone-shaped dynamic traditional Chinese medicine extraction equipment

Bo Liu[1,a], Haining Wang[2,b]

[1]Insisitute of information science and engineering,Central South University,Changsha, China

[2]Institute of Modern Industrial Design, Zhejiang University, Hangzhou 310027, China

[a]cocoaliu07@gmail.com, [b]wanghn@zju.edu.cn

Keywords: traditional Chinese medicine; extract; automatic; cone-shaped

Abstract. In order to solve the low automatic degree and poor quality of extract production to traditional Chinese medicine, a new automatic cone-shaped dynamic traditional Chinese medicine extraction equipment is designed. This equipment has innovation research in extraction process, structure design aspects, to achieve high efficiency, saving energy and automatic extraction purpose.

Introduction

Traditional Chinese medicine is a treasure of Chinese nation, in the Chinese nation of development for thousands of years, it played a major role for human disease prevention and control[1]. With the development of Chinese medicine formulations diversity, the requirements to the quality control during extraction process have become more sophisticated, how to make a leap of extraction development has been a topic of concern in the field of modern medicine[2-4].

The extraction equipment is the key equipment in extraction process, largely determine the extraction efficiency and quality of medicine[5,6]. Traditional clamp cover heating extraction installment material easy coking, easy sealing, low extraction efficiency, it is difficult to achieve high-quality assembly line. In this case, it's necessary to research and develop an efficient intelligent dynamic extraction equipment[7,8]. To the existing problems, this paper studies a new traditional Chinese medicine extraction equipment, the equipment in extraction technology, it made innovative research in extraction and structural design, to achieve high efficiency, saving energy and automatic extraction purpose.

System Overview

For the current problems of traditional extraction equipment exists materials easy coking, easy sealing, low extraction efficiency, difficult to achieve automate production and so on, combined with the foreign advanced technology of cone-shaped dynamic extraction, to put forward a new extraction equipment which combined the heating cycle extraction process, cone-shaped extraction tank design, automatic control technology. System consists of three components: heating system, extraction system and control system. Specific technical route system flow chart shown in Figure 1.

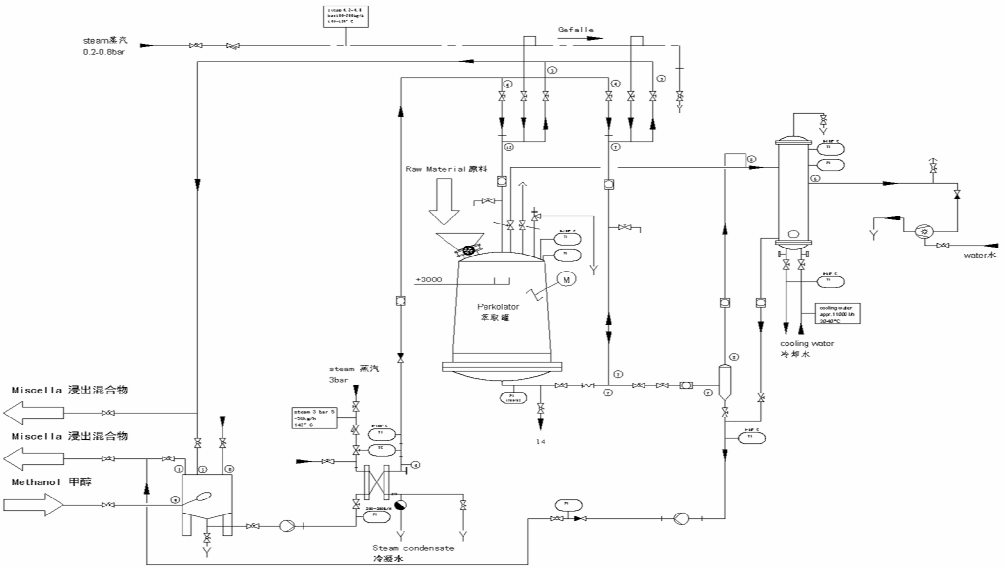


Figure 1 Outer loop extraction system flow chart

The following extract, respectively for the three components of the system described in detail.

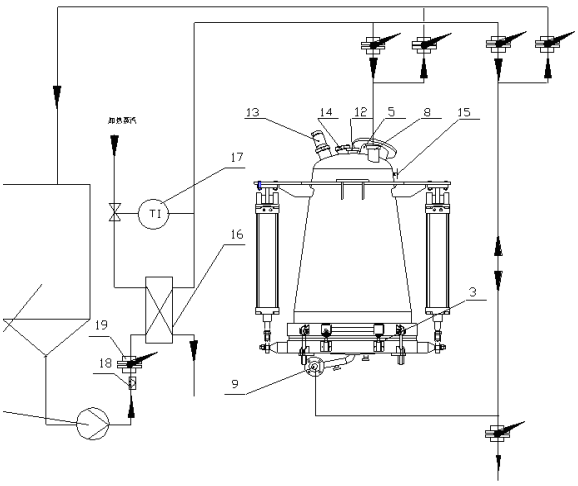
Heating System

The traditional clamp cover heating extraction installment, is directly in withdraws the tank on the tube wall carries on the heating with the aid of the steam or other heating medium. As the solid material in the extraction tank direct contact with the heating tube wall, extremely easy to generate overheating, coking, bonding and so on. Especially regarding heat-sensitive Chinese herbal medicine (such as sea-buckthorn) extraction, this kind of phenomenon is especially serious, has greatly affected the quality which extracts.

In order to avoid material coking phenomenon, in the selection of heating system, this paper use external heating system for heating the extracted supernatant. Cone-shaped dynamic extraction equipment extraction by external temperature automatic control and circulation system, not only rapidly heat extract in extraction tank to desired extraction temperature, but also automatically maintain the extraction temperature, avoid the results that the extracted solid material and the heating surface contact directly which would cause overheating, coking, bonding effects, which are not conducive to extraction and affect the quality of extract. Since the presence of extract buffer tank and circulation pump, it makes the whole process completed in a premise of high ratio of liquid and solid and the condition of extract forced circulation, the extraction equipment has advantages of high extraction rate and thermal efficiency, good extracting quality, short operation time etc.

Extraction System

This is the core of the system equipment, including: inverted cone-shaped tank, sieve plate, the control cylinder, feeder hopper, valve, outer loop heating, thermostat, flow meter, temporary storage tank, circulating pump and other components , the structure shown in Figure 2 and Figure 3.



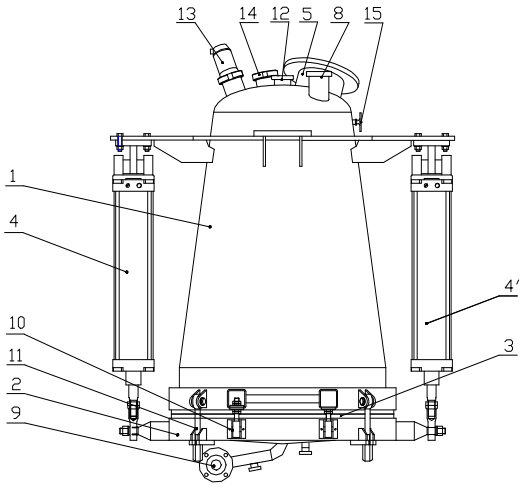


Figure2 Extraction system and the heating system structure Figure3 Extracting tank structure

1 cone-shaped tank 2 end open door 3 sieve plate 4 control cylinder 5 feeder inlet 6 feeder hopper 7 valve 8 top liquid entrance 9 bottom liquid entrance 10 locate pin 11 locking nut 12 steam outlet 13 endoscopic light 14 viewer 15 thermometer 16 outer loop heating 17 thermostat 18 flow meter 19 control valve 20 temporary storage tank 21 circulating pump

There are two innovations:

(1)Materials in extraction tank for forward and reverse extraction

The top and bottom liquid entrance of cone-shaped dynamic extraction tank and outside circulate temperature control and heating circulate system are designed to carry the material in extraction tank on forward (extract enters from tank deck, flows material level through the sieve of bottom door and flow out by tank bottom) and the reversion (extract enters from tank bottom, flows material level through the sieve of bottom door and flow out by tank deck) extraction. The merit of use the reversion extraction is: regarding the material which not easy to seep or fine-grained material, uses the reversion extraction, under the same pump pressure, extract affects bottom-up in the material, because the pump pressure direction and the material gravity strength direction is opposite, cannot exist the situation that when the material being extracted under the action of the pump is tight and lead to more pressure the more flow out of date extract increased resistance. Therefore, with the aid of the circulating pump power may guarantee extracts circular flow.

(2)Extraction tank for the inverted cone-shaped

Inverted cone-shaped is designed to take full advantage of the tank weight of material and facilitate product out. Cone-shaped dynamic extraction tank tapered inverted cone angle is generally (5 ~ 40) °, On the theoretical calculations and numerous experiments by the final selection to (10 ~ 30) °. This choice, not only the product shape is more beautiful at the same time facilitate the material by weight of gravity to the bottom of the tank. In cone-shaped dynamic extraction tank, bottom door diameter larger than tank. Inverted cone-shaped tank design makes the reverse extraction, extract flow from the tank bottom-up process shrinking, thereby increasing its velocity, and does not produce short extract affected the results of extraction efficiency. Inverted cone-shaped tank and large diameter bottom door not only make product out easily, but also when the material dry the problem of extract short-circuit will not happen.

Control system

The equipment is done automatically by a control system for the production process. Main achieved:

(1)Data acquisition

Control system automatic collect technology parameter of liquid level, pressure, density, temperature and so on, automatic collect running performance on the valves, motors, pumps etc, transmit data to main control room, while receiving the instruction from the host machine.

(2)Automatic control

Control system work automatically during the production process, such as extraction temperature, pressure, concentration control. The process also can control the automatic opening and closing the technology valve, order control, interlock control, etc.

(3)Troubleshooting

Trouble appears inevitable. When equipment occurs error, if the production process continued, will inevitably affect the quality of extract. The control system has corresponding interlock shutdown function in design of troubleshooting, which is a link in the chain fails, the corresponding other part will stop. By the technical staff for troubleshooting, find the error and production to be continued.

Conclusions

The device used by the trial, Compared with the traditional clamp cover heating extraction installment, extraction rate increased 20%, 80% efficiency, lower processing costs by 50%. Seen from using automatic cone-shaped dynamic traditional Chinese medicine extraction equipment, will not only improve the efficiency of traditional Chinese medicine production, but also greatly reduce production costs. It's significance is not only equipment performance in line with industry development, improve the technological content of China pharmaceutical machinery and equipment, and to promote modernization industrial development of traditional Chinese medicine, improve employment opportunities and promoting economic growth will produce huge economic and social relevance.

References：

[1] Xianjie Song, Xiaoping Shen, Fahuan Ge, Xiaongyong Zhao. Chinese medicine extract process automation and control technology. Journal of Chinese medicine materials, 2006(9):984-986

[2] Yuan Yingjin, Liu Mingyan, Dong Anjie. Key modern production technology of Chinese medicine. Beijing: Chemical industry press. 2002.

[3] Bai Yu. Automatic control for the extracting process of the production of traditional Chinese medicine. Automation & instrumentation, 2007(3):69-72

[4] Liu Xuhai, Chen Xiaorong, Wei Xiaohua. Extraction and automatic control of modern Chinese

medicine. Chinese journal of information on traditional Chinese medicine, 2004（10）:887

[5] Liu Qingge, Wang Jianzhong, Shi Jianping, Wang Jingyi. Automatic production of traditional Chinese medicine. Chinese Journal of Pharmaceuticals, 2003 34（8）

[6] Liu Genqiang, Liu Cuiping, Chen Hongqiang. Extract automatic control system of traditional Chinese medicine. Qilu Pharmaceutical Affairs, 2006 25(12)

[7] Wang Chenggang, Yang Hongm Wei Huazhong, Zhang Minshu. Multi-function extractor Computer Monitoring System Design and Implementation. Wuhan Institute of Chemical Technology, 2006 28(1):71-74

[8] Zhou Jie, Ye Ping. Process Design of Feeding and Discharge in Chinese Medicine Extraction. Pharmaceutical & engineering design, 2003 24(1):4-6

doi:10.4028/www.scientific.net/AMM.50-51.368

Study of Milling Force in High Speed Milling Carbon Fibre Reinforced Plastics(CFRP) Using Diamond Tool

Dong Liu [1, a], Huaiying Wang[2, b] Honghai Xu [1,c], Chaoying Zhang [1,d]

[1] College of Mechanical Electronical and Engineering, North China University of Technology, Beijing, 100144, China

[2] Beijing Institute of Aerospace Testing Technology,Beijing 100074, China,

[a] liudong@ncut.edu.cn,,[c]xuhh@ncut.edu.cn, [d]zhangchaoying@ncut.edu.cn

Keywords: Milling force, Carbon firbre reinforced plastics, Diamond tool, High speed Milling

Abstract. The Composites are difficult machining materials which widely used in aerospace industry due to their excellent mechanical properties. Tool wear and delamination are considered the major concern in manufacture the parts and assembly. The orthogonal experiment method was used to investigate the milling force of carbon fibred composite with diamond mill tool. The experiments were carried out under air cooling cutting conditions. The experimental results indicated that the feed rate was the most significant fact which affect the average peak value of the milling force. the average milling force decreased with the cutting speed increased and increased with the feed rate and width of cut increased.

Introduction

Composite materials such as fibre-reinforced plastics(CFRP) are extensively used in aerospace, automotive and civil applications due to their excellent mechanical properties. About 30% of the external surface area of the Boeing 767[1] and 50% of the primary structure in Boeing 787[2] program consists of composites. These materials made by using carbon fibres for reinforcing plastic resin matrices, such as epoxy, are characterized by having higher specific strenghth(up to 4500MPa), higher fracture toughness, excellent corrosion resistance, higher specific stiffness coupled with favorable damping properties and zero thermal expansion coefficient. These properties made them especially attractive for aerospace and civil applications.

Milling is the machining operation most frequently used in manufacturing composite materials. The machinability of fiber reinforce plastics is strongly influenced by the type of fiber embedded in the composites and its properties. The carbon fibre reinforced plastics exhibit totally different milling results as compared to those of milling common metals and other materials. Tool wear and delamination at the entrance and exit surface of the workpiece were two major challenges while milling composite materials. Over the years, many researchers have studied the cutting mechanics during machining CFRP composite materials, such as the correlation between tool wear, delamiation damage and surface roughness[2], the influence of process parameters on tool life and workpiece damage[3], the thrust force and delamination during drilling CFPR using core-saw drill[4]. In recent years, manufacture requirements have directed greater emphasis on the development on machining composite materials with new tools and method such as saw drill, core drill and step drill and diamond tools[5]. In this article, the diamond mill tool was used to mill the CFRP composite materials. The thrust force and milling force were measured and the experimental results were analyzed using orthogonal method.

Experimental Setup

Tool and Materials. In this study, CFRP laminates which having thickness of 3.5mm was used as workpiece material. This carbon fibre composite material was produced by autoclave with a fibre orientation of 0/90° , as can be observed in Fig.1.

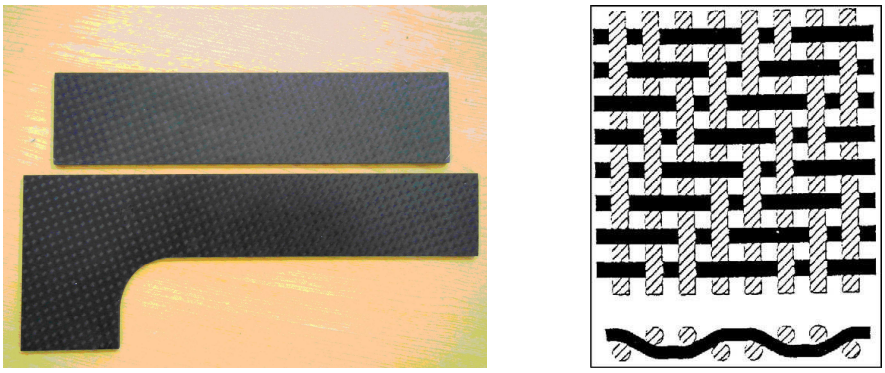


Fig.1. CFRP composite material with a fibre orientation of 0/90°

The carbide drill brazed diamond in its flank and half core drill were utilized in this research. The two flute of carbide drill presents the following geometry: the diameter was ϕ5mm, a helix angle of 30° and flute length of 20mm. The carbide drill and half core drill were presented in Fig.2 and Fig.3.

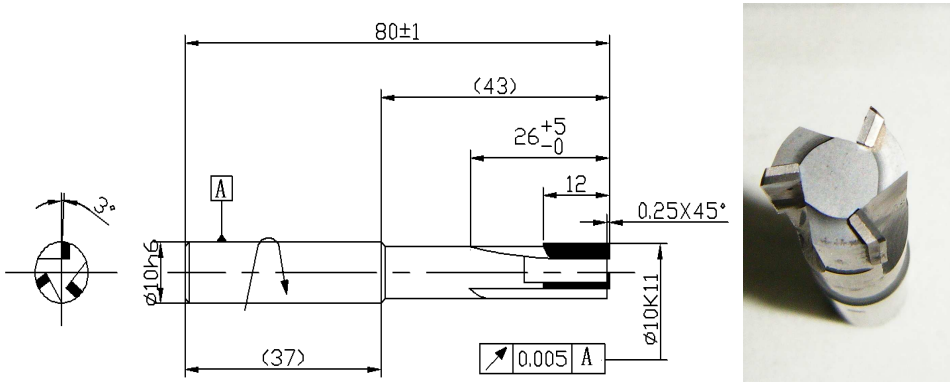


Fig.2 miill tool brazed diamond

All the milling tests were carried out on a 5-axis vertical CNC machining centre DMU 60P hi-dyn with a maximum spindle and power of 24,000rpm and 56kw. The four direction precision dynamometer used in the test was made by the Beijing University of Aeronautics and Astronautics. the electric bridge box, the YD28 dynamic strain gauge, A/D Data acquisition card and computer were selected as the cutting force measurement system.

Plan of experiments

The dynamometer was mounted on the work table of the machining center via rigid clamping accessories to counter vibrations.. The CFRP composite material laminate was clamped on the dynamometer. The experiments were carried out under air cooling cutting conditions. The milling setup for the conducted tests is shown in Fig.4. During the experiments, the mechanical loads signals (milling force) were measured using dynamometer and its complementary charge amplifier. Signal were transmitted using a highly insulated cable further through an A/D card converter to data acquisition and storage software called DASP. Before experiments, the whole setup was calibrated thoroughly.

The spindle speeds used in the experiments were 10000rpm, 14000rpm, 18000rpm. The feed speed were 50mm/min, 80mm/min, 120mm/min and the width cut were 1mm, 3mm, 5mm. The level of the milling parameters of orthogonal experimental plan was shown as table 1.

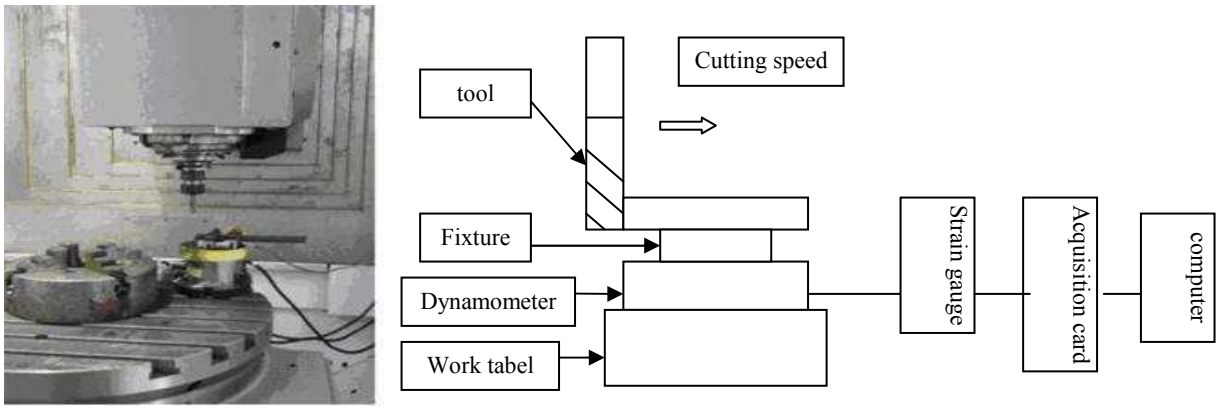


Fig.3 Schematic diagram of the experiment

Table 1 Assignment of the levels to the factors

Level	1	2	3
Revolution (rpm)	10000	14000	18000
Width of cut (mm)	1	3	5
Feed rate (mm/z)	0.01	0.02	0.03

Results and discussion

The values of the milling force from the recorded data were show in table 2.

Table 2 experimental results of milling force

	Cutting speed (m/min)		Feed rate (mm/z)		Width of cut (mm)		Milling force F_{xmax} (N)	Milling force F_{ymax} (N)
1	314.2		0.01		1		362.35	252.46
2	314.2		0.02		3		432.56	312.56
3	314.2		0.03		5		459.13	345.13
4	439.8		0.01		3		368.2	257.20
5	439.8		0.02		5		405.3	279.33
6	439.8		0.03		1		422.2	331.25
7	565.56		0.01		5		364.7	264.75
8	565.56		0.02		1		365.3	282.3
9	565.56		0.03		3		375.4	315.4
	F_{xmax} (N)	F_{ymax} (N)	F_{xmax} (N)	F_{ymax} (N)	F_{xmax} (N)	F_{ymax} (N)		
Average of level 1	418.03	303.38	365.08	258.14	383.28	288.67		
Average of level 2	398.56	289.26	401.05	291.40	392.05	295.05		
Average of level 3	368.46	287.48	418.91	330.59	409.71	296.40		
Extreme difference	49.55	15.90	53.83	72.46	26.42	7.73		

The extreme difference reflected the impact of cutting parameters on the milling force. The cutting parameter was the greater effect on the milling force while the extreme difference of this cutting

parameter was greater. Fig.4～Fig.6 shows the variation of extreme difference of cutting speed, feed rate, width of the cut respectively. It can be seen from the table.2 that the feed rate was the most significant fact which affect the average peak value of the milling force, the extreme difference of F_x was 53.83N and the extreme difference of F_y was 26.42, the following was the cutting speed which extreme difference of F_x was 49.55 and the extreme difference of F_y was 15.90, and the last was the width of cut which extreme difference of F_x was 26.42 and the extreme difference of F_y was 7.73.

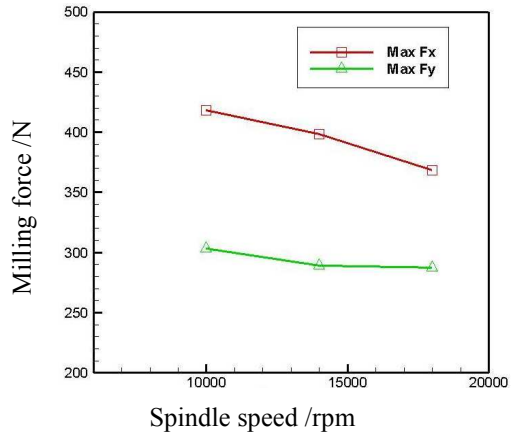


Fig.4 Variation of milling force with cutting speed

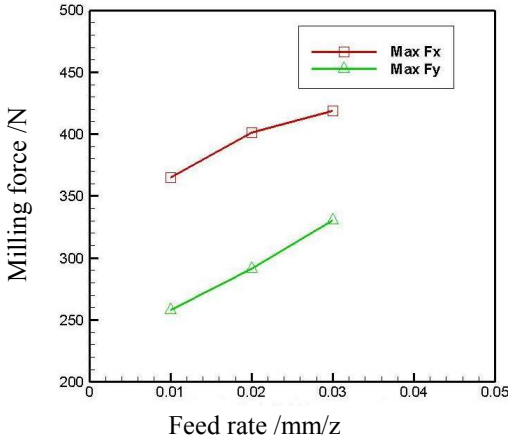


Fig.5 Variation of milling force with feed rate

It can be seen form the Fig.4～Fig.6 that the average milling force decreased with the cutting speed increased and increased with the feed rate and width of cut increased. The width of cut has little effect on the milling force and the feed rate has more notable effect on the milling force.

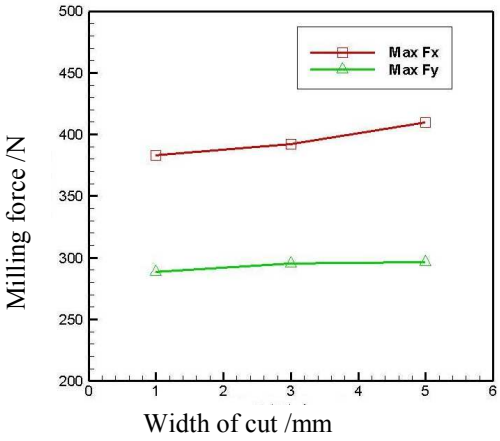


Fig.6 Variation of milling force with width of cut

Summary

The diamond tool was used to mill the CFRP composite materials. The experiments were carried out under air cooling conditions. the regulation of the milling force influenced by the cutting velocity, feed speed and width of cut was obtained. The experimental results indicated that the feed rate was the most significant fact which affect the average peak value of the milling force. the average milling force decreased with the cutting speed increased and increased with the feed rate and width of cut increased.

References

[1] S.R.Ravishankar, C.R.L. Murthy, Proceedings of 14th Conference on Non-Destructive testing, Oxford and IBH Publishing, (1996), p. 8-13.

[2] Sanjay Rawat, Helmi Attia, Wear , Vol.267(2009), p. 1022-1030

[3] I.S.Shyha, D.K.Aapinwall, S.L.Soo, International Journal of Machine Tools & Manufacture, Vol.49(2009), pp. 1008-1013

[4] C.C.Tsao, Int J Adv Manuf Technol, Vol.37(2008), p.23-28

[5] H.Hochen, C.C.Tsao, Journal of Materials Processing Technology, Vol.167(2005), p. 251-261

[6] Liu Dong, Xu Honghai, Zhang Chaoying, . Advanced Materials Research, v 102-104(2010), p. 729-732

©(2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.372

Generalized Refinable Function Vectors With Hermite Interpolating Property

Libo Cheng

School of Science,Changchun University of Science and Technology, Changchun 130022,China

clbyy@126.com

Keywords: generalized refinable function vector; Hermite interpolating property; mask

Abstract. Wavelet analysis has many applications in scientific areas such as computer graphics, image processing, numerical algorithms and signal denoising. In general, a wavelet is derived from a refinable function vector via a multiresolution analysis. In this paper, we presented a novel notion of generalized Hermite interpolating refinable function vector. In terms of its mask, several properties (such as interpolation property, symmetry property and approximation property) with respect to generalized Hermite interpolating refinable function vector. We shall present an example at the end of this paper.

1. Introduction

Interpolating refinable function vector was discussed in [2,3,4], Hermite interpolating refinable function vector was discussed in [1], in this paper, we shall present a generalized Hermite interpolating refinable function vector in section 2,in section 3,we shall respect several properties of the generalized Hermite interpolating refinable function vector in terms of its mask.

Before proceeding further, let us recall some necessary notions and auxiliary results. R denotes the real plane. $\mathbb{Z}$ denotes the set of all integers. C denotes the complex plane. $L_2(R)$ denotes the set of all Lebesgue measurable functions f such that $\|f\|_{L_2(R)}^2 = \int_R |f(x)|^2 dx < \infty$. And $C^{r\times s}$ denotes the linear space of all finitely supported sequences of $r\times s$ matrices on C. And $(l_0(\mathbb{Z}))^{r\times s}$ denotes the linear space of all finitely supported sequences of $r\times s$ matrices on $\mathbb{Z}$.

Definition 1.1 We call $\phi = [\varphi_1,\cdots,\varphi_r]^T : R \mapsto C^{r\times 1}$ is a *refinable function vector* if ϕ satisfies the following *vector refinement equation*

$$\phi(x) = 2\sum_{k\in Z} a(k)\phi(2x-k), \quad a.e. \quad x\in R \tag{1.1}$$

$a : Z \mapsto C^{r\times r}$ is a finitely supported sequence of $r\times r$ matrices on $\mathbb{Z}$, called the *matrix mask* for the refinable function vector ϕ.

In the frequency domain, the vector refinement equation in (1.1) can be rewritten as

$$\hat{\phi}(2\xi) = \hat{a}(\xi)\hat{\phi}(\xi), \qquad \xi\in R \tag{1.2}$$

Where

$$\hat{a}(\xi) = \sum_{k\in Z} a(k)e^{-ik\xi}, \qquad \xi\in R \tag{1.3}$$

And the Fourier transform of a function $f\in L_1(R)$ is defined to be

$$\hat{f}(\xi)=\int_R f(x)e^{-ix\xi}dx, \qquad \xi\in R \tag{1.4}$$

When the multiplicity $r=1$, the function vector ϕ is simply a scalar function and therefore for the case $r=1$, ϕ is called a scalar refinable function.

Definition 1.2 We call ϕ is a *interpolating refinable function* if ϕ is continuously and for $k\in Z$,

$$\phi(k)=\delta_k \tag{1.5}$$

Where δ is a Dirac sequence such that $\delta_0=1$ and $\delta_k=0$ for all $k\neq 0$.

The concept of interpolating refinable function vectors was introduced by [2,3].

Definition 1.3 We call $\phi=[\varphi_1,\varphi_2]^T$ is a *refinable function vector* if ϕ is finitely supported and continuously and for $k\in Z$,

$$\varphi_1\left(\frac{k}{2}\right)=\delta_k, \qquad \varphi_2\left(\frac{k}{2}\right)=\delta_{k-1} \tag{1.6}$$

2. Definition and characterization of generalized refinable function vectors with Hermite interpolating

As an important subdivision scheme, Hermite interpolants are always of interest in the computer graphics, in this paper, we present a generalized refinable function vector with Hermite interpolating.

Definition 2.1 We call a vector $\phi=[\varphi_1,\cdots\varphi_{2r}]^T\in L_2(R)$ is *generalized Hermite interpolating refinable function vector* if ϕ is continuously differentiable and for $k\in Z$, ϕ satisfies the following interpolation condition

$$\begin{pmatrix} \varphi_1(k) & \varphi_1^{(1)}(k) & \cdots & \varphi_1\left(k+\frac{r-1}{r}\right) & \varphi_1^{(1)}\left(k+\frac{r-1}{r}\right) \\ \varphi_2(k) & \varphi_2^{(1)}(k) & \cdots & \varphi_2\left(k+\frac{r-1}{r}\right) & \varphi_2^{(1)}\left(k+\frac{r-1}{r}\right) \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \varphi_{2r}(k) & \varphi_{2r}^{(1)}(k) & \cdots & \varphi_{2r}\left(k+\frac{r-1}{r}\right) & \varphi_{2r}^{(1)}\left(k+\frac{r-1}{r}\right) \end{pmatrix} = \delta_k I_{2r} \tag{2.1}$$

$r=2,3,\cdots$

Where δ is a Dirac sequence such that $\delta_0=1$ and $\delta_k=0$ for all $k\neq 0$, for a smooth function f, $f^{(j)}(\cdot)$ denotes the $j-$th derivative of $f(\cdot)$, I_{2r} is a $2r\times 2r$ unit matrix.

Note 2.1 when $r=2$ and $k=0$, $\phi=[\varphi_1,\cdots,\varphi_4]^T\in L_2(R)$,then (2.1) becomes the following

$$\begin{pmatrix} \varphi_1(0) & \varphi_1^{(1)}(0) & \varphi_1\left(\frac{1}{2}\right) & \varphi_1^{(1)}\left(\frac{1}{2}\right) \\ \varphi_2(0) & \varphi_2^{(1)}(0) & \varphi_2\left(\frac{1}{2}\right) & \varphi_2^{(1)}\left(\frac{1}{2}\right) \\ \varphi_3(0) & \varphi_3^{(1)}(0) & \varphi_3\left(\frac{1}{2}\right) & \varphi_3^{(1)}\left(\frac{1}{2}\right) \\ \varphi_4(0) & \varphi_4^{(1)}(0) & \varphi_4\left(\frac{1}{2}\right) & \varphi_4^{(1)}\left(\frac{1}{2}\right) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$ please see [1].

Considering some important and necessary properties of wavelets analysis, we will discuss the following properties of the Hermite interpolating refinable function vectors.

Theorem 2.1 If ϕ is a generalized Hermite interpolating refinable function vector, that is (2.1) holds, then the shifts of ϕ are linearly independent.

In fact, suppose that exists the sequences $c_1,c_2,\cdots,c_{2r}:Z\mapsto C$ (at least one nonzero) such that

$$\sum_{l=1}^{2r}\sum_{k\in Z}c_l(k)\varphi_l(x-k)=0 \tag{2.2}$$

Since ϕ is interpolating, we have $c_l(k)=0, l=1,2,\cdots,2r$.

Theorem 2.2 For any compactly supported function $f:R\mapsto C$, if denote $\tilde{f}$ as follows

$$\tilde{f}(x)=\sum_{k\in Z}\left[f(k),f^{(1)}(k),f\left(k+\frac{1}{r}\right),f^{(1)}\left(k+\frac{1}{r}\right),\cdots,f\left(k+\frac{r-1}{r}\right),f^{(1)}\left(k+\frac{r-1}{r}\right)\right]\phi(x-k)$$

$$=\sum_{k\in Z}\left[f(k),f^{(1)}(k),f\left(k+\frac{1}{r}\right),f^{(1)}\left(k+\frac{1}{r}\right),\cdots,f\left(k+\frac{r-1}{r}\right),f^{(1)}\left(k+\frac{r-1}{r}\right)\right]\begin{pmatrix}\varphi_1(x-k)\\ \varphi_2(x-k)\\ \vdots\\ \varphi_{2r}(x-k)\end{pmatrix} \tag{2.3}$$

Then as $j=0,1$, $\tilde{f}^{(j)}(x)$ agree with $f^{(j)}(x)$ on $r^{-1}Z$.

3. Properties of the mask

In the section, we shall provide a complete mathematical characterization for Hermite interpolating refinable function vectors with compact support in terms of their masks.

For simplicity of discussion, let $r=2$, we have the following results.

Theorem 3.1 Let $\phi=[\varphi_1,\cdots,\varphi_4]^T\in L_2(R)$ be a compactly supported refinable function vector such that $\hat{\phi}(2\xi)=\hat{a}(\xi)\hat{\phi}(\xi)$, a is the mask of ϕ. If ϕ is Hermite interpolating, that is, ϕ is continuously differentiable and (2.1) holds, then a satisfies

$$\hat{a}(z)=\begin{pmatrix} \frac{1}{2} & 0 & * & * \\ 0 & \frac{1}{4} & * & * \\ \frac{1}{2}z & 0 & * & * \\ 0 & \frac{1}{4}z & * & * \end{pmatrix}. \qquad \text{Here } z=e^{-i\xi}.$$

For a matrix mask a with multiplicity $2r$, we need the following definition in [7,8,9].

Definition 3.1 We say that a satisfies the *sum rules* of order $k+1$ with a dilation factor d if there exists a sequence $y\in\left(l_0(Z)\right)^{1\times r}$ such that $\hat{y}(0)\neq 0$ and

$$\left[\hat{y}(d\cdot)\hat{a}(\cdot)\right]^{(j)}(2\pi m/d)=\delta_m\hat{y}^{(j)}(0) \quad , \quad \forall j=0,\cdots,k, m=0,\cdots,d-1 \tag{3.1}$$

For the rest of the paper, we shall discuss the case $d=2$.

The following result determines the $\hat{y}$ vector in the definition of sum rules in (3.1) for Hermite interpolatory masks.

Theorem 3.2 Let $\phi=[\varphi_1,\cdots,\varphi_4]^T\in L_2(R)$ be a compactly supported refinable function vector such that $\hat{\phi}(2\xi)=\hat{a}(\xi)\hat{\phi}(\xi)$, a is the mask of ϕ. If ϕ is Hermite interpolating, that is, ϕ is continuously differentiable and (2.1) holds, then there exists a nonzero sequence $y\in\left(l_0(Z)\right)^{1\times 4}$ which satisfies

$$\hat{y}(0)=(1,\ 0,\ 1,\ 0),\hat{y}^{(1)}(0)=(0,\ i,\ i/2,\ i) \tag{3.2}$$

where $i^2=-1$, and a satisfies the sum rules of 2.

Corollary 3.1 Let $\phi=[\varphi_1,\cdots,\varphi_4]^T\in L_2(R)$ be a compactly supported refinable function vector such that $\hat{\phi}(2\xi)=\hat{a}(\xi)\hat{\phi}(\xi)$, a is the mask of ϕ.If ϕ is Hermite interpolating, that is, ϕ is continuously differentiable and (2.1) holds, then

$$\int_R\left[\varphi_1(x)+\varphi_3(x)\right]dx=1 \tag{3.3}$$

4. Example

In this section, we shall present an example of Hermite interpolating refinable function vector .

Example 4.1 Let a be supported in $[-1,1]$. Then we may obtain Hermite interpolating refinable function vector(see Fig.1)and mask a (satisfies the sum rule of order4) followed by

$$a(-1)=\begin{pmatrix} 0 & 0 & 0.25 & 0.75 \\ 0 & 0 & -0.03125 & -0.0625 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},\quad a(0)=\begin{pmatrix} 0.5 & 0 & 0.25 & -0.75 \\ 0 & 0.25 & 0.03125 & -0.0625 \\ 0 & 0 & 0.25 & 0.75 \\ 0 & 0 & -0.03125 & -0.0625 \end{pmatrix},$$

$$a(1)=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.25 & -0.75 \\ 0 & 0.25 & 0.03125 & -0.0625 \end{pmatrix}.$$

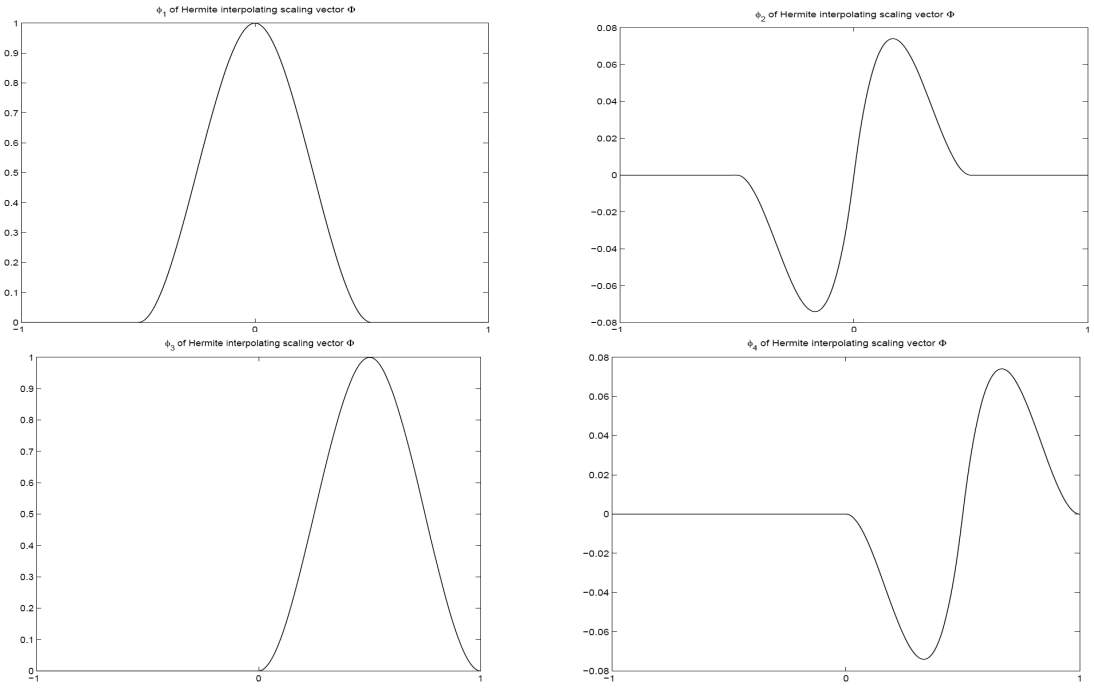


Fig.1 Hermite interpolating refinable function vector $\phi=[\varphi_1,\cdots,\varphi_4]^T$

References

[1] CHENG Li-bo,SUN Jia-ning,LIANG Xue-zhang,WEN Xue-bing. Refinable Function Vectors with Hermite Interpolating Property. Journal of Jilin University Science Edition.Vol. 47 (2009), 1155-1159

[2] Koch K. Interpolating Scaling Vectors. Int J Wavelets Multiresolut Info Process. Vol. 3 (2005), 389-416

[3] ZHOU Ding-xuan. Interpolatory Orthogonal Multiwavelets and Refinable Functions. IEEE Trans on Signal Process , 2002,50 (3):520-527

[4] Selesnick I W. Interpolating Multiwavelet Bases and the Sampling Theorem. IEEE Trans on Signal Process , 1999,47 (6): 1615-1621

[5] Conti C, Cotronei M, Sauer T. Full Rank Interpolatory Subdivision Schemes:Kronecker,Filters and Multiresolution. J Comput Appl Math,Vol. 225 (2009), 219-239

[6] HAN Bin, Kwon S G, ZHUANG Xiao-sheng. Generalized Interpolating Refinable Function Vectors. J Comput Appl Math,Vol. 227 (2009), 254-270

[7] HAN Bin. Vector Cascade Algorithms and Refinable Function Vectors in Sobolev Spaces. J Approx Theory ,Vol. 124 (2003), 44-88

[8] JIANG Qing-tang,Oswald P, Riemenschneider S D. $\sqrt{3}$-Subdivision Schemes:Maximal Sum Rule Orders. Constru Approx ,Vol. 19 (2003), 437-463

[9] Koch K. Multivariate Orthonormal Interpolating Scaling Vectors. Appl Comput Harmon Anal, Vol. 22 (2007), 198-216

doi:10.4028/www.scientific.net/AMM.50-51.377

A note on *q*-derivative operator and divided difference

Feng Cui

College of Statistics and Mathematics,

Zhejiang Gongshang University, Hangzhou 310018, China

fcui@zjgsu.edu.cn

Keywords: Divided difference; q -Derivative; q -Leibnitz formula

Abstract. The *q*-analogue of the derivative operator is playing a more and more important role in mathematics and physics. Moreover, the divided difference, as an important and classical mathematical tool with a close relation to the derivative, is also used in many fields. In this paper, the connection between the *q*-derivative and the divided difference is investigated such that the *q*-derivative can be understood better. The *q*-derivative of higher order of a function *f* can be represented by the divided difference of *f* at some special nodes. Furthermore, the result is used to provide a new and easier proof of *q*-Leibnitz formula and its generalization.

Introduction

Recently, the *q*-analogue of the derivative operator is paid more and more attention ever since the pioneering work of Jackson [1]. Usually, it is defined by

$$D_q f(x) := \frac{f(x)-f(qx)}{(1-q)x}, \tag{1}$$

where q is fixed and $0<q<1$. We call this the *q*-derivative of $f(x)$.

As J. Koekoek and R. Koekoek pointed out in [2], the definition as Eq. 1 is not valid for $x=0$. Hence, they gave a new definition as follows,

$$D_q f(x) =: \begin{cases} \dfrac{f(x)-f(qx)}{(1-q)x}, & x\neq 0, \\ f'(0), & x=0, \end{cases} \tag{2}$$

where q is a fixed real number not equal to 1 and f is differentiable at $x=0$. Moreover, they also presented the expression for the high order *q*-derivative of f at $x=0$. The divided difference, as an important mathematical tool with a close relation to the derivative, is used not only in the interpolation [3], numerical differentiation [4], but also in the combinatorics [5], etc.

In this paper, we investigate the connection between the *q*-derivative of higher order and the divided difference for any $x, q\in\mathbb{R}$ but $q\neq 1$, and provide the divided difference form representation for the high order *q*-derivative.

Preliminaries

Firstly, we introduce the requisite notation. Given an array of nodes $x_0\le x_1\le\cdots\le x_n$, the divided differences of f are defined recursively by

$$f[x_0,x_1,\ldots,x_n]=\begin{cases}\dfrac{f[x_1,x_2,\ldots,x_n]-f[x_0,x_1,\ldots,x_{n-1}]}{x_n-x_0}, & x_n\neq x_0,\\ \dfrac{1}{n!}f^{(n)}(x_0), & x_n=x_0,\end{cases}$$

where $f[x_0]=f(x_0)$.

We put $[n]=(1-q^n)/(1-q)=1+q+\cdots+q^{n-1}$, and define the q-factorial $n!_q=[1][2]\cdots[n]$, where $0!_q=1$. Then the q-multinomial coefficient may be defined by

$$\begin{bmatrix} n \\ k_1,k_2,\cdots,k_m \end{bmatrix}=\frac{n!_q}{(k_1)!_q(k_2)!_q\cdots(k_m)!_q}.$$

where $k_1+k_2+\cdots+k_m=n$. Mostly, we will use the q-binomial coefficient $\begin{bmatrix} n \\ k,n-k \end{bmatrix}$, which is written as $\begin{bmatrix} n \\ k \end{bmatrix}$. As for $\binom{n}{k}$, it is the binomial coefficient which is $\dfrac{n!}{k!(n-k)!}$.

Now, we shall state some lemmas which will be needed in the proofs of our theorems.

Lemma 1 For $q\neq 0,\pm 1$, we have

$$\sum_{i=0}^{n-1}\frac{q^{ki}}{\prod_{j\neq i}(q^i-q^j)}=\begin{cases}0, & k=0,1,\ldots,n-2,\\ 1, & k=n-1,\\ [n], & k=n.\end{cases}$$

Proof Considering the divided difference of the function f at a sequence of nodes $\{q^i\}_{i=0}^{n-1}$, it is well known that

$$f[1,q,\ldots,q^{n-1}]=\sum_{i=0}^{n-1}\frac{f(q^i)}{\prod_{j\neq i}(q^i-q^j)}. \tag{3}$$

Let $f(x)=1,x,\ldots,x^{n-1}$, respectively. According to [6], for $k=0,1,\ldots,n-2$, the divided differences of x^k at $\{q^i\}_{i=0}^{n-1}$ are zeros, and the divided difference of x^{n-1} at $\{q^i\}_{i=0}^{n-1}$ is equal to 1. Moreover, let $f(x)=x^n$, then we have

$$f[1,q,\ldots,q^{n-1}]=\sum_{i=0}^{n-1}q^i=\frac{1-q^n}{1-q}.$$

Therefore, combine the above equation with Eq. 3, and we obtain the lemma.

Lemma 2 (see [7]) Given an array of nodes $x_0<x_1<\cdots<x_n$, we have

$$f[x_0^{k_0+1},x_1^{k_1+1},\ldots,x_n^{k_n+1}]=\frac{1}{k_0!k_1!\cdots k_n!}\cdot\frac{\partial^{k_0}}{\partial x_0^{k_0}}\frac{\partial^{k_1}}{\partial x_1^{k_1}}\cdots\frac{\partial^{k_n}}{\partial x_n^{k_n}}f[x_0,x_1,\ldots,x_n],$$

where $x_i^{k_i+1}$ means that the point x_i is repeated k_i+1 times.

Lemma 3 For $x\neq y$, suppose that $f^{(k)}(x)$ and $f^{(m)}(y)$ exist, then

$$f[x^{k+1}, y^{m+1}] = \sum_{v=0}^{k} (-1)^{k-v} \binom{m+k-v}{k-v} \frac{f^{(v)}(x)}{v!(x-y)^{m+k+1-v}} + \sum_{v=0}^{m} (-1)^{m-v} \binom{m+k-v}{m-v} \frac{f^{(v)}(y)}{v!(x-y)^{m+k+1-v}}. \quad (4)$$

In particular, for $m = 0$, we have

$$f[x^{k+1}, y] = (y-x)^{-k-1} \left(f(y) - \sum_{v=0}^{k} \frac{(y-x)^v}{v!} f^{(v)}(x) \right).$$

Proof From Lemma 2, we have

$$f[x^{k+1}, y^{m+1}] = \frac{1}{k!m!} \cdot \frac{\partial^k}{\partial x^k} \frac{\partial^m}{\partial y^m} \left(\frac{f(x)}{x-y} + \frac{f(y)}{y-x} \right)$$

$$= \frac{1}{k!} \sum_{v=0}^{k} \binom{k}{v} f^{(v)}(x) \left(\frac{1}{(x-y)^{m+1}} \right)^{(k-v)} + \frac{1}{m!} \sum_{v=0}^{m} \binom{m}{v} f^{(v)}(y) \left(\frac{1}{(y-x)^{k+1}} \right)^{(m-v)}.$$

Hence, we immediately obtain Eq. 4.

Main Results

In this section, we present our theorems.

Theorem 1 If assume that $f^{(n)}(x)$ exists at $x = 0$ for $q \neq 1$, then the following formula holds

$$D_q^n f(x) = n!_q \, f[x, qx, \ldots, q^n x], \quad (5)$$

where $D_q^n f(x)$ is the *n*th order *q*-derivative of $f(x)$.

Proof We will prove the theorem by induction for $q \neq \pm 1$, because for $q = -1$, the result is trivial. Clearly, according to the definition of the divided difference, it is true for $n = 1$. Assuming that Eq. 5 are valid for $n-1$, we show it also holds for *n*. Applying the definition of *q*-derivative of high order, we have

$$D_q^n f(x) = D_q(D_q^{n-1} f(x)) = \begin{cases} \dfrac{D_q^{n-1} f(x) - D_q^{n-1} f(qx)}{(1-q)x}, & x \neq 0, \\ (D_q^{n-1} f(x))' \,|_{x=0}, & x = 0 \end{cases}$$

$$= \begin{cases} (n-1)!_q \dfrac{f[x, qx, \ldots, q^{n-1}x] - f[qx, q^2x, \ldots, q^n x]}{(1-q)x}, & x \neq 0, \\ (n-1)!_q \lim\limits_{t \to 0} \dfrac{f[t, qt, \ldots, q^{n-1}t] - f[0, 0, \ldots, 0]}{t}, & x = 0. \end{cases}$$

Firstly, we consider the case of $x \neq 0$. According to the recurrence relations of the divided difference, we have

$$f[x, qx, \ldots, q^n x] = \frac{f[x, qx, \ldots, q^{n-1}x] - f[qx, q^2x, \ldots, q^n x]}{(1-q^n)x}.$$

Thus, for $x \neq 0$, it follows that

$$D_q^n f(x) = (n-1)!_q \frac{1-q^n}{1-q} f[x, qx, \ldots, q^n x] = n!_q f[x, qx, \ldots, q^n x].$$

For $x = 0$, we divide the proof into two steps.

Step 1: Suppose that $q = 0$. By the definition of the divided difference and Lemma 3, we have the followings,

$$\frac{f[t, 0^{n-1}] - f[0^n]}{x} = \frac{\frac{1}{x^{n-1}}\left(f(t) - \sum_{v=0}^{n-2} \frac{f^{(v)}(0)}{v!} t^v\right) - \frac{f^{(n-1)}(0)}{(n-1)!}}{x} = \frac{f(t) - \sum_{v=0}^{n-1} \frac{f^{(v)}(0)}{v!} t^v}{t^n}$$

$$= \frac{\frac{f^{(n)}(0)}{n!} t^n + o(t^n)}{t^n} = f[0^{n+1}] + o(1), \quad t \to 0.$$

Therefore, from $(n-1)!_q = n!_q = 1$, it follows that $D_0^n f(x) = n!_q f[0^{n+1}]$.

Step 2: In the case of $q \neq 0$, it follows from Taylor's expansions that

$$f(q^i t) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} \left(q^i t\right)^k + o(t^n), \quad t \to 0, \quad i = 0, 1, \ldots n-1. \tag{6}$$

Hence,

$$\frac{f[t, qt, \cdots, q^{n-1}t] - f[0^n]}{t} = \frac{\sum_{i=0}^{n-1} \frac{f(q^i t)}{\prod_{j \neq i} (q^i t - q^j t)} - f[0^n]}{t}.$$

Furthermore, it follows from Eq. 6 that

$$\frac{f[t, qt, \cdots, q^{n-1}t] - f[0^n]}{t} = \frac{1}{t^n}\left(\sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} \sum_{i=0}^{n-1} \frac{q^{ik}}{\prod_{j \neq i} (q^i - q^j)} - f[0^n] t^{n-1} + o(t^n)\right), \quad t \to 0.$$

Due to Lemma 1, we have

$$\frac{f[t, qt, \ldots, q^{n-1}t] - f[0^n]}{t} = \frac{f^{(n)}(0)}{n!} \sum_{i=0}^{n-1} \frac{q^{in}}{\prod_{j \neq i} (q^i - q^j)} + o(1) = [n] f[0^{n+1}] + o(1), \quad t \to 0.$$

Therefore the proof is completed.

It is very useful that $D_q^n f(x)$ is represented by the divided difference of f at the nodes $x, qx, \ldots, q^n x$. As an application of Theorem 1, we can easily verify the q-Leibnitz rule and its generalization in [8].

Example 1 $D_q^n \prod_{i=1}^{m} f_i(x) = \sum_{k_1 + \cdots + k_m = n} \prod_{i=1}^{m} D_q^{k_i} f_i(q^{k_0 + k_1 + \cdots + k_{i-1}} x), \quad k_0 = 0.$

In particular, we have the q-Analogue of Leibnitz rule

$$D_q^n[f(x)\cdot g(x)] = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} D_q^k f(x)\cdot D_q^{n-k} g(q^k x).$$

In fact, if let $H = \prod_{i=1}^{m} f_i$, it follows from Theorem 1 that

$$D_q^n H(x) = n!_q\, H[x, qx, \ldots, q^n x].$$

By the generalization of the Steffensen formula [9], we know that

$$H[x, qx, \ldots, q^n x] = \sum_{0=k_0<k_1<\cdots<k_{m-1}<k_m=n} \prod_{i=1}^{m} f_i[q^{k_{i-1}} x, \ldots, q^{k_i} x].$$

Furthermore, applying Theorem 1 to the above formula again, we obtain

$$H[x, qx, \ldots, q^n x] = \sum_{0=k_0<k_1<\cdots<k_{m-1}<k_m=n} \prod_{i=1}^{m} \frac{1}{(k_i - k_{i-1})!_q} D_q^{k_i-k_{i-1}} f_i(q^{k_i} x)$$

$$= \frac{1}{n!_q} \sum_{0=k_0<k_1<\cdots<k_{m-1}<k_m=n} \begin{bmatrix} n \\ k_1, k_2-k_1, \ldots, k_m - k_{m-1} \end{bmatrix} \times \prod_{i=1}^{m} D_q^{k_i-k_{i-1}} f_i(q^{k_i} x).$$

If let $v_0 = 0, v_i = k_i - k_{i-1}, i = 1, 2, \ldots, m$, we get

$$H[x, qx, \ldots, q^n x] = \sum_{v_1+v_2+\cdots+v_m=n} \begin{bmatrix} n \\ v_1, v_2, \ldots, v_m \end{bmatrix} \prod_{i=1}^{m} D_q^{v_i} (q^{v_0+\cdots+v_{i-1}} x).$$

Acknowledgement

The work is supported by Zhejiang Provincial Natural Science Foundation of China (Y6100126).

References

[1] F.H. Jackson: Trans. R. Soc. Edin. Vol. 46 (1908), pp. 253--281; F.H. Jackson: Quart. J. Math. Oxford Ser. Vol. 2 (1951), pp. 1--16.

[2] J. Koekoek and R. Koekoek: J. Math. Anal. Appl. Vol. 176 (1993), pp. 627--634.

[3] M. El-Mikkawy and B. El-Desouky: Appl. Math. Comput. Vol. 138 (2003), pp. 375--385.

[4] H.Y. Wang, F. Cui and X.H. Wang: Acta. Math Sin (Engl Ser), Vol. 23 (2007), pp. 365--372.

[5] X.H. Wang and A.M. Xu: J. Comput. Math. Vol. 25 (2007), pp. 697--704.

[6] C. de Boor: Surv. Approx. Theory Vol. 1 (2005), pp. 46--69.

[7] E. Isaacson and H. B. Keller: Analysis of Numerical Methods (Wiley, New York 1966).

[8] W.P. Johnson: J. Combin. Theory Ser. A Vol. 76 (1996), pp. 305--314.

[9] J.F. Steffensen: Danske Vid. Selsk. Math.-Fys. Medd. Vol. 17 (1939), pp. 1--12.

©2010 Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.382

Simulation of General 3D Virtual Stochastic Road Model

Min Lin[1,a] , Xiangwei Zhang[1,b] and Siyuan Cheng[2,c]

[1] Department of Automobile Engineering, Faculty of Electromechanical Engineering, Guangdong University of Technology, Guangzhou 510006, China

[2] Department of Electromechanical Engineering, Faculty of Electromechanical Engineering, Guangdong University of Technology, Guangzhou 510006, China

[a]linmin3000@163.com, [b]xzxx@gdut.edu.cn, [c]imdesign@gdut.edu.cn

Keywords: Pseudo Excitation Method, Mathematical model, 3D virtual stochastic road, Ride performance simulation.

Abstract. In order to provide the research of virtual prototype a complex 3-D virtual road environment based on a new method，a 3-D Pseudo Excitation Method stochastic road model was established based on the single-point Pseudo Excitation Method model, and then the simulation result was gained by Matlab.What's more，based on the theory of triangle network，a general 3-D virtual road model with the arithmetic of nodes and elements was created.Then a specified 3-D virtual road file was imported into Adams to form the 3-D virtua1 road. Finally, a C-level virtual road was generated for the riding performance simulation of virtual vehicle.By comparing the simulation result with the GB / T 13441-92，the virtual vehicle's reasonable time under the ride performance requirement was obtained.

Introduction

Along with the modern hydrodynamics' development and integration application, the modern vehicle systems usually manifest nonlinear property. There are lots of random nonlinear problems in the nonlinear vehicle constitutive system and random road. As the important index in the cars comfort, many scholars have made progress in the riding comfort at home and abroad. At the same time, virtual prototype technology has been applied to the great extent, which is one of the most important methods. Meanwhile most of the virtual prototype technology usually focus on the design and performance of the prototype, the road condition of the virtual prototype is less researched[1]. At present, the authors have not studied on the complex 3-Dimensional road model universality, especially the one which is based on the Pseudo Excitation Method.

The roughness of road models were studied in the process, which went through from linear to nonlinear, from frequency domain to time domain, from mathematic model to simulation output . For example, the time frequency model was studied from one point to multi-point, from one rut to double rut, from two-dimension to three-dimension process, at home and abroad. In China, there are many series of deep research about multi-point and double-rut model, especially the three-dimension random road model research which brought many useful productions[1]. But the 3-Dimension random road model system is still need to consummate and further. In this paper, the authors want to give a new kind of system with 3-Dimension random road, which can be more easily and quickly set up on Matlab. On the other hands, according to the vehicle riding performance, the authors finally also give the virtual prototype simulation in the condition of 3-Dimensional road model.

PEM between is accurate because the cross-correlation terms between the participant modes and between the excitations are both included in the results. It is also convenient to use because the stationary random analyses transformed accurately into harmonic non-stationary random easily analyses are ones, while transformed accurately into very solved deterministic transient ones.

Pseudo excitation method on the 3-dimensional random road

Method introduction. There were four methods of Time Domain Model to describe the road roughness: the white noise method[2], the harmonic superpose methods[3], the disperse sampling methods based on the PSD[4], and the disperse time sequence method based on AR/ARMA[5]. But all of these methods could not obtain exact road roughness, the authors introduced a new method which could process directly the given power spectral density. Considering the effect of the vehicle natural vibration frequency and velocity completely, the simulation and analysis were done based on the C grade of the road which is used by this method. The road roughness was calculated by the Pseudo Excitation Method, which was also called fast complete quadric combination method. The pseudo-excitation method, in which random excitations are mathematically represented by a series of harmonic loads, is developed for vibration analysis of vehicle of closely spaced natural frequencies or equipped with discrete damping devices. By using this method, all cross-correlation terms between normal modes in a structural response can be retained naturally, and the non-orthogonal structural damping or stillness properties can be handled accurately. The accuracy of the method is validated which has proved an effective method by comparing with the transfer matrix method. A road roughness simulation is to be set to demonstrate the advantages of the suggested method.
The altitude of one random sample of the road is showed below (Eq.1)

$$q(n)=\sum_{k=0}^{N_r-1}\sqrt{G_{qq}(f=k\Delta f)}e^{i(\frac{2\pi kn+\phi_n}{N_r})} \tag{1}$$

where $G_{qq}(f=k\Delta f)=\dfrac{\alpha\beta^2}{u\pi(\alpha^2+(f/u)^2)}$ and u is the velocity, α,β is the constant of the road grade.

There is an example of road roughness simulation, the α =0.12m^{-1}, β =0.006m, u =50km/h=13.89m/s. The simulation extension and spatial frequency of statistical analysis are limited to $0.011\text{m}^{-1}\le\Omega\le2.83\text{m}^{-1}$, which is corresponding to each other[1].

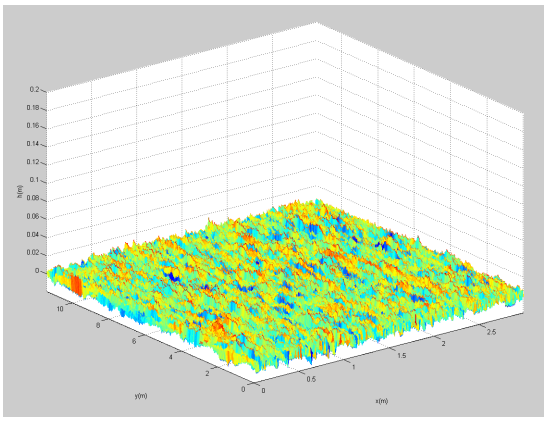


Fig. 1 3-Dimension road roughness

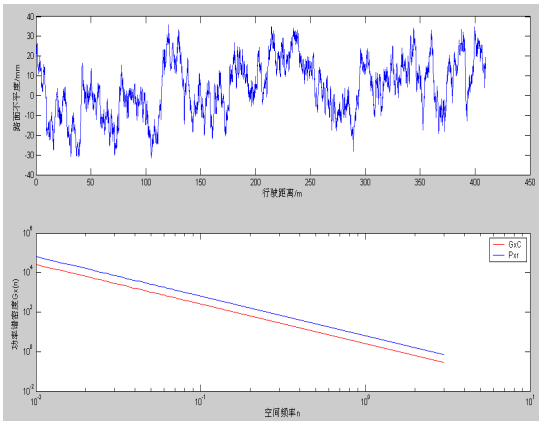


Table.1 z-direction roughness & standard model

The differences from the PEM internal forces at the road roughness are those listed in the captions of Fig. 1. Not only did such peak values change by 21% for SH waves or 35% for P waves, but also the original symmetry in the responses was lost. This is because the symmetric modes do not participate in the seismic response when all supports move uniformly, for which the loading is anti-symmetric, whereas all modes participate in the seismic responses when the wave passage effect is taken into account, so that the support move ments do not exhibit any symmetric or antisymmetric behaviour.

The above computations were executed on a Pentium-4 personal computer with main frequency 2.2 GHZ and the z-direction roughness & standard model built by the three methods are listed in Table 1. When the ground motion is assumed to be uniform, it can be seen that although PEM and RSM

produce very close results, PEM is more excient than RSM. However, for the wave passage effect is very important and so the use of conventional RSM, which ignores it, is inappropriate. Although THM can account for this effect, computation is required for quite a few samples in order to obtain a reliable statistical average, which is very time consuming, whereas PEM is both reliable and excient. Comparatively, the incoherence effect is not so important as the wave passage effect, whereas it is computationally very expensive. Therefore for ordinary road roughness which are not of very special importance, this factor can be neglected, but the wave passage effect must be considered.

For multiple excitation problems, the PSD matrices of the ground acceleration excitations have the form

$$[S_{XX}(w)]=\begin{bmatrix} S_{\ddot{X}_1\ddot{X}_1}(w) & S_{\ddot{X}_1\ddot{X}_2}(w) & \bullet\bullet\bullet\bullet & S_{\ddot{X}_1\ddot{X}_n}(w) \\ \bullet\bullet\bullet\bullet & \bullet\bullet\bullet\bullet & \bullet\bullet\bullet\bullet & \bullet\bullet\bullet\bullet \\ S_{\ddot{X}_1\ddot{X}_1}(w) & S_{\ddot{X}_1\ddot{X}_n}(w) & \bullet\bullet\bullet\bullet & S_{\ddot{X}_n\ddot{X}_n}(w) \end{bmatrix} \tag{2}$$

This new method is frequency field simulation of pseudo-excitation method(PEM) and the method of mathematical modeling for automobile vibration structure based on ideas of finite element method and the numerical simulation method of time field based on variation theories. They have been applied to study modeling and simulation of automobile ride comfort, good conclusions have been gained[6-8].

The basic principles of PEM, particularly the analysis for spatial effects, are presented in this paper. Additionally, numerical results are given for a real road, and these are compared with those obtained from the conventional FFT. It is shown (in Table.1)that if all ground nodes are assumed to move uniformly, the two methods give quite close results. However, when the wave passage effect is taken into account, the results given by FFT and by PEM are quite close together. Steps of pseudo-excitation method applied are introduced for simulation of automobile ride comfort. Different characteristics between FFT and PEM are compared between pseudo-excitation method and Fourier analysis method. Effectiveness of pseudo-excitation method is discussed for simulation of full automobile ride comfort[9-12]. Results show that pseudo-excitation method is equal to Fourier analysis method for simulation of automobile ride comfort, but it is better than Fourier analysis method. Moreover, the results given by the conventional FFT can deviate from them considerably because RSM ignores the spatial effects. Of the two methods, PEM alone is both quite easy to use and very efficient[13].

Stationary random vibration pseudo-excitation method is a new method for simulation of time-frequency field, but it is seldom applied in auto domain. Based on different characteristics of pseudo excitation, pseudo excitation can be noted as single-point pseudo excitation and generalized-point pseudo excitation and multi-point pseudo excitation. Basic theories of stationary random vibration pseudo-excitation method are discussed from simple to more difficult ones, rigorous derivations are given, which further enrich and perfect theory of pseudo-excitation method[14-16].

Based on ideas of finite element method automobile vibration structure is divided, element and global analysis are done, bind disposals were disposed as forced displacements, then, matrixes of factors for mathematical model of automobile vibration structures, assembly method of matrixes of factors is introduced. Mathematical model of automobile vibration structures is assembled by hand, and effectiveness of method is tested for mathematical modeling[17]. Modules are programmed for element analysis and global analysis and bind disposals. So the matrixes of factors of mathematical model can be assembled not only by hand but also by program, flexibility of method is indicated, the problems are resolved for complex of mathematical model of automobile vibration structures.

Summary

This paper discusses virtual excitation method and constructs virtual road excitation for discussing the feasibility of this method applied to automotive random vibration. According to one quarter freedom model of automobile,the frequency-response characteristics of the system and system vibration responses are deduced based on virtual excitation method,a new method is put forward for calculating power spectrum density of system vibration responses,and an example is given for automotive random vibration.The results show that virtual excitation method is a more simple research method in comparison with Fourier analysis method. In this result, we can see the road roughness can be set up in the 3-Dimension. The result is compared with standard picture, and they are close to each other.

References

[1] XuYanhai.Computer simulation on stochastic road rregularities[J].Transactions of the Chinese Society for Agricultural Machinery, Vol.38(1)(2007), p.33-36.
[2] Zhang Yonglin.Research on time series modeling method of 3D stochastic road[J]. Journal of Wuhan University of Technology, Vol.29(8) (2007), p.92-95.
[3] Zhang Jiafan.Numerical simulation of stochastic road process using white noise filtration[J]. Mechanical Systems and Signal Processing, Vol.20(2) (2006), p.363-372.
[4] Yoshimura A.Semi-active suspension of passenger cars using fussy reasoning and the field testing[J].Int.J.of Vehicle Designs,Vol.19(2) (1998), p.150-166.
[5] Zhang Yonglin.Time domain model of road irregularities simulated using the harmony superposition method[J].Transactions of the Chinese Society of Agricultural Engineering, Vol.19(6) (2003), p.32-35.
[6] Jiang Li-biao,Wang Deng-feng, Xie Dong et al.Buildup and simulation of a truck model with rear air suspension using leaf spring as guiding rod[J]. Journal of Harbin Institute of Technology(new series), Vol.13(6) (2006), p.654-657.
[7] Lin Jiahao,Williams F W.Computation and analysis of multi-excitation random seismic responses [J]Engineering Computations, Vol.9(5) (1992), p.561-574.
[8] Carassale L, Tubino F, Solari G. Seismic response of multi-sup-ported structures by proper orthogonal decomposition. Proceed-ings of International Conference on Advances in Structural Dynamics (ASD2000), Hong Kong. Elsevier Science Ltd, (2000), p. 827-34.
[9] Wang ML, Heo G, Satpathi D. Dynamic characterization of a long span bridge: a finite element based approach. Soil Dynam Earthq Eng, Vol. 16(7-8)(1997), p.503-12.
[10]Xue SD, Cao Z, Wang XS. Random vibration study of struc- tures under multi-component seismic excitations. Adv Struct Eng, Vol. 5(3)(2002), p.185-92.
[11]Lee M, Penzien J. Stochastic analysis of structures and piping systems subjected to stationary multiple support excitations. Earthq Eng Struct Dynam, Vol. 11(1)(1983), p.91-110.
[12]Berrah M, Kausel E. Response spectrum analysis of structures subjected to spatially varying motions. Earthq Eng Struct Dynam, Vol. 21(6)(1992), p.70-81.
[13]Lin YK, Zhang R, Yong Y. Multiply supported pipeline under seismic wave excitations. J Eng Mech,Vol. 116(6)(1990), p.108-109.
[14]Lin JH, Zhang WS, Li JJ. Structural responses to arbitrarily coherent stationary random excitations. Comp Struct, Vol. 50(5)(1994), p.629-33.
[15]Clough RW, Penzien J. Dynamics of structures. New York: McGraw-Hill Inc,(1993).
[16]European Committee for Standardization, Eurocode 8: Structures in seismic regions-design. Part 2: Bridge. Brussels, (1995).
[17]National Standard of The People's Republic of China, Code for seismic design of buildings GBJ 11-89, Beijing, (1989) (in Chinese).

©2010 Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.386

Searching Best Strategies Algorithm For The No Balance Assignment Problem

Maoyan Fang [a], Minle Wang [b], Yiming Bi [c]

The Secondary Artillery Engineer Institute, Shanxi Xi'an, China

[a]fmy19791019@sohu.com, [b] wang-ml@sohu.com, [c] bjym@cetin.net.cn

Key words: Overall situation searching; Assignment problem; Cost matrix

Abstract: The No Balance Assignment Problem (NBAP) is mainly resolved by changing it into Balance Assignment Problem (BAP) and using classical algorithm to deal with it now. This paper proposed Searching Best strategies Algorithm (SBSA) to resolve this problem, and it needn't to change NBAP into BAP. SBSA resolves NBAP based on searching the best answer of the cost matrix. This algorithm's theory is simple, and it is easy to operate. The result of the research indicate that the algorithm not only can deal with NBAP, but also can deal with BAP and other problems such as translation problem.

Introduction

The assignment problem (AP) is that some jobs be assigned to some workers so that the cost (or income) is smallest (or largest). Suppose there are m jobs and n workers. AP is named no balance assignment problem (NBAP), when $m \neq n$. When $m = n$, AP is named balance assignment problem (BAP). Now the method often be used to deal with NBAP is changing the NBAP into BAP, and so resolved by Hungarian Approach (HA) or GA.

Hungarian Approach has two defects[1].

The first is in the second step of HA. It uses the method that "try to assign" to find the independence zero elements. The theory is simple and getting the elements is quickly, but there are $n!$ perhaps lines for these elements to a $n \times n$ matrix.

The second is in the third step of HA. In this step, using the least number of lines cover all the zero elements. One time the rows are covered, one time the columns are covered in the course, especially the theory of the method is hard to understand for most of the people. They have to do the job according to the book step by step, if there are not book about HA, people hardly finish the job.

GA is fast in the large-scale calculation, but the answer perhaps is not the best .

So this paper proposed Searching Best Strategies Algorithm (SBSA). This algorithm's theory is simple, and it is easy to operate and program by compute.

NBAP's SBSA

Let us suppose that there are m jobs and n workers, $n \geq m$, and one worker does one job or doesn't do job, and every job must be done. The cost that job i is assigned to worker j is c_{ij}, and looking for an assignment strategy make the cost is least.

The SBSA is an advanced greedy method. Because every job must be done by one worker, so if the children assignment strategy's cost is least, the cost is the least. The algorithm is that look for

the least cost in every cost matrix. Since $n \geq m$, and every job must be taken, so when a job only can be done by one worker, the job must be assigned to the worker. According to this principle, and we got the SBSA's process.

The SBSA's Process

Step1: Building the cost matrix C, and the cost job i be assigned to worker j is c_{ij}. If worker j can't undertake job i, $c_{ij} = 0$.

Step2: Judging whether there are rows that only one element is not zero, if there are these rows, and the position of the not zero element is (i, j), the job i would be assigned to worker j. All elements of the row and column that the not zero element in were changed into zero.

Step3: Looking for the minimal elements of the matrix, and marked the position of the elements. If only one minimal element, we would choice the method of the element. If there were more than one minimal elements, for example there were n minimal elements, put all the elements of the row and column that the minimal element in into a row from little to large, then we would get n rows. Compared the elements in the same position of the rows from first position,, until only one maximal element in the same position of the rows or reach the last position, then stop comparing, and choice the method of the element that its row has the maximal element, and all elements of the row and column that the element in were changed into zero.

Step4: do step2 and step3, until all jobs were assigned.

2.2 The Advantage Of SBSA

Comparing to HA and GA, SBSA has its advantage.

First, the theory of SBSA is simple, and it's easy to operate. It needs to create the cost matrix only, and get the strategy by operating the matrix. It is simpler than HA and GA.

Second, the answer get from SBSA is better than other algorithms. From the step3, SBSA looks for the minimal element from overall situation, while get rid of the largest element, makes the later elements almost little, so get the best answer.

Third, this algorithm has wild use. The algorithm could deal with every assignment problem, and this algorithm can also deal with other problems, such as translation problem.

Simulation

For example, there were 7 jobs and 13 workers. Every job only was done by one worker, the cost matrix was:

$$C = \begin{bmatrix} 2 & 5 & 3 & 6 & 7 & 5 & 0 & 4 & 3 & 0 & 2 & 3.5 & 4 \\ 3 & 4 & 2 & 1 & 6 & 4 & 2 & 0 & 3 & 7 & 4 & 5 & 0 \\ 2.5 & 6 & 4 & 2 & 5 & 4.5 & 6 & 5 & 3 & 2 & 0 & 0 & 0 \\ 0 & 3 & 5 & 4 & 2 & 1.5 & 4 & 3 & 2 & 0 & 1.5 & 4 & 3 \\ 5 & 2 & 7 & 0 & 4 & 6 & 8 & 2 & 3.5 & 4 & 0 & 3 & 4 \\ 4 & 3 & 3.5 & 2 & 0 & 5 & 0 & 6 & 2 & 8 & 4 & 3 & 2 \\ 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 5 & 3 & 0 & 0 \end{bmatrix}$$

First, judging whether there were rows that only one element was not zero, there was not this row. Then looking for the minimal element, there was only one minimal element in matrix C, that is 1,

the position was (2,4), so job 2 was assigned to worker 4, the cost is 1, and the elements of row 2 and column 4 were changed into zeros. Then the cost matrix was:

$$C = \begin{bmatrix} 2 & 5 & 3 & 0 & 7 & 5 & 0 & 4 & 3 & 0 & 2 & 3.5 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2.5 & 6 & 4 & 0 & 5 & 4.5 & 6 & 5 & 3 & 2 & 0 & 0 & 0 \\ 0 & 3 & 5 & 0 & 2 & 1.5 & 4 & 3 & 2 & 0 & 1.5 & 4 & 3 \\ 5 & 2 & 7 & 0 & 4 & 6 & 8 & 2 & 3.5 & 4 & 0 & 3 & 4 \\ 4 & 3 & 3.5 & 0 & 0 & 5 & 0 & 6 & 2 & 8 & 4 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 3 & 0 & 0 \end{bmatrix}$$

Continue to judge whether there were rows that only one element was not zero, there was not this row. Then looking for the minimal element, the minimal element was 1.5, and there were two, the positions were (4, 6), (4,11), and they were in the same row, so we only need to put the elements of the columns that they were in into two rows, one was 4.5, 5, 5, 6, the other was 2, 3, 4. Comparing the elements of the first position, because 4.5>2, we selected the first, so the job 4 was assigned to worker 6, the cost was 1.5, and the elements of row 4 and column 6 were changed into zeros. Then the cost matrix was:

$$C = \begin{bmatrix} 2 & 5 & 3 & 0 & 7 & 0 & 0 & 4 & 3 & 0 & 2 & 3.5 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2.5 & 6 & 4 & 0 & 5 & 0 & 6 & 5 & 3 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 2 & 7 & 0 & 4 & 0 & 8 & 2 & 3.5 & 4 & 0 & 3 & 4 \\ 4 & 3 & 3.5 & 0 & 0 & 0 & 0 & 6 & 2 & 8 & 4 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 3 & 0 & 0 \end{bmatrix}$$

Continue to judge whether there were rows that only one element was not zero, there was not this row. Then looking for the minimal element, the minimal element was 2, and there were seven minimal elements, the positions were (1,1), (1, 11), (3, 10), (5, 2), (5,8), (6, 9), (6, 13), the elements of the row and the column that the minimal elements in put into seven rows, that was:

2	2.5	3	3	3.5	4	4	4	5	5	7
2	3	3	3	3.5	4	4	4	5	7	
2.5	3	4	4	5	5	5	6	6	8	
2	3	3	3.5	4	4	4	5	5	6	7	8
2	3	3.5	4	4	4	4	5	5	6	7	8
2	3	3	3	3	3.5	3.5	4	4	6	8	
2	3	3	3.5	4	4	4	4	6	8		

Continue to comparing the elements in the first position of these rows, 2.5 was the maximal element, and only one maximal elements, so we selected (3,10) as the choice. Since, the job 3 was assigned to worker 10, and the cost was 2, and the elements of row 3 and column 10 were changed into zeros. Then the cost matrix was:

$$C = \begin{bmatrix} 2 & 5 & 3 & 0 & 7 & 0 & 0 & 4 & 3 & 0 & 2 & 3.5 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 2 & 7 & 0 & 4 & 0 & 8 & 2 & 3.5 & 0 & 0 & 3 & 4 \\ 4 & 3 & 3.5 & 0 & 0 & 0 & 0 & 6 & 2 & 0 & 4 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \end{bmatrix}$$

Continue to judge whether there were rows that only one element was not zero, there was one row——seventh row, the non zero element was 3, position was (7,11), so job 7 was assigned to worker 11, the cost was 3, and the elements of row 7 and column 11 were changed into zeros. Then the cost matrix was:

$$C = \begin{bmatrix} 2 & 5 & 3 & 0 & 7 & 0 & 0 & 4 & 3 & 0 & 0 & 3.5 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 2 & 7 & 0 & 4 & 0 & 8 & 2 & 3.5 & 0 & 0 & 3 & 4 \\ 4 & 3 & 3.5 & 0 & 0 & 0 & 0 & 6 & 2 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

As this way, we got the strategy that was (2, 4), (4, 6), (3, 10), (7, 11), (1, 1), (5, 8), (6, 13), and the cost of the children strategy were 1, 1.5, 2, 3, 2, 2, 2, the cost of the strategy was 13.5.

Next, we used GA to resolve the problem. In order to operate easily, the strings of the population were made up of integers, and the length of the strings is 20. The integers in the first seven positions expressed jobs, and the integers in the next thirteen positions expressed the workers. The strategy of the strings was that the job that was expressed by the integer in the first position was assigned to the worker that was expressed by the integer in the eighth position, and the job that was expressed by the integer in the second position was assigned to the worker that was expressed by the integer in the ninth position, and so on. The population size was 30, and the crossover probability was 0.8, and the mutation probability was 0.01, and the generation was 50, and the first population was random produced by computer. By this algorithm, the least cost was 15, and the strategy was (1,1), (3,10), (7,11), (4,5), (2,3), (5,2), (6,9).

The SBSA's answer was better than GA's from the simulation, and the time that SBSA used by computer was shorter than GA's. For this example, SBSA was better than GA.

Summary

This paper proposed SBSA for NBAP, the algorithm's theory is simpler than other algorithms' that deal with NBAP, and the operation and programming by compute is easier than other algorithm. The simulation proved that SBSA is better than GA, and when the cost matrix is not too large, SBSA's calculating speed is similar to GA, even more quickly than GA, and the algorithm also can deal with BAP and other AP.

References

[1] ZhANG Lianpeng. Two Improvements on the Hungarian Assignment Approches[J]. Journal of Xi'an Aerotechnical college, 2007(1):64-66

[2] XIA Yong. Improved Gilmore-Lawler Bound for Quadratic Assignment Problems[J]. Chinese Journal of Engineer Mathmatics, 2007(6):401-413

[3] QIAN Songdi. Operational Research[M]. Tsing Hua University Press,1990:128-134

[4] LIU Yang, FAN Zhiping. A Method for Solving The Multi-objective Assignment Problem with Multiple Format Information, 2008(5):73-78

[5] LI Yan, GUO Qiang. Algorithm for uncertainty assignment problem. Computer Engineering and Applications,2009(15)

[6] ZHOU Kang, TONG Xiaojun, XUN Jin. Algorithm of DNA computing on optimal assignment Problems[J]. Systems Engineering and Electronics, 2007(7):1183-1187

[7] LI Yueqiu, YANG Yaqin, LIU Wenqi. A new approach of optimalassignment proble[J]. Journal of Science of Teachers' Colege and Universit, 2008(3):15-17

[8] CHI Jie. A Not of the Constrained Assignment Problem [J]. CoLLEGE MATHEMATICS, 2009(4):66-69

[9] YU Yingzi, ZHANG Qiang. An Efficient Algorithm for a Class of Generalized Assignment Problem. MATHEMATICS IN PRACTICE AND THEORY, 2008(2):86-91

[10] XU Xiaolin. Study of Multi-type Vehicle Scheduling Problem Based on Hungary Algorithm[J]. Fire Control & Command Control, 2009(2):137-139

doi:10.4028/www.scientific.net/AMM.50-51.391

Hermitian Solutions to a Quaternion Matrix Equation*

Ning Li[1,2, a], Jing Jiang[3, b] and Wenfeng Wang[4, c]

[1] School of Statistics and Mathematics, Shandong University of Finance, Jinan 250014, China

[2] Department of Mathematics, Shanghai University, Shanghai 200444, China

[3] Department of Mathematics, Qilu Normal University, Jinan 250013, China

[4] Department of Automation, Binzhou University, Binzhou 256603, China

[a] lnsdfi@163.com, [b] jing5099@163.com, [c] wfwang801@163.com

Keywords: quaternion matrix equation; Hermitian solution; skew-Hermitian solution; Hermitian generalized inverse.

Abstract. In this paper, we consider Hermitian and skew-Hermitian solutions to a certain matrix equation over quaternion algebra H. Necessary and sufficient conditions are obtained for the quaternion matrix equation $AXA^* + BYB^* = E$ to have Hermitian and skew-Hermitian solutions, and the expressions of such solutions are also given. As an application, the common skew-Hermitian g-inverse of quaternion matrix A and B is considered.

Introduction

In 1976, Khatri and Mitia [1] gave the general Hermitian solution to the matrix equation

$$AXA^* = B \tag{1}$$

over the complex field, necessary and sufficient conditions for (1) to have a Hermitian solution are derived. There have been many papers to discuss the equation (1) (e.g., [1-9]). For instance, Groß [4] investigated it in 1998. Wei and Wang [7] considered the rank-constrained least square Hermitian nonnegative-definite solution of (1). Zhang and Cheng [8] studied rank-constrained Hermitian nonnegative-definite solution of (1) by using *SVDs*. Recently, Liu, Tian and Takane [9] present explicit formulas for the maximal and minimal ranks of Hermitian and skew-Hermitian solutions of (1) over the complex field.

Hermitian and skew-Hermitian solutions have been widely discussed in matrix theory, which are very useful in engineering problems, linear system theory and numerical analysis, and others. Investigating Hermitian and skew-Hermitian solutions to certain matrix equations over quaternion algebra H is a significant part of matrix theory. Motivated by the work mentioned above, and keeping the applications and interests of quaternion matrix in view (e.g., [10-13]), in this paper we consider the generalization of the equation (1), and obtain Hermitian and skew-Hermitian solutions to certain matrix equation over quaternion algebra H. In Section 2, we give necessary and sufficient conditions for the existence and the expressions of Hermitian and skew-Hermitian solutions to the quaternion matrix equation

$$AXA^* + BYB^* = E, \tag{2}$$

where $A, B \in H^{m\times n}$, $E \in H^{m\times m}$ and $X \in H^{n\times n}$, $Y \in H^{n\times n}$ are unknown. As an application, the common skew-Hermitian g-inverse of quaternion matrix A and B is considered.

Throughout this paper, we denote the real number field by R, the complex number field by C, the set of all $m\times n$ matrices over the quaternion algebra

$$H = \{a_0 + a_1 i + a_2 j + a_3 k \mid i^2 = j^2 = k^2 = ijk = -1,\ a_0, a_1, a_2, a_3 \in R\}$$

* This research was supported by the Research Fund Special Contract of Binzhou University (BZXYKJ0802), and the Binzhou University Field Grade Bilingual Curriculum Project (2008 No. 01).

by $H^{m\times n}$, the identity matrix with the appropriate size by I, the transpose of a matrix A by A^T, the column right space of A over H by $R(A)$, and the row left space of A over H by $N(A)$. $A^\dagger$ stands for the Moore-penrose inverse of a matrix A, which is defined to be the unique matrix X satisfying the four matrix equations

$(i)\ AXA = A,\ (ii)\ XAX = X,\ (iii)\ (AX)^* = AX,\ (iv)\ (XA)^* = XA.$

Moreover, R_A and L_A stand for the two projectors $R_A = I - AA^\dagger$, $L_A = I - A^\dagger A$ induced by A.

(Skew-) Hermitian solutions of the quaternion matrix equation (2)

Lemma 1. (see Lemma 2.1 in [11]) Let $A \in H^{m\times n}$, $C \in H^{m\times q}$ be known and $X \in H^{m\times q}$ unknown. Then the matrix equation

$$AX = C$$

has a solution if and only if

$$AA^\dagger C = C.$$

In this case, the general solution is

$$X = A^\dagger C + U - A^\dagger AU,$$

where $U \in H^{n\times q}$ is any matrix.

The next lemma is due to Khatri and Mitia [1], which can be generalized to H.

Lemma 2. Let $A \in H^{m\times n}$, $B \in H^{m\times m}$ and $B^* = B$ be known and $X \in H^{n\times n}$ unknown. Then the matrix equation (1) has a Hermitian solution if and only if

$$AA^\dagger B = B.$$

In that case, the general Hermitian solution of the matrix equation (1) is

$$X = A^\dagger B(A^\dagger)^* + L_A V + V^* L_A,$$

where $V \in H^{n\times n}$ is any matrix.

The following Lemmas are due to Wang [12], which can be generalized to H.

Lemma 3. Let $A \in H^{m\times n}$, $B \in H^{p\times q}$, $C \in H^{m\times r}$, $D \in H^{s\times q}$, $E \in H^{m\times q}$ and $X \in H^{n\times q}, Y \in H^{r\times s}$ are unknown. Let $M = R_A C,\ N = DL_B,\ S = CL_M,\ T = R_D N,\ G = R_S C.$ Then the matrix equation

$$AXB + CYD = E$$

is consistent if and only if

$$MM^\dagger R_A E D^\dagger D = R_A E,\ CC^\dagger E L_B N^\dagger N = E L_B.$$

In that case, the general solution of the matrix equation is of the form

$$X = A^\dagger(E - CYD)B^\dagger + L_A U + ZR_B$$

$$Y = M^\dagger R_A E D^\dagger + L_M(V - S^\dagger SVNN^\dagger) - L_M S^\dagger C L_G W T N^\dagger + (W - G^\dagger GWTT^\dagger)R_D$$

where U, V, W, Z are any matrices over H with appropriate sizes.

Lemma 4. Let $A \in H^{m\times n}$, $C \in H^{n\times n}$ be known, and $X \in H^{n\times m}$ unknown. Then the matrix equation

$$A^*X \pm X^*A = B \tag{3}$$

has a solution if and only if

$$B = \pm B^*,\ F_A B F_A = 0.$$

If the matrix equation (3) has a solution, then the general solution is

$$X = (A^\dagger)^* B - \frac{1}{2}(A^\dagger)^* B\ A^\dagger A - AA^\dagger W + \frac{1}{2}AA^\dagger W A^\dagger A + W \mp \frac{1}{2}(A^\dagger)^* W^* A \tag{4}$$

where $W \in H^{n\times m}$ is any matrix.

Proof. This proof is just like [10], so we omit it here.

Theorem 5. Let $A,\ B \in H^{m\times n}, E \in H^{m\times m}$ and $X \in H^{n\times n},\ Y \in H^{n\times n}$ are unknown. Let

$$M = R_A B, N = R_A C, G = L_M B^*,$$

$T = (BM^{\dagger}N)^{*} - B\ M^{\dagger}N$.

Then system (2) has a Hermitian solution if and only if

$$E = E^{*},\ MM^{\dagger}NBB^{\dagger} = N, \tag{5}$$

$$L_G T L_G = 0. \tag{6}$$

In that case, the Hermitian solutions of system (2) over H can be expressed as the following

$$X = A^{\dagger}(E - BYB^{*})\ (A^{\dagger})^{*} + L_A U + U^{*} L_A$$

$$Y = (M^{\dagger}N + L_M V)\ (B^{*})^{\dagger} + B^{\dagger}(M^{\dagger}N + L_M V)^{*} L_B + L_B P L_B, \tag{7}$$

where $U \in H^{n\times n}$ is any matrix, P is an arbitrary Hermitian matrix over H, and

$$V = (G^{\dagger})^{*} T - \frac{1}{2}(G^{\dagger})^{*} T\ G^{\dagger}G - GG^{\dagger}W + \frac{1}{2}GG^{\dagger}WG^{\dagger}G + W - \frac{1}{2}(G^{\dagger})^{*} W^{*} G, \tag{8}$$

where $W \in H^{n\times n}$ is any matrix.

Proof. If the matrix equation (2) has Hermitian solutions $X_0 = X_0^{*}$, $Y_0 = Y_0^{*}$, then

$$AX_0A^{*} + BY_0B^{*} = E\ .$$

Obviously, $E = E^{*}$. It follows from Lemma 3 that (5) holds and

$$AX_0A^{*} = E - BY_0B^{*} \tag{9}$$

has a Hermitian solution X_0. By Lemma 2, (9) has a Hermitian solution, yields

$$MY_0B^{*} = N. \tag{10}$$

Since

$$F_M F_M^{\dagger}(Y_0B^{*} - M^{\dagger}N) = F_M Y_0B^{*} = Y_0B^{*} - M^{\dagger}MY_0B^{*} = Y_0B^{*} - M^{\dagger}N,$$

by Lemma 1,

$$F_M V = Y_0B^{*} - M^{\dagger}N$$

has a solution. There exists V_0 such that

$$F_M V_0 = Y_0B^{*} - M^{\dagger}N$$

i.e.,

$$BF_M V_0 = BY_0B^{*} - B\ M^{\dagger}N.$$

Hence,

$$G^{*}V_0 - V_0G^{*} = T \tag{11}$$

has a solution V_0, yields $L_G T L_G = 0$ by Lemma 4.

Conversely, suppose that (5) and (6) hold, it can be verified that the matrix X,Y of the form (7) are solutions of (2). Now we show that X, Y in (7) are Hermitian matrices. From $E = E^{*}$, we only need to show Y is a Hermitian matrix.

Let

$$S = (M^{\dagger}N + L_M V)\ (B^{*})^{\dagger} + B^{\dagger}(M^{\dagger}N + L_M V)^{*} L_B,$$

i.e.,

$$S^{*} = B^{\dagger}\ (M^{\dagger}N + L_M V)^{*} + L_B\ (M^{\dagger}N + L_M V)\ (B^{*})^{\dagger}.$$

Then S is a Hermitian matrix if and only if

$$B^{\dagger}BF_M\ V(B^{*})^{\dagger} - (B^{\dagger}BF_M V(B^{*})^{\dagger})^{*} = (B^{\dagger}BM^{\dagger}N(B^{*})^{\dagger})^{*} - B^{\dagger}BM^{\dagger}N(B^{*})^{\dagger}. \tag{12}$$

By (6) and Lemma 4, matrix equation (12) has a solution and the general solution is of the form (8). For such solution V of (8), we obtain

$$B^{\dagger}(G^{*}V - VG^{*})(B^{*})^{\dagger} = B^{\dagger}T(B^{*})^{\dagger} = (B^{\dagger}BM^{\dagger}N(B^{*})^{\dagger})^{*} - B^{\dagger}BM^{\dagger}N(B^{*})^{\dagger}, \tag{13}$$

which is equal to (12), yields $S = S^{*}$.

From $Y = S + L_B P L_B$, $S = S^{*}$,

Y is a Hermitian matrix. Obviously X is a Hermitian matrix.

Finally, let X_0, Y_0 be arbitrary Hermitian solutions of (2), then there exists V_0 of the form (11), such that $F_M V_0 = Y_0 B^* - M^\dagger N$.

Let $U = X_0$, $V = V_0$, $P = Y_0$ in (7), we have that X_0, Y_0 can be expressed as the following

$$X_0 = A^\dagger (E - BY_0 B^*)\,(A^\dagger)^* + L_A X_0 + X_0 L_A,$$

$$Y_0 = (M^\dagger N + L_M V_0)\,(B^*)^\dagger + B^\dagger (M^\dagger N + L_M V_0)^* L_B + L_B Y_0 L_B.$$

When $E = -E^*$, we can get some similar result, the proof is omitted.

Theorem 6. Let $A,\ B \in H^{m\times n}, E \in H^{m\times m}$ and $X \in H^{n\times n},\ Y \in H^{n\times n}$ are unknown. Let

$$M = R_A B, N = R_A C, G = L_M B^*,$$

$$T = (BM^\dagger N)^* + B\,M^\dagger N.$$

Then system (2) has a skew-Hermitian solution if and only if

$$E = -E^*,\ MM^\dagger NBB^\dagger = N,$$

$$L_G T L_G = 0.$$

In that case, the skew-Hermitian solutions of system (2) over H can be expressed as the following

$$X = A^\dagger (E - BYB^*)\,(A^\dagger)^* + L_A U - U^* L_A$$

$$Y = (M^\dagger N + L_M V)\,(B^*)^\dagger - B^\dagger (M^\dagger N + L_M V)^* L_B + L_B P L_B,$$

where $U \in H^{n\times n}$ is any matrix, P is an arbitrary skew-Hermitian matrix over H, and

$$V = (G^\dagger)^* T - \frac{1}{2}(G^\dagger)^* T\ G^\dagger G - GG^\dagger W + \frac{1}{2} GG^\dagger W G^\dagger G + W + \frac{1}{2}(G^\dagger)^* W^* G,$$

where $W \in H^{n\times n}$ is any matrix.

Let $A = A_1 + A_2 i + A_3 j + A_4 k$ be a square matrix over H. A matrix X is called a (skew-) Hermitian g-inverse of A, denoted by $A_h^-\,(A_{sh}^-)$ if it satisfies $AXA = A$ and $X = X^*\,(X = -X^*)$.

Corollary 7. Let $A,\ B \in H^{m\times n}$ be given. Let $E = BL_A,\ G = R_E B,\ T = B^\dagger E,\ F = BA^\dagger B^* - D.$ Then A, B have a common skew-Hermitian g-inverse, if and only if

$$A = -A^*,\ B = -B^*,\ BB^\dagger FEE^\dagger + EE^\dagger FBB^\dagger - EE^\dagger FEE^\dagger - F = 0.$$

In that case, the common skew-Hermitian g-inverse of A, B can be expressed as

$$X = A^\dagger + F_A V_0^* - V_0 F_A + F_A Z^* - ZF_A + B^\dagger BZE^\dagger E - E^\dagger EZ^*\ B^\dagger B$$

$$+\frac{1}{2}(T^* ZT^* - TZ^* T) + \frac{1}{2}\,(E^\dagger EZ^* T\ T^\dagger - TT^\dagger ZE^\dagger E), \tag{14}$$

where $Z \in H^{n\times n}$ is arbitrary, and V_0 is defined by

$$V_0 = \frac{1}{2} B^\dagger F(E^\dagger)^* - \frac{1}{2}\,B^\dagger EE^\dagger F(E^\dagger BG^\dagger)^* + \frac{1}{2} G^\dagger F(E^\dagger)^*.$$

Summary

In this paper, we have derived necessary and sufficient conditions for the existence and the expression for the general (skew-) Hermitian solutions to system (2) over H. As an application, we have presented the common skew-Hermitian g-inverse of quaternion matrix A and B. Some known results can be viewed as the special cases of this paper. Moreover, investigating the maximal and minimal ranks of the Hermitian solutions to system (2) over H is also an interesting and important further research topic. We will solve this problem in another paper.

References

[1] C.G. Khatri, S.K. Mitra: SIAM J. Appl. Math. Vol. 31 (1976), p. 579.

[2] J.K. Baksalary: Linear and Multi-linear Algebra, Vol. 16 (1984), p. 133.
[3] H. Dai: Linear Algebra Appl. Vol. 131 (1990), p. 1.

[4] J. Groß: Bull. Malaysian Math. Sci. Soc. Vol. 21 (1998), p. 57.

[5] H. Dai, P. Lancaster: Linear Algebra Appl. Vol. 246 (1996), p. 31.

[6] J. Groß: Linear Algebra Appl. Vol. 321 (2000), p. 123.

[7] M. Wei, Q. Wang: Int. J. Comput. Math. Vol. 84 (2007), p. 945.

[8] X. Zhang, M. Cheng: Linear Algebra Appl. Vol. 370 (2003), p. 163.

[9] Y. Liu, Y. Tian, Y. Takane: Linear Algebra Appl. Vol. 431 (2009), p. 2359.

[10] S.L. Adler: *Quaternionic Quantum Mechanics and Quantum Fields* (Oxford University Press, UK 1995).

[11] Q.W. Wang: Comput. Math. Appl. Vol. 49 (2005), p. 665.

[12] Q.W. Wang: Linear Algebra Appl. Vol. 384 (2004), p. 43.

[13] Q.W. Wang, C.K. Li: Linear Algebra Appl. Vol. 430 (2009), p. 1626.

doi:10.4028/www.scientific.net/AMM.50-51.396

Crack Pattern and Ductility of Connections Composed of CSSEUHSC Columns and SEC Beams under Cycle Loads

YAN Changwang[1, 2,a], ZHANG Ju [1,b] and JIA Jinqing [2,c]

[1] Inner Mongolia University of Technology, Huhhot 010051, China

[2] Dalian University of Technology, Dalian 116024, China

[a] ycw20031013@126.com, [b] ycw20031013@hotmail.com, [c] ycw970741@126.com

Keywords: seismic damage; crack pattern; CSSEUHSC column; cycle loads

Abstract. In order to investigate the crack pattern and ductility of connection composed of cross shaped steel encased ultra high strength concrete (CSSEUHSC) columns and steel encased concrete (SEC) beams under cycle loads, six interior connection specimens were tested in the laboratory. It was found that cracks appeared in the connection core regions and at the beam end for all specimens, and all the connections behaved in a ductile manner and failed in bending with a beam plastic hinge and shearing in the connection core region. The experimental results indicated that test parameters of connection composed of CSSEUHSC columns and SEC beams with good crack resistance performance may be referred for engineering application.

Introduction

After cracks appeared in beam-column connection region under earthquake load, columns and beams adjacent to the connection can not continue to work effectively. Especially for the structures in the corrosive environment, crack resistance capacity is required highly. So it was necessary to study the behavior and calculation of crack resistance in the connection core region. Steel encased ultra high strength concrete (100MPa) structure is more and more widely used in the high-rise and long-span buildings [1]. But many researches [2, 3] on crack resistance of beam-column connection are focused on the traditional concrete ranged 30 MPa to 70 MPa. So the writers will study crack pattern of connection composed of CSSEUHSC columns and SEC beams under cycle load based on performed experiments [4,5] in this paper.

Experimental program

Description of the specimens. A total of six specimens were designed to represent approximately a half-scale model of prototype connection composed of CSSEUHSC columns and SEC beams used in high-rise buildings. The square column of six specimens had the same cross-section of 220 × 220 mm and height of 1500 mm, and the rectangular beams of six specimens had the same cross-section of 180 × 250 mm and length of 2220 mm. The specimens consisted of the shaped steel, longitudinal reinforcement, hoop reinforcement and concrete. The configuration of specimen is shown in Fig.1.

Material properties. The ultra high strength concrete (UHSC) mix was designed for the connection region and column. The mix proportions for 1 m^3 UHSC consisted of 144 kg water, 420 kg Type PII 52.5R cement, 60 kg silica fume, 120 kg Class-I fly ash, 1155 kg coarse aggregate, and 495 kg sand. The water-to- cementitious materials ratio was 24%. Type Sika-3301E superplasticizer was also used to improve workability and setting time. The cube compressive strength values of concrete used in the connection core region and column are list in Table 1. The cube compressive values used in the beam were approximately 35 MPa.

Steel coupons cut from the specimens were tested to determine their yield strength and ultimate strength of the shaped steel. Three coupons were taken from the flange and web plates of the section. The strengths of the longitudinal reinforcement and hoop reinforcement were also tested. The average test values are summarized in Table 2.

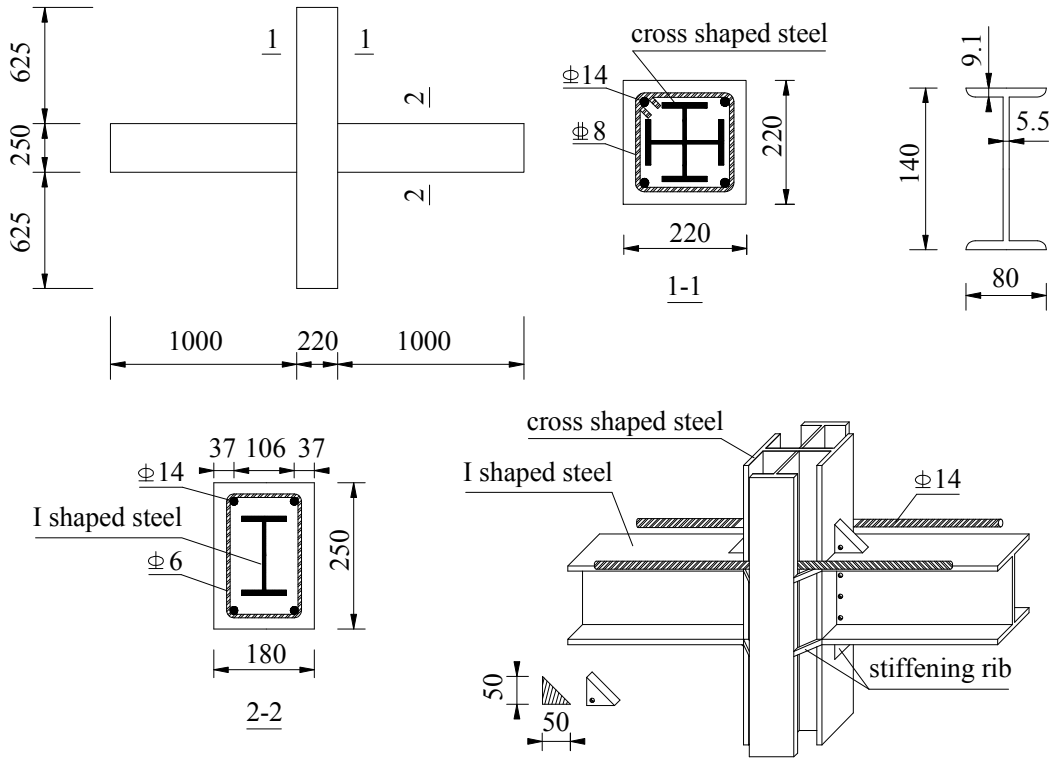


Fig.1 Connection configuration (unit: mm)

Table 1 Summary of specimen information

Specimen designation	Cube compressive strength used in the connection core region [MPa]	Stirrup spacing [mm]	Volumetric stirrup ratio [%]	Applied axial load ratio	μ
SSRC-II-40-1.0	96.3	119	1.0	0.40	3.04
SSRC-II-35-1.6	102	75	1.6	0.35	3.65
SSRC-II-40-1.6	105.4	75	1.6	0.40	3.50
SSRC-II-45-1.6	98.7	75	1.6	0.45	3.23
SSRC-II-40-2.2	103.8	54	2.2	0.40	3.56
SSRC-II-45-2.2	103.8	54	2.2	0.45	3.42

Table 2 Measured mechanical properties of structural steel and reinforcement

Material	Yield strength [MPa]	Ultimate strength [MPa]
Φ6	382	551
Φ8	470	661
Φ14	357	469
Φ16	391	571
I_{14}	254	368

Loading history. The lateral loading history of the connection was generally based on the JGJ101-96[6] for cyclic testing. The axial load was applied up to the design level on the top of the column and was kept constant during the test. The cyclic lateral loads were imposed vertically at the ends of the beams using force-control before the specimen yielded, then using displacement-control

after the specimen yielded. Cycles were then taken to lateral yielding displacement levels of Δ_y, $2\Delta_y$, $3\Delta_y$, $4\Delta_y$, et al. Loading was repeated only once at each control point before the specimen yielded and repeated cubic at each control point to obtain the degraded curve of restoring force after the specimen yielded. The adopted loading history is shown in Fig. 2.

Test setup. The test setup for simulating the loading state of connection is illustrated in Fig.3.

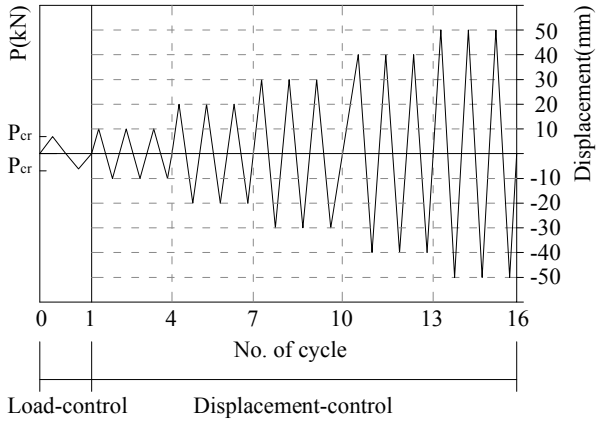


Fig.2 Typical loading history

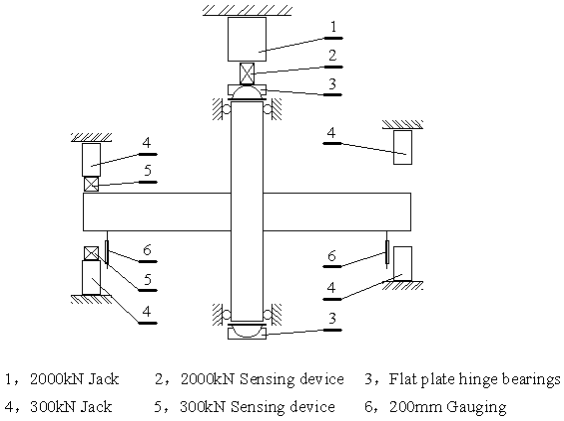


Fig.3 Test setup and measurement

Ductility

The displacement ductility is required to ensure ductile response and evaluate the deformation capacity. A displacement ductility coefficient (μ) is defined as the ratio of the ultimate displacement (Δ_u) corresponding to 15% drop in the strength from the peak value to the yielding displacement (Δ_y), namely $\mu=\Delta_u/\Delta_y$. The ultimate displacement and yielding displacement of connections were determined according to "energy method" shown in Fig.4. Table 1 shows the displacement ductility coefficients of all specimens.

It was shown that displacement ductility coefficient increased with increasing volumetric stirrup ratio and decreased with increasing applied axial load.

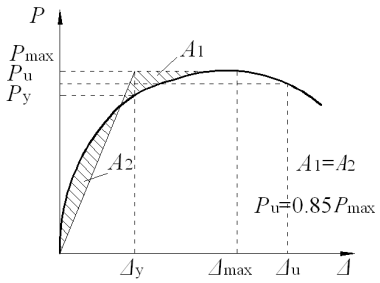


Fig.4 Definitions of yield and failure points

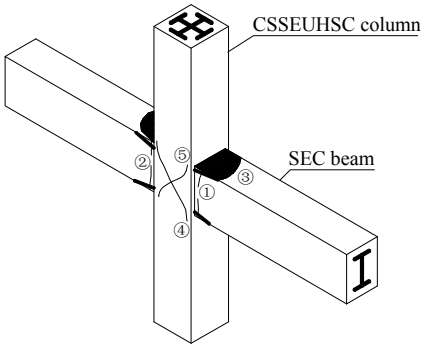


Fig.5 Main cracks progression

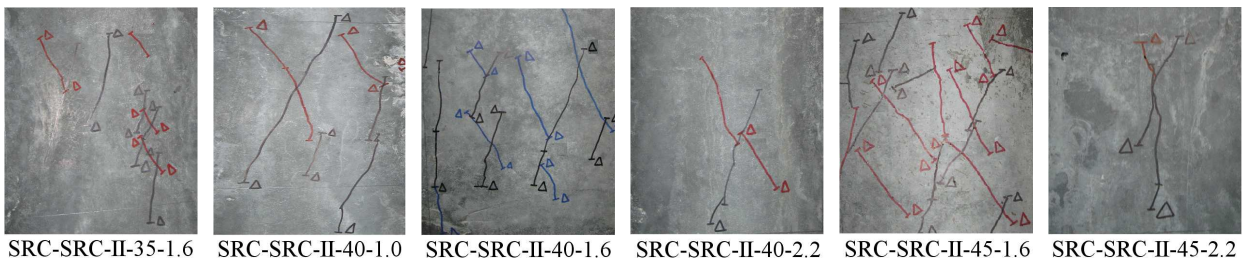


Fig.6 Crack patterns of specimens at a drift angle of 1%

Crack pattern

Cracking pattern can provide useful information regarding the failure mechanism of each specimen, so the cracking pattern was carefully marked and recorded during the experiments. For all specimens, the main progression of cracks is demonstrated in Fig.5. The primary crack ① and ② formed in the portion of the beams adjacent to the beam-column interface and at the beam end, while the cyclic lateral load reached at approximately 13% of peak load. With an increase of reversal cycle load, the concrete ③ was crushed for most specimens during the 3% and 4% drift angle stages. At approximately 64% of peak load, some cracks in the connection core region started and spread in diagonal direction, and crack ④and crack ⑤ formed during the 2% and 3% drift angle stages.

The crack patterns of specimen at a drift angle of 1% are shown in Fig.6. During the 3% and 4% drift angle stages, the cyclic lateral load became maximum (namely peak load, approximately 95 kN), and no cracks started and the width of cracks relatively stopped increasing. Major cracks were concentrated in the connection core region and and at the beam end. during the 5% drift angle cycles, the cyclic lateral load decreased to 85%～90% of its peak load.

Conclusions

This paper had described a series of new tests for connection composed of CSSEUHSC columns and SEC beams. A discussion on the ductility and crack pattern had presented. From the results of this paper, the following conclusions may be drawn:

(1) The displacement ductility coefficient of connection composed of CSSEUHSC columns and SEC beams ranged from 3.04 to 3.65, while displacement ductility coefficients of the concrete connection and steel reinforced concrete connection are 0.2 and 0.4 approximately.

(2) All the connections behaved in a ductile manner and failed in bending with a beam plastic hinge and shearing in the connection core region as the traditional reinforced concrete connection subjected to reversal cycle load.

Acknowledgements

The research reported in the paper is a part of the Project 50878037 supported by National Natural Science Foundation of China, Project LP1015 supported by Open Foundation of State Key Laboratory of Coastal and Offshore Engineering and Project 20100471445 supported by China Postdoctoral Science Foundation. Their financial supports are highly appreciated.

References

[1] J. Hegger. High-strength Concrete for a 186 m High Office Building in Frankfurt, Germany. Engineering Structures, Vol.18 (1996). p.850

[2] J. R. Tang, X. H. Chen. Behavior and Calculation of Shear Capacity of Steel Reinforced Concrete Beam-column Joint. Journal of Building Structures, Vol.11(1990). p.28

[3] H. X. Zhang, L. G. Wang, W. N. Wang. Crack Resistance Behavior of the Joint of HSRC Column and HRC Beam under Low Reversed Cyclic Horizontal Loading. Journal of Shenyang Architecture and Civil Engineering University, Vol.19(2003). p.178

[4] C. W. Yan. Study on Seismic Performance of Steel Reinforced Ultra High Strength Concrete Frame Connection. PHD Thesis. Dalian University of Technology, China (2009)

[5] C. W. Yan, J. Q. Jia, J Zhang. Crack pattern and ductility of steel reinforced ultra high strength concrete composite joint subjected to reversal cycle load. Key Engineering Materials, Vol.417-418(2010). p.845

[6] JGJ101-96. Specification for test methods of seismic buildings. Beijing: Architecture Industrial Press of China (1997)

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.400

Kinematic/Dynamic Simulation of Gear System of Tunnel Boring Machine

Lida Zhu[1,2], Haifeng Zhao[2], Mingliang Zhang[3], and Wanshan Wang[1]

[1]School of Mechanical Engineering and Automation, Northeastern University, 110004, China

[2]Northern Heavy Industries Group Co., Ltd,Shenyang, 110027,China

[3]Shenyang Yuanda Aluminum Industry Engineering Co., Ltd,Shenyang, 110027,China

zld1999@gmail.com

Keywords: Tunnel boring machine; Kinematics; Dynamics; Gear System; Simulation

Abstract. Gear system of tunnel boring machine is simulated and analyzed in kinematic/dynamic based on virtual prototype to ensure stability in motion process. The three-dimensional model is built by using SolidWorks software and gear system of virtual prototype is simulated by using ADAMS. The angular velocity of gears and force curves between gears and shaft are got by simulation to analyze kinematically and dynamically, so the method can provide reference of dynamic optimization design.

Introduction

Tunnel boring machine is of necessary for underground construction and water conservancy and electricity. The design theory is not enough perfect and manufacturing technology is not reasonable in tunnel boring machine; therefore the key component need be researched. Gear transmission is a very important component in tunnel boring machine. Deformation and vibration will be emerged by support load and dynamic transmission process in Gear system of tunnel boring machine, which directly affect motion stability. Therefore, it is necessary of simulation analysis in kinematic and dynamic for gear system to forecast and improve before physical prototype manufacturing to ensure design feasibility and save development time [1-3].

The gear virtual prototype model is built using co-simulation method in SolidWorks and ADAMS software. Firstly, three-dimensional model is created by SolidWorks software, and then transformed ADAMS to virtual prototype model after adjusting model. Finally, the kinematic/dynamic simulation is realized to get angle velocity and force curves, which provides reference for motion stability [5-6].

Virtual prototype modeling of gear system

There are many CAD/CAM softwares applying industry at present, such as SolidWorks, Pro/E, UG, CATIA. SolidWorks has many advantages, which makes the best of graphical user interface in Microsoft Windows and provides convenience in design.Table.1 shows some basic parameters of gear, such as gear number, modulus, pressure angles and width of gear.

Involutes parameters equation is as following:

$$\begin{cases} x = r_b(\cos\alpha + \alpha\sin\alpha) \\ y = r_b(\sin\alpha - \alpha\cos\alpha) \end{cases} \quad (1)$$

Where, x, y is respective coordinate value in involutes. r_b is basic radius. α is angle of involute in basic circle.

According to index and parameter gear in design, three-dimensional modeling of gear and shaft are created by using extend and removal Command in SolidWorks, and figure.1 shows assemble of gear.

Tab.1 Geometry parameters of gear

	Gear number	Modular	Width of gear	Pressure angle
Small gear	25	4	22	20
Big gear	42	4	22	20

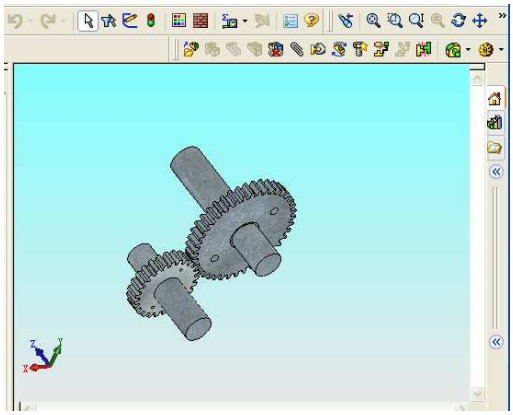
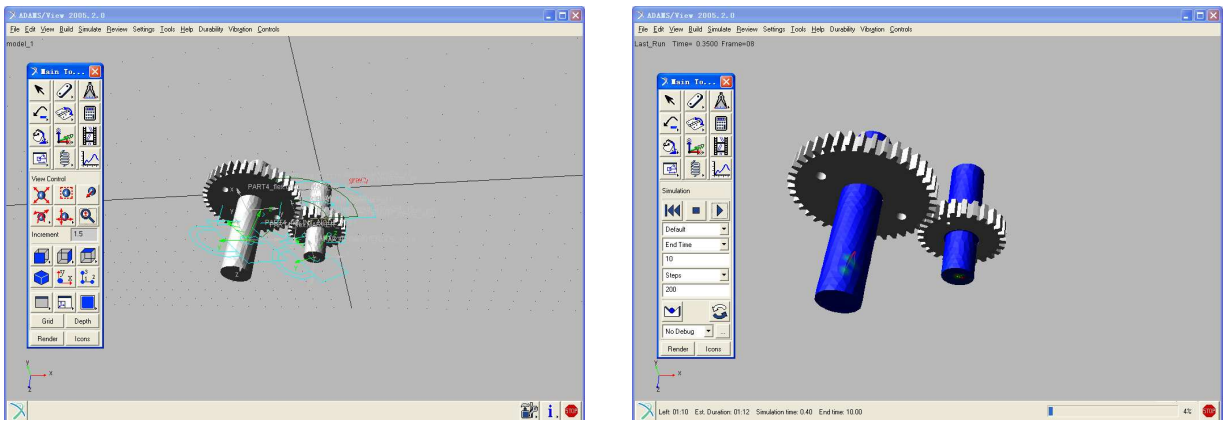


Fig.1 Assemble model of gear

The virtual prototype model of gear is realized in ADAMS software. Firstly, three-dimensional model is created by SolidWorks software, and then transformed ADAMS to virtual prototype model. There are three graphics interchange formats that are respectively STEP, IGES and Parasolid. Parasolid format can avoid data lost during exchange of file, which is very important for validity of simulation results[7]. Fig.2 (a) shows three-dimensional model of gear into ADAMS/View using Parasolid format.



(a) (b)

Fig.2 Virtual prototype mode of gear system

The steps of modeling of gear system:

1. Some parts need be renamed, such as small gear, big gear and shaft.

2. Material property is added to part. For example, material of gear is defined steel. In order to simulate real environment, aggravation will be set in ADAMS.

3. Constraint relation is added between parts, such as shifting pair, revolute joint ,hinges and so on.

4. When you create or modify a gear, you specify or change the revolute, or cylindrical joints located the marker defining the point of contact between the geared parts.

5. Self-checking results are very important in ADAMS/View before adjusting relative among gears and shafts. Constraint relation and the degree of freedom are validity by self-checking results.

At present, the broadly applied software is ADAMS in virtual prototype. Three dimensions FEM of

spindle is divided into thousands of elements with proper element type, and set its attributes and corresponding parameters. For connecting rigid bodies with flexible bodies, the exterior nodes of spindle are set, which are built on the center of relative gyration of flexible body to other parts in mechanical system, and dealt with by setting rigid region. The simulation is executed in Figure 2(b), which includes the information of the flexible body, such as mass, center of mass, moment of inertia, frequency, mode shapes and participation factors. The motion pairs are set up between exterior nodes and rigid bodies to connect flexible bodies and rigid bodies.

Kinematic /dynamic simulation of gear system

The motion process of gear system virtual prototype is realized based on motion function by user-defined in ADAMS. First step: gear begin acceleration to 25(deg/s) from 0 to 1s. Second step: gear can motion level off in 25 25(deg/s) from 1s to 9s. Last step: gear begins deceleration to 0 from 9s to 10s. The angle velocity function of Small gear is as following:
if(time-1:25d*time,25d,if(time-9:25d,25d,if(time-10:250d-25d*time,0,0)))

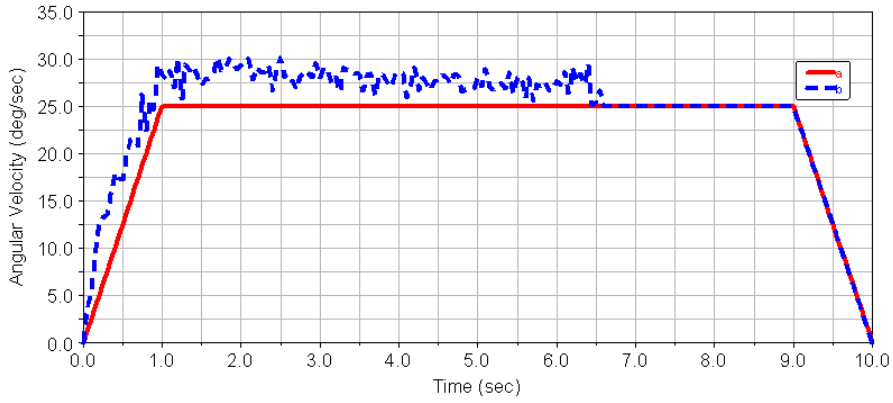


Fig.3 Compare curves of small angle velocity in rigid shaft and flexible shaft

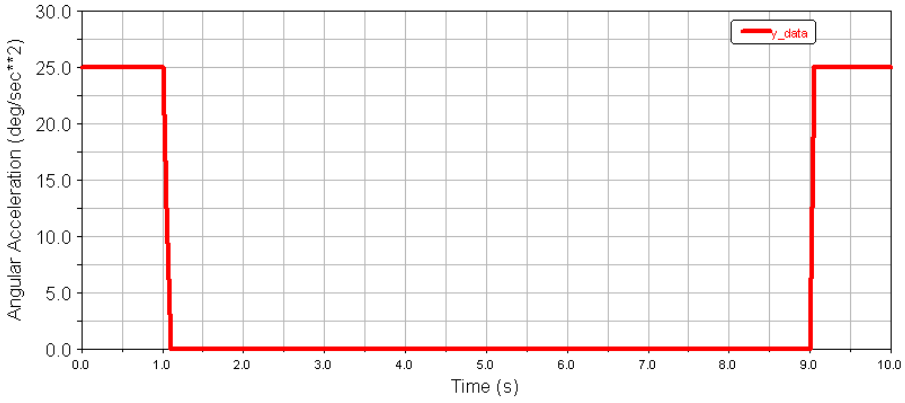


Fig.4 Angle acceleration curves of small gear

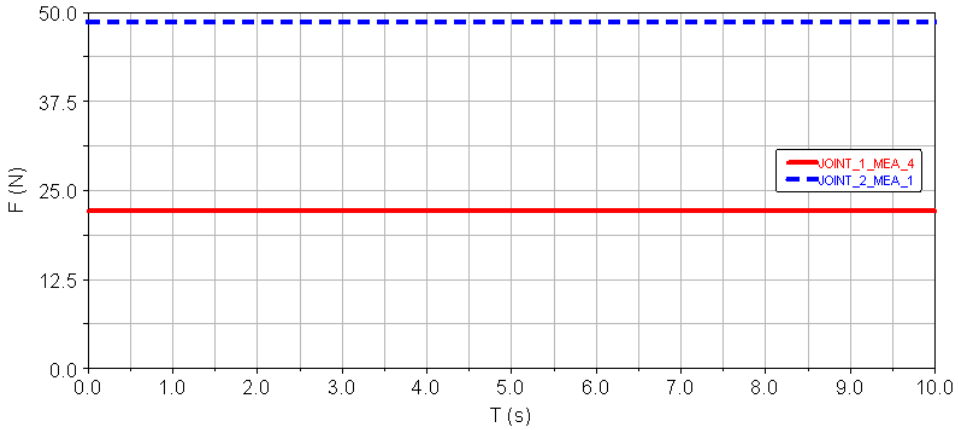


Fig.5 Force curves between gears and shaft

Figure.3 shows Compare curves of small angle velocity in rigid shaft and flexible shaft, and get the same results with mathematics calculation. Figure.5 shows forces between gears and shaft in simulation, which is respectively 24.7N and 49.4N.

Conclusion

Kinematic/dynamic simulation of gear system in tunnel boring machine is very important to reliability design and failure diagnose. Angle velocity and angle acceleration of gears and forces between gears and shafts can get from simulation curves, which can improve mechanical dynamic property and key part and components design. Therefore, it can provide some references of dynamic optimization design.

References

[1] Altintas, Y., Brecher, C., Weck, M., Witt, S., 2005, Virtual machine tool, Annals of CIRP, 54/2: 651-670.
[2] Zaeh, M. Siedl, D,2007, A new method for simulation of machining performance by integrating finite element and multi-body simulation fortunnel boring machines, Annals of the CIRP.383-386. Vol.56/1,
[3] Zhu,C X, Zhu L D, 2008,Modeling of Parallel Robots in Coordination with Flexible Multibody System and Dynamic Simulation, Journal of Northeastern University (Natural Science), Vol.29.No3. 366-370.(In Chinese)
[4] LI San-qun; JIA Chang-zhi; WU Cai-gang; LIU Hai-ping, 2007,Dynamic Simulation Study of Gear Meshing Based on Virtual Prototyping Technology, Journal of System Simulation, Vol.19,No.4:901-904.
[5] Huang Zeping; Ma Jisheng; Wu Dalin; 2002, Simulation Study on Contact Stress of Gear Footh, Journal of Mechanical Transmission, Vol.31, No. 4: 26-28.
[6] LONG Kai;CHENG Ying,2002,The Research of Parameters by The Simulation of Exciting Force in Gears, Computer Simulation,Vol.19,No.6:87-91.
[7] Zheng Y. R. ADAMS Virtual prototype technology introduce and improvement[M],Beijing: Mechanical Industry Press.
[8] Simeon. B, 2006, “On Lagrange multipliers in flexible multi-body dynamics,” Computer methods in applied mechanicsan demerging, vol.195, pp.6993-7005.
[9] Xin Z, Yang R Q, 2002,Theory and application of dynamic modeling of multi-flexible System[J], Mechanical Science and Technology,21(5):387-389.
[10] Zr H, Oertli, T, 2004, Finite Element Modeling of Ball Screw Feed Drive Systems, Annals of the CIRP, 53/1:289-294.

©(2014) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.404

Principal Component Analysis the Economy of Shanghai in 2010

Xiaoqiang Guo [1], Zhendong Li [2],

Dongdong Hao[3], Xin Xie[3] and Jianmin Wang[3]

[1] College of Science,Hebei United University, Tangshan 063009, Hebei Province, China

[2] Basic Education Department, Tangshan College, P.R. China, 063000

[3] Institute of Light Industy, Hebei United University, Tangshan 063009, China

guoxiaoqiang@heut.edu.cn

Keywords: principal component analysis, Shanghai World Expo, index

Abstract. This paper from the economic analysis, quantitative evaluation of the 2010 Shanghai World Exop impact. First, from the short-term and long-term benefits of the two considerations, the loss of earnings, base construction costs on the percentage of total funding, permanent building retained, the number of daily tours, the number of participating countries for the evaluation index, subjectively weight to the five indicators，calculate its scores to rank for five World Expos including Shanghai World Expo. Second, using principal component analysis, we get the five indicators of objective weighting and ranking for above five World Expos. The results show that the Shanghai World Expo will boost the economic development and has a huge influence on the economy.

1. Introduction

World Expo is sponsored by the government of a country, a number of countries or international organizations in order to show the human in the social, economic, cultural and technological achievements in the field of large-scale international exhibition. Its characteristics are held for a long time, large-scale exhibition, participating countries more and far-reaching. From 1851's London "The Great Exhibition", the Expo is becoming an important stage of the exchange of peoples history, culture, exhibition and technological achievements, embodied the spirit of cooperation and look forward to for future development. Shanghai has been successful to applicate World Expo in 2002, indicating that the rise of China's comprehensive strength, and bringing great economic benefits to Shanghai. Choose a side interested in this issue, build a mathematical model, use the internet data, quantitative assessment of the 2010 Shanghai World Expo impact.

This paper consider its influence from economic point of view. World Expo impact on the ecnomy, you can use a number of indicators to measure. The main are: participating countries and international organizations, visitors, base construction costs on the percentage of total funding, whether the balance of payments, whether to retain the building. Participating countries and international organizations reflects the economic development of the host city. The number of visitors reflects the development of tertiary industry. Base construction costs on the percentage of total funding and whether to retain the building reflect the city's long-term economic benefits. Whether the balance of payments reflects the short-term economic benefits. But the World Expo impact on each indicator is different. In this thesis, for the difference, we use principal component analysis to get Expo impact on each index, and then compare the recent five Expo to rank.

2. Model and Solution

In order to better solve the problem, we assume:

I) Dose not take into account the different monetary values in different years;

II) It is not considered separately the number of tours on the impact of holidays;

III) Search data can be a very good response to the actual situation.

In the following discussion, we use a_{ij}, $\overline{a_j}$, σ_j, R, λ, p, Z, M, N_i, f_i to denote evaluation of i-Expo countries j-index value of indexes, the average of the variable j, the standard deviation of the variable j, correlation matrix, eigenvalue, the number of indicators, loss of earnings States, permanent building to retain, principal component, i-indicator respectively.

2.1 From the short-term and long-term benefits of the economic impact of the Expo

We select comprehensive exposition of nearly four times the data overseas and identify five indicators: loss of earnings, base construction costs on the percentage of total funding, permanent building to retain, the number of daily tours and the number of participating countries. These five indicators are weighted, so its rank, to see its influence.

From the short-term benefits, above five indicators give weight: 0.3, 0.2, 0.1, 0.25, 0.15.

Set if loss $Z=0$; if not $Z=1$. And if retain $M=1$; if not $M=0$.

We need to be standard raw data. Let $a_{ij}=\dfrac{a_{ij}-\overline{a_j}}{\sigma_j}$, $where\ \overline{a_j}=\dfrac{1}{n}\sum_{i=1}^{n}a_{ij}$,

$\sigma_j=\sqrt{\dfrac{1}{n-1}\sum_{i=1}^{n}(a_{ij}-\overline{a_j})^2}$, i=1,2,…,n; j=1,2,…,p. The results shown in Table 1.

Table 1: five World Expo and its datas needed of five indicators

Year	Name	Loss of earning	Base construction costs on the percentage of total	Permanent building to retain	The number of daily tours	The number of countries participating
1970	Japan	0.7303	-0.8151	0.44721	0.77901	-1.24098
1992	Seville	-1.09545	1.27746	0.44721	0.04565	0.04337
2000	Hanover	-1.09545	-0.21723	0.44721	-1.29277	0.09341
2005	Aichi	0.7303	0.77923	-1.78885	-0.66323	-0.40699
2010	Shanghai	0.7303	-1.02436	0.44721	1.13135	1.5112

According to the datas in Table 1 and respective weight vector data, we have the scores and ranking of the World Expos, shown in Table 2.

Table 2: the short-term benefits and ranking of theWorld Expos

World Expo	1970	1992	2000	2005	2010
Score	0.1094	-0.0105	-0.6365	-0.0308	0.5685
Ranking	2	3	5	4	1

From the long-term benefits, above five indicators give weight: 0.1, 0.3, 0.25, 0.2, 0.15. Ranking shown in Table 3.

Table 3: the long-term benefits and ranking of theWorld Expos

World Expo	1970	1992	2000	2005	2010
Score	-0.090	0.4011	-0.3075	-0.3341	0.3305
Ranking	3	1	4	5	2

From Table 2 and Table 3 datas, Shanghai World Expo will have the economic impact of top-ranking, and explain Shanghai World Expo a great influence on the economy.

2.2 Principal component analysis

2.2.1 Model building

Specific steps are as follows:

1) Standardization of data, process is the same as 2.1.

2) Find the correlation matrix of standardized data $R=(r_{ij})_{p\times p}, where\ r_{ij}=\frac{1}{n-1}\sum_{k=1}^{n}a_{ki}a_{kj}.$

3) Requirements related to eigenvalues and eigenvectors of R. Let |R- λ I|=0, obtain the characteristic roots, eigenvalue contrbution rate λ_i / p , cumulative contrbution rate $\sum_{i=1}^{m}\lambda_i / p$, where $p=\sum_{i=1}^{p}\lambda_i$; Let $(R-\lambda_i)\mu=0$， obtain the corresponding eigenvectors.

4) Determine the number of principal components. Principal components usually take the cumulative number of m makes the 80% contribution rate.

5) Write m principal components $N_i=\mu_i a, a=(a_1,a_2,\cdots,a_p),\ \mu_i$ be eigenvector of λ_i.

6) Calculate principal components of the load. That is principal components N_i with the original variables a_i a correlation coefficient.

7) Establishment of a comprehensive evaluation function to rank World Expos.

2.2.2 Solution

According to Table 1, write the correlation matrix:

$$\begin{bmatrix} 1.000 & -0.484 & -0.408 & 0.569 & -0.062 \\ -0.484 & 1.000 & -0.43 & -0.493 & -0.205 \\ -0.408 & -0.436 & 1.000 & 0.371 & 0.228 \\ 0.569 & -0.493 & 0.371 & 1.000 & 0.224 \\ -0.062 & -0.205 & 0.228 & 0.224 & 1.000 \end{bmatrix}$$

We have its eigenvalues, the contrbution rate and the cumulative contrbution rate shown in Table 4.

Table 4: eigenvalues, the contrbution rate and the cumulative contrbution rate

Indicators	Eigenvalue	The contrbution rate (%)	The cumulative contrbution rate (%)
Principal Component 1	2.1668	43.3360	43.3360
Principal Component 2	1.4976	29.9522	73.2882
Principal Component 3	0.8365	16.7295	90.0177
Principal Component 4	0.4991	9.9813	99．999
Principal Component 5	0.0001	0.0012	100.00

As seen from Table 4, the cumulative contrbution rate of the first three characteristics to achieve 90.0177%, indicate that the first three principal components contain the basic information on the five indicators. So take the first three eigenvalues, their corresponding eigenvectors are:

$$（0.4034 \quad -0.5760 \quad 0.3104 \quad 0.5874 \quad 0.2532）^T$$

$$（0.6537 \quad 0.0187 \quad -0.6618 \quad 0.0739 \quad -0.3591）^T$$

$$（0.0896 \quad 0.1968 \quad -0.3948 \quad -0.0443 \quad 0.8918）^T$$

So the first three principal components:

Principal Component 1: $N_1 = 0.4034 f_1 - 0.5760 f_2 + 0.3104 f_3 + 0.5874 f_4 + 0.2532 f_5$

Principal Component 2: $N_2 = 0.6537 f_1 + 0.0187 f_2 - 0.6618 f_3 + 0.0739 f_4 - 0.3591 f_5$

Principal Component 3: $N_3 = 0.0896 f_1 + 0.1968 f_2 - 0.3948 f_3 - 0.0443 f_4 + 0.8918 f_5$

The first three principal components obtained eigenvalue contribution rate of weight coefficient, weight to the three principal components, calculate a composite score, that is

$$N = 0.4334 N_1 + 0.2995 N_2 + 0.1673 N_3 .$$

According to the scores to rank, as shown in Table 5.

Table 5: comprehensive benefits and ranking of the World Expos

World Expo	1970	1992	2000	2005	2010
Score	0.4176	-0.7312	-0.7686	0.1090	0.9732
Ranking	2	4	5	3	1

As seen from Table 5, 2010 Shanghai World Expo ranked first, so the greatest influence on the economy.

3. Conclusion

Expo has impact on many aspects of the economy and different levels. We are taking a weighted means analysis, and all three different methods reflect the Shanghai World Expo will have great influence on the economy. Using principal component analysis avoid the subjective judgments in the weight, so that the results fit well with the fact.

References

[1] S.Zh. Cai: *Principle and Methods of Mathematical Modeling*, Beijing, Marine Press（2000）.

[2] D.Y. Zhu: *Featured Case of Mathematical Modeling*, Beijing, Science Press(2003).

[3] Zh.F. Yuan and Sh.D. Song: *Multivariate Statistical Analysis*, Beijing, Science Press(2009).

[4] D.J. Xiang: *Practical Multivariate Statistical Analysis*, Wuhan, Geosciences University of China Press(2005).

[5] J.P. Li and B.Ch. X: *Methods and Application of Multivariate Statistical Analysis*, Beijing, Renmin University of China Press(2008).

[6] Old Shanghai Science Association: *Went into the Expo: come and meet the century's great opportunity*, Shanghai Science and Technology Literature Press(2007).

[7] Shanghai Statistical Yearbook, http://www.stats-sh.gov.cn/2004shtj/tjnj/tjnj2010.htm, 2010.9 .

© (2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.409

Game Analysis of Preventing Conspiracy in Construction Project

Lichao Feng, Chunfeng Liu, Shaohong Yan, Aimin Yang and Jinpeng Wang

College of Science, Hebei Polytechnic University, NO.46 Xinhua West Street, Tangshan 063009, Hebei Province, China

flch19820520@yahoo.cn

Keywords: construction project; conspiracy; game theory; equilibrium probability

Abstract. With rapid growth of national economy, infrastructure construction maintains its vigorous development, but there are some serious problems in this field, for example, construction unit and supervision unit may conspire together to swindle governmental capital. This paper introduces some background which exists in construction project management system, analyses relation and benefit of project's tripartite main body, and establishes conspiracy game's utility matrix of construction unit and supervision unit by introducing game theory. By establishing conspiracy game model between construction unit and supervision unit, the paper mainly discusses conspiracy action which may occur between construction unit and supervision unit (rent-seeking and rent-creating), gets equilibrium probabilities of seeking rent and creating rent, and make some further analysis.

1 Introduction

Project risk management, as a very important aspect of management research, is development and application of risk management in the field of project management. In recent years, with the prosperous development of project construction, project risk management has become the focus problem in the field of management research at home and abroad. To deal with global financial crisis, China put forward a series of polices, some of which are focused on infrastructure construction. However, the phenomena of conspiracies between construction units and supervision units emerge in endlessly, resulting in a lot of jerry-built projects[1,2].

In the process of governmental construction project, government, supervision unit and construction unit take important roles and form a certain degree of economic partnership[3-5]. The relation between government and supervision unit is the relation of entrusting and being entrusted, and the relation between supervision unit and construction unit is the relation of supervising and being supervised; construction unit may conspire with supervision unit to purse higher benefit, so supervision unit entrusted by government supervise the performance of construction unit strictly. At the same time, to pursue additional benefit, supervision unit has the possibility to conspire with construction unit, which influence the performance of project seriously [6-7]. This paper establishes a rational incentive mechanism to constrain supervision unit and construction unit by introducing game theory.

2 Relation Analysis and Benefit Analysis of Project's Tripartite Main Body

By signing a contract, government and construction unit form principal-agent relationship. According to the contract, as party A of project, government presents the project to construction unit to perform; construction unit, as party B of project, is responsible for the performance including cost, completion period and quality of the project. The relation of government and construction unit is made by supervision contract, and is realized by performing the contract, which is the relation of entrusting and being entrusted. As soon as supervising contract is signed, the relation of entrusting and entrusted is made, supervising contract must strive to protect interest of owner, and government would give some supports to supervision unit to assist supervision unit and pay the cost of supervision according to contract. The relation between supervision unit and construction unit is relation of supervising and

being supervised; construction unit would perform project according to contract, and the whole process of performance would be supervised by supervision unit.

As party A of project, the foremost concern of government are the cost, completion period and quality of the project, and its aim is to use least cost to complete the project with high quality and short completion period. As party B of project, the aim of construction unit is to complete the project and get the fund of government within the stipulated time. In the process of performance, construction unit has the possibility to pursue the benefit of maximization by jerry-build. Supervision unit, as third party of project, is to gain the benefit by supervising the completion of project. At the same time, supervision unit has the possibility to pursue its own benefit of maximization by conspiring with construction unit.

3 Game Model and Equilibrium Analysis of Conspiracy between Construction Unit and Supervision Unit

3.1 Game Model of Conspiracy between Construction Unit and Supervision Unit

In Chinese imperfect project management system, construction unit and supervision unit have the possibility to reach a natural motive to conspire. It is because that the work of construction unit must be reflected through the supervision units eventually, and supervision unit has the objective need to share fund to maximize its own interests, construction unit and supervision unit are highly interdependent in gaining additional private benefits.

In fact, whether construction unit and supervision unit choose to conspire with each other or not is the course of mutual gaming. Game model is as follows:

When construction unit transmits the information of conspiracy earlier to supervision unit to seek rent, if supervision unit accept and create rent then conspiracy comes into being, and supervision unit does not need to supervise construction unit resulting in loss of government benefit; if supervision unit doesn't accept, supervision unit would perform stricter supervision to construction unit. In the same way, conspiracy may come into being when supervision unit create rent to construction unit earlier to transmit the intention of conspiracy, and construction unit is faced with two alternatives: accept or not accept. When accepting, conspiracy is generated; otherwise, the supervision of construction unit would be strengthened.

Make some assumptions:

1) Construction unit and supervision unit are rational, and risk neutral;

2) law-abiding construction unit would not seek rent to supervision unit, law-abiding supervision unit would not create rent to construction unit;

3) For the sake of convenience, let all utility functions $u(x)$ s of supervision unit and construction unit be x ($u(x) = x$);

Let fixed net benefit of construction unit completing project be S_1 (S_1>0); let trust fund of supervision unit completing supervision be S_2 (S_2>0, not including governmental additional reward), let cost of supervision be C_J ($C_J < S_2$), then net benefit of supervision unit is $U_J = S_2 - C_J$; so when conspiracy between construction unit and supervision unit doesn't succeed, construction unit and supervision unit perform to pursue their benefits of maximization, and the utilities of construction unit and supervision unit are $U_S = S_1$, $U_J = S_2 - C_J$ respectively.

Assume that the probability of construction unit seeking rent is P_1, that the probability of supervision unit creating rent is P_2, and that capital of government is R which would be carved up by construction unit and supervision unit when conspiracy between them is achieved. The curve-up ratio is L:(L-1), namely, construction unit would get LR and supervision unit would get (1-L)R. In fact, conspiracy behavior between construction unit and supervision unit may be audited out (the probability is α, $0 < \alpha < 1$), and construction unit and supervision unit would be punished V_S, V_J

respectively ($V_S > LR, V_J > (1-L)R$). When construction unit seeks rent but law-abiding supervision unit doesn't accept, supervision unit would strengthen supervision force, resulting in benefit lost S_1' of construction unit and additional supervising cost C_J' of supervision unit. When supervision unit creates rent but law-abiding construction unit doesn't accept, supervision unit would also strengthen supervision force, resulting in benefit lost S_1'' of construction unit and additional supervising cost C_J'' of supervision unit. So we could get conspiracy game's utility matrix as Table 1

Table 1 conspiracy game's utility matrix of construction unit and supervision unit

Construction unit / Supervision unit	Rent-seeking P_1	No rent-seeking $1-P_1$
Rent-seeking P_2	$S_1 + LR - \alpha V_S$ $S_2 + (L-1)R - \alpha V_J$	$S_1 - S_1''$ $S_2 - C_J - C_J''$
No Rent-seeking $1-P_2$	$S_1 - S_1'$ $S_2 - C_J - C_J'$	S_1 $S_2 - C_J$

3.2 Equilibrium Analysis of Conspiracy between Construction Unit and Supervision Unit

Let expected utility of construction unit be μ_1 and expected utility of supervision unit be μ_2, so we could get that

$$\mu_1 = P_1\left[P_2(S_1 + LR - \alpha V_S) + (1-P_2)(S_1 - S_1')\right] + (1-P_1)\left[P_2(S_1 - S_1'') + (1-P_2)S_1\right] \quad \text{.....(1)}$$

$$\mu_2 = P_2\left\{P_1\left[S_2 + (L-1)R - \alpha V_J\right] + (1-P_1)\left[S_2 - C_J - C_J''\right]\right\} + (1-P_2)[P_1(S_2 - C_J - C_J') + (1-P_1)(S_2 - C_J)] \quad \text{.....(2)}$$

Do the first order to formulas (1) and (2)

$$\frac{\partial \mu_1}{\partial P_2} = P_1(S_1 + LR - \alpha V_S) - P_1(S_1 - S_1') + (1-P_1)(S_1 - S_1'') - (1-P_1)S_1 = 0$$

We could get

$$P_1^* = \frac{S_1''}{LR - \alpha V_S + S_1' + S_1''} \quad \text{.....(3)}$$

$$\frac{\partial \mu_2}{\partial P_1} = P_2[(S_2 + (L-1)R - \alpha V_J) - (S_2 - C_J - C_J'')] + (1-P_2)[(S_2 - C_J - C_J') - (S_2 - C_J)] = 0$$

We could get

$$P_2^* = \frac{C_J'}{(L-1)R - \alpha V_J + C_J + C_J' - C_J''} \quad \text{.....(4)}$$

Take formula (3) into formula (1) and take formula (4) into formula (2), we could obtain maximal expected utilities of construction unit and supervision unit μ_1^*, μ_2^*.

$$\mu_1^* = \frac{S_1''}{LR - \alpha V_S + S_1' + S_1''}\left[\frac{C_J'}{(L-1)R - \alpha V_J + C_J + C_J' - C_J''}(S_1 + LR - \alpha V_S)\right.$$

$$+(1-\frac{C_J'}{(L-1)R-\alpha V_J+C_J+C_J'-C_J''})\ (S_1-S_1')\ ]+(1-\frac{S_1''}{LR-\alpha V_S+S_1'+S_1''})[$$

$$\frac{C_J'}{(L-1)R-\alpha V_J+C_J+C_J'-C_J''}\ (S_1\text{-}S_1'')+(1-\frac{C_J'}{(L-1)R-\alpha V_J+C_J+C_J'-C_J''})S_1\ ]\text{.....(5)}$$

$$\mu_2^*=\frac{C_J'}{(L-1)R-\alpha V_J+C_J+C_J'-C_J''}\{\frac{S_1''}{LR-\alpha V_S+S_1'+S_1''}[\ S_2+(\mathrm{L}-1)\,\mathrm{R}-\alpha\ V_J\ ]+(1$$

$$-\frac{S_1''}{LR-\alpha V_S+S_1'+S_1''})\ (S_2-C_J-C_J'')\}+(1-\frac{C_J'}{(L-1)R-\alpha V_J+C_J+C_J'-C_J''}$$

$$[\frac{S_1''}{LR-\alpha V_S+S_1'+S_1''}\ (S_2-C_J-C_J')+(1-\frac{S_1''}{LR-\alpha V_S+S_1'+S_1''})\ (S_2-C_J)\ ]\text{.....(6)}$$

From formula (3), we could get that equilibrium probability P_1^* of rent-seeking is related with conspiring benefit LR of construction unit, punishment $\alpha\ V_S$, and loss S_1'. In fact, the probability P_1^* is unrelated with S_1'', because S_1'' is benefit lost of construction unit when construction unit does not accept rent-creating of supervision unit. Namely, construction unit is law-abiding and does not seek rent naturally, so S_1'' means nothing to equilibrium probability P_1^*.

From formula (4), we could get that equilibrium probability P_2^* of rent-creating is related with conspiring benefit (1-L)R of supervision unit, punishment αV_J, and two kind of costs C_J, C_J''. In fact, the probability P_2^* is unrelated with C_J', because C_J' is additional cost of supervision unit when supervision unit does not accept rent-seeking of construction unit. Namely, supervision unit is law-abiding and doesn't create rent naturally, so C_J' means nothing to equilibrium probability P_2^*.

From analysis above, we could obtain that if and only if probability of rent- seeking is equal or greater than equilibrium probability P_1^* ($P_1 \geq P_1^*$) construction unit has the possibility to seek rent, and that if and only if probability of rent-creating is equal or greater than equilibrium probability P_2^* ($P_2 \geq P_2^*$) supervision unit has the possibility to create rent. In a summary, if and only if $P_1 \geq P_1^*$, $P_2 \geq P_2^*$, conspiracy between supervision unit and construction unit occurs successfully.

4 Conclusion

This paper introduces game theory into the analysis process, forecasting process and decision-making process of construction project, to prevent occurrrence of conspiracy. The paper establishes game model of Project's Tripartite Main Body, discusses optimal strategy under equilibrium statue, and gets the conclusion that controlling the auditing probability and punishment would prevent conspiracy from occurring efficiently, which has the certain economical practical significance. Of course, in order that government encourage supervision unit to supervise construction unit by adjusting reward and decreases the probability of conspiracy between supervision unit and construction unit, the paper needs to design governmental incentive contract furtherly, and calculate optimal incentive reward which maximize governmental expected utility.

References

[1] D.P. Fang, M.G. Li, and L.Y. Shen: Risks in Chinese Construction Market Contractors' Perspective. American Society of Civil Engineers Vol. 130- 6(2010), p. 853-861
[2] H. Chris and A. Tung: Project management for construction: fundamental concepts for owners, engineers, architects, and builders. Pittsburgh (1998)
[3] D.O. Garold: Project management for engineering and construction. New York: McGraw-Hill (1993)
[4] R.H. Wan and S.Z. Zhao: DESIGN AND IMPLEMENTATION OF PROJECT CONTRACT MANAGEMENT INFORMATION SYSTEM BASED ON UML. Computer Applications and Software Vol.8 (2006), p. 45-47
[5] Y.H. Yang and Y.L. Wang: A Game Analysis of the Collusion of Players Involved in Large Construction Project. Journal of Industrial Engineering Management Vol.20-2 (2006),p.126-128
[6] H.J. Zhang, X.Z. Yu and X.C. He: Analysis on the tripartite game of owner and supervision cahoots program. SHANXI ARCHITECTURE Vol.37-7(2008) p.23～24
[7] J. Xu: Three-side Dynamic Game Analysis on Macro-Controlling in the Domain of Real Estate. Journal of UESTC(Social Sciences Edition) Vo1.9-6(2007) p.18～21
[8] Y. Sun, D. P. Fang and S. Q. Wang: Safety risk identification and assessment for Beijing Olympics venues construction. Journal of management in engineering Vol.24-1(2008), p.40-47
[9] D. Fudenberg and J.Triole: Game theory. The MLT press (1991)
[10] W.Y. Zhang: Game Theory and Information Economics. Shanghai People's Publishing House (1996)

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.414

Incentive Contract Designing of Preventing Conspiracy in Construction Project

Lichao Feng[1], Zhiyou Liu[2], Xinghuo Wan [1], Jingguo Qu [3]and Yamian Peng[1]

[1] College of Science, Hebei Polytechnic University, NO.46 Xinhua West Street, Tangshan 063009, Hebei Province, China

[2] Affiliated Hospital, North China Coal Medical College, Tangshan 063000, Hebei Province, China
[3] College of Light Industry, Hebei Polytechnic University, Tangshan 063000, Hebei Province, China

flch19820520@yahoo.cn

Keywords: construction project; conspiracy; game theory; incentive contract

Abstract. With rapid growth of national economy, infrastructure construction maintains vigorous development, but there are some serious problems in this field such as the conspiracy between construction unit and supervision unit. This paper introduces some background of construction project management system, analyses relation and benefit of project's tripartite main body, and establishes game model between government and supervision unit. Finally, the paper designs governmental incentive contract, and calculates optimal incentive reward which maximize governmental expected utility, which could encourage supervision unit to supervise construction unit by adjusting reward, and decrease probability of conspiracy between supervision unit and construction unit.

1 Introduction

Project risk management, as a very important aspect of management research, is development and application of risk management in the field of project management. In recent years, with the prosperous development of project construction, project risk management has become the focus problem in the field of management research at home and abroad. To deal with global financial crisis, China put forward a series of polices, some of which are focused on infrastructure construction. However, the phenomena of conspiracies between construction units and supervision units emerge in endlessly, resulting in a lot of jerry-built projects[1,2].

In the process of governmental construction project, government, supervision unit and construction unit take important roles and form a certain degree of economic partnership[3-5]. The relation between government and supervision unit is the relation of entrusting and being entrusted, and the relation between supervision unit and construction unit is the relation of supervising and being supervised; construction unit may conspire with supervision unit to purse higher benefit, so supervision unit entrusted by government supervise the performance of construction unit strictly. At the same time, to pursue additional benefit, supervision unit has the possibility to conspire with construction unit, which influence the performance of project seriously [6-7]. This paper establishes a rational incentive mechanism to constrain supervision unit and construction unit by introducing game theory.

2 Relation Analysis and Benefit Analysis of Project's Tripartite Main Body

By signing a contract, government and construction unit form principal-agent relationship. According to the contract, as party A of project, government presents the project to construction unit to perform; construction unit, as party B of project, is responsible for the performance including cost, completion period and quality of the project. The relation of government and construction unit is made by supervision contract, and is realized by performing the contract, which is the relation of entrusting and being entrusted. As soon as supervising contract is signed, the relation of entrusting and entrusted is

made, supervising contract must strive to protect interest of owner, and government would give some supports to supervision unit to assist supervision unit and pay the cost of supervision according to contract. The relation between supervision unit and construction unit is relation of supervising and being supervised; construction unit would perform project according to contract, and the whole process of performance would be supervised by supervision unit.

As party A of project, the foremost concern of government are the cost, completion period and quality of the project, and its aim is to use least cost to complete the project with high quality and short completion period. As party B of project, the aim of construction unit is to complete the project and get the fund of government within the stipulated time. In the process of performance, construction unit has the possibility to pursue the benefit of maximization by jerry-build. Supervision unit, as third party of project, is to gain the benefit by supervising the completion of project. At the same time, supervision unit has the possibility to pursue its own benefit of maximization by conspiring with construction unit.

3 Game Model and Inventive Contract between Government and Supervision Unit

For preventing conspiracy occurrence, government must need to fully consider game strategy between self and supervision unit, and need to design incentive contract, assuring that supervision unit could bear project, and encourage supervision unit to complete project to maximize expected utility.

3.1 Game Model between Government and Supervision Unit

Make some assumptions:

1) Let expected benefit of government's investment be 0, and its aim is to pursue expected utility of maximization ;

2) Let supervision units be risk neutral; let supervision unit's utility function be $u(x)$;

3) Let government be risk averse, for the sake of convenience, let governmental utility function $u_1(x)$ be x ($u_1(x) = x$);

4) Assume that the probability of construction unit seeking rent is P_1, that the probability of supervision unit creating rent is P_2, and that capital of government is R which would be carved up by construction unit and supervision unit when conspiracy between them is achieved. The curve-up ratio is L:(L-1), namely, construction unit would get LR and supervision unit would get (1-L)R. In fact, conspiracy behavior between construction unit and supervision unit may be audited out (the probability is α , $0 < \alpha < 1$), and construction unit and supervision unit would be punished V_S, V_J respectively ($V_S > LR, V_J > (1-L)R$). let trust fund of supervision unit completing supervision be S_2 (S_2 >0, not including governmental additional reward).

Game process is designed as follows:

1) Government entrusts supervision unit with supervision of construction unit;

2) Government signs an incentive contact with supervision unit;

3) Supervision unit could choose his own risk status, namely, create rent or not create rent. The choice of supervision unit is unknown to government, but it is known that the probabilities of creating rent and non-creating rent are P_2 and 1- P_2 respectively. From analysis above, illegal benefit of supervision unit is (1-L)R under the condition of rent-creating, illegal benefit of supervision unit is 0 under the condition of non-rent-creating.

4) The probability of conspiring successfully is P_1P_2 , the probability of conspiring unsuccessfully is 1- P_1P_2 . If conspiracy succeeds, government would lose R , and government would give supervision unit S_2 as entrusted reward.

5) The probability of government auditing supervision unit's risk statue successfully is α , and cost of auditing is D. If the result of auditing is that supervision unit conspires unsuccessfully, government

would think conspiring unsuccessfully attribute to supervision unit's non-creating rent, so supervision unit would get $S_2^{'} = S_2 + B$, where B is the incentive reward for non-creating rent; but if the result of auditing is just reverse, government would retrieve illegal profit and supervision unit would be punished V_J.

From analysis above, we could get game matrix between government and supervision unit as Table 2:

Table 2 game matrix between government and supervision unit

government / supervision unit	Auditing successfully α	Auditing unsuccessfully 1-α
Conspiring successfully P_1P_2	$R-D$ $u(S_2-V_J)$	$-R-D-B$ $u[S_2+B+(L-1)R]$
Conspiring unsuccessfully $1-P_1P_2$	$R-D-B$ $u(S_2+B)$	$R-D$ $u(S_2-V_J-B)$

3.2 Equilibrium Analysis of Game Model

According to game matrix above, we could get expected utility of supervision unit under the condition of conspiring successfully is

$$E_1 = P_1P_2\{\alpha u(S_2-V_J)+(1-\alpha)u[S_2+B+(L-1)R]\};$$

expected utility of supervision unit under the condition of conspiring unsuccessfully is

$$E_2 = (1-P_1P_2)[\alpha u(S_2+B)+(1-\alpha)u(S_2-V_J-B)].$$

When expected utilities of supervision unit under the condition of conspiring successfully and unsuccessfully are indifferent, namely, $E_1 = E_2$

$$P_1P_2\{\alpha u(S_2-V_J)+(1-\alpha)u[S_2+B+(L-1)R]\} = (1-P_1P_2)[\alpha u(S_2+B)+(1-\alpha)u(S_2-V_J-B)] \quad ....(1)$$

We obtain equilibrium probability α that

$$\alpha^* = \frac{(1-P_1P_2)u(S_2-V_J-B)-P_1P_2u[S_2+B+(L-1)R]}{P_1P_2\{u(S_2-V_J)-u[S_2+B+(L-1)R]\}-(1-P_1P_2)[u(S_2+B)-u(S_2-V_J-B)]} \quad .....(2)$$

If the probability of governmental auditing successfully $\alpha > \alpha^*$, then optimal choice of supervision unit is conspiring unsuccessfully; conversely, optimal choice is conspiring successfully. To prevent the occurrence of conspiracy, government must strengthen supervising force to improve supervising quality.

3.3 Designing Incentive Contract

In fact, the incentive reward B needs to be designed rationally by government; otherwise, it would influence the effect of preventing occurrence of conspiracy. So government would establish an incentive contract to maximize its expected utility. For the sake of convenience,let $u(x) = x$. We could establish the model of governmental expected utility $Eu(B)$ of maximization (incentive reward B is unknown):

$$\max_B Eu(B) = \alpha[P_1P_2(R-D)+(1-P_1P_2)(R-D-B)] + (1-\alpha)[P_1P_2(-R-D-B)+(1-P_1P_2)(R-D)]$$

$$\text{s.t.}\quad P_1P_2\{\alpha(S_2-V_J)+(1-\alpha)[S_2+B+(L-1)R]\}+(1-P_1P_2)[\alpha(S_2+B)+(1-\alpha)(S_2-V_J-B)] \geq \mu_2^* \quad .....(3)$$

Where $Eu(B)$ is governmental expected utility, $Eu(B)$ is at least 0, and constraint condition (3) is incentive constraint which denotes expected benefit of supervision unit under the condition of incentive contract is equal or greater than that under the condition of non-incentive contract μ_2^*. To solve the problem, simplify the constraint condition (3):

$$(2\alpha-1)(1-P_1P_2)B \geq \mu_2^* - [\alpha+(1+P_1P_2)(1-\alpha)]S_2 - [(1-\alpha)(L-1)R-\alpha V_J]P_1P_2 + (1-\alpha)V_J$$

As mentioned above, α is the successful auditing probability when conspiracy between construction unit and supervision unit occurs. However, in fact, at present our construction project management is not enough strict, supervising rules are imperfect, and force of attacking conspiracy is not enough. So the probability α is always smaller than 10%, so

$$B \geq \frac{\mu_2^* - [\alpha+(1+P_1P_2)(1-\alpha)]S_2 - [(1-\alpha)(L-1)R-\alpha V_J]P_1P_2 + (1-\alpha)V_J}{(2\alpha-1)(1-P_1P_2)} \quad \text{.....(4)}$$

calculate governmental expected utility, we could get

$$Eu(B) = -[\alpha(1-P_1P_2)+(1-\alpha)P_1P_2]B + \alpha(R-D) + (1-\alpha)[(1-2P_1P_2)R-D].$$

It is worthy to note that the function is linear function of incentive reward B, and that the function is monotonic decreasing about incentive reward B. So we could obtain governmental maximal expected utility, if and only if B gets its minimum.

$$\min B = \frac{\mu_2^* - [\alpha+(1+P_1P_2)(1-\alpha)]S_2 - [(1-\alpha)(L-1)R-\alpha V_J]P_1P_2 + (1-\alpha)V_J}{(2\alpha-1)(1-P_1P_2)} \quad \text{.....(5)}$$

So we could obtain governmental maximal expected utility as follows:

$$MaxEu(S_2', B) = \alpha(R-D) + (1-\alpha)[(1-2P_1P_2)R-D] - [\alpha(1-P_1P_2)+(1-\alpha)P_1P_2] \times \frac{\mu_2^* - [\alpha+(1+P_1P_2)(1-\alpha)]S_2 - [(1-\alpha)(L-1)R-\alpha V_J]P_1P_2 + (1-\alpha)V_J}{(2\alpha-1)(1-P_1P_2)} \quad \text{.....(6)}$$

4 Conclusion

This paper introduces game theory into the analysis process, forecasting process and decision-making process of construction project, to prevent conspiracy from occurring. The paper introduces some background of construction project management system, analyses relation and benefit of project's tripartite main body, and establishes game model between government and supervision unit. In addition, the paper designs governmental incentive contract, and calculates optimal incentive reward which maximize governmental expected utility. So government encourage supervision unit to supervise construction unit by modifying reward B, decreasing the probability of conspiracy between supervision unit and construction unit, which has the certain economical practical significance.

References

[1] D.P. Fang, M.G. Li, and L.Y. Shen: Risks in Chinese Construction Market Contractors' Perspective. American Society of Civil Engineers Vol. 130- 6(2010), p. 853-861

[2] H. Chris and A. Tung: Project management for construction: fundamental concepts for owners, engineers, architects, and builders. Pittsburgh (1998)

[3] D.O. Garold: Project management for engineering and construction. New York: McGraw-Hill (1993)

[4] R.H. Wan and S.Z. Zhao: DESIGN AND IMPLEMENTATION OF PROJECT CONTRACT MANAGEMENT INFORMATION SYSTEM BASED ON UML. Computer Applications and Software Vol.8 (2006), p. 45-47

[5] Y.H. Yang and Y.L. Wang: A Game Analysis of the Collusion of Players Involved in Large Construction Project. Journal of Industrial Engineering Management Vol.20-2 (2006),p.126-128

[6] J. Xu: Three-side Dynamic Game Analysis on Macro-Controlling in the Domain of Real Estate. Journal of UESTC(Social Sciences Edition) Vo1.9-6(2007) p.18～21

[7] Y. Sun, D. P. Fang and S. Q. Wang: Safety risk identification and assessment for Beijing Olympics venues construction. Journal of management in engineering Vol.24-1(2008), p.40-47

[8] W.Y. Zhang: Game Theory and Information Economics. Shanghai People's Publishing House (1996)

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.418

A Review on Development Course of Formula for Roots of Equation

Xinghua Ma[1, a], Fengtong Yue[2, b], Shujuan Yuan[3, c], and Dongmei Li[3]

[1] College of Science, Hebei Polytechnic University, Tangshan, 063009,China

[2] Hebei University of Science and Technology Tangshan Branch , Tangshan, 063020,China

[3] College of Light Industry, Hebei Polytechnic University, Tangshan, 063000, China

[a]mxh332@163.com, [b]yueft0722@vip.163.com, [c]yuanshujuan1980@163.com

Keywords: Radical formula, Development course, Depressed

Abstract. This essay starting with the origin of thought about equation, discusses the development course of radical formula for equation from simple to higher degree, introduces the derivation method of radical formula for cubic equation and quartic equation, and compares two methods of radical formula for quartic equation(Ferrali method and Descartes method). There is no general formula of root for equation of 5-th order or higher degree. Eventually, this essay expounds the development process of equation's contributions to algebra field.

1 Introduction

With the further development of science, technology and all aspects of society, the research of various subjects were more and more in-depth, more and more inosculated. Many researches involved and delved into quantitative relations and geometric figures. Equation is such a good mathematical tools, which could explain the relationship between variables in realistic problems clearly and solve practical problems excellently.

At one time a great philosopher brought up such a proposal: random question→mathematical problem→algebra problems→solving system of equations→solving algebraic equations with a single variable. Formula for roots of equation is the core to solve many practical problems, so we should realize the importance of solving the equation. Never forget where one's happiness comes from, never forget one's own origin. We memorize predecessor 'great achievements,

which could make us widen sight, enlighten thinking, explore the law of development, and take comprehensive knowledge and profound understanding of mathematics.

1.1 The origin of solving equations in ancient times

Greece was the holy places in Mathematician and philosopher hearts, and equation theory was center topics of classical algebraic. Until the mid-19th century, algebra was still a mathematical subject with equation theory as the center. Solving algebraic equations remained the basic question, especially for solving equations with radicals. So-called equation has real roots, which could be understood that solution could be expressed through finite four operations and extraction to equation coefficients.

Formula Solution of equation had been used to solve quadratic, such as $ax^2 + bx + c = 0$,by Greek mathematician Helen and his former Diophantine since A.D. 2nd century. Until A.D. 7th century Indian mathematician Brahmagupta investigated several different forms of quadratic such as $ax^2 = c$, $ax^2 + bx = c$ and gave roots formula respectively, but he only solved the situation that coefficient was positive.

The Arab has absorbed Greek and India's achievement, and impelled the equation rapid development. In A.D. 820 years, the Arabic mathematicians wrote a great influence book, "Algebra", which introduced the main content about solving equations, and proposed the method that equals to

"moving items " and " merging similar items". Precisely as a result of the Arabic mathematician's original development, the name of algebra evolved from this book.

2 Solving formula of quadratic equation and the simple application

The formula solution of equation had been greatly boosted by the Indian mathematician GhasKara(1114~1185) during the 12th century. First, expressed the formula named SamaGupta clearer; Second, made the quadratic equation unified, and then gave the extract formula for Roots of $ax^2+bx+c=0$;Third, confirmed that quadratic equation has two roots, and proposed the existence of irrational roots.

The methods to solve quadratic equations were completing the square, formula method, factorization method and so on. Completing the square is the basic of formula method. Arbitrarily a quadratic equation can become a general form $x_1=-\frac{b}{2a}+\frac{\sqrt{b^2-4ac}}{2a}$, $x_2=-\frac{b}{2a}-\frac{\sqrt{b^2-4ac}}{2a}$.

Any quadratic equation changed into general form could be directly transformed into formula to solve. We lead factorization method into some special equations. This kind of thought that declines power is a kind of importance thought which handle higher power equations. This thought of reducing the equation order is a kind of importance thought which can handle higher power equations.

3 The formula for roots of cubic equation

Since the 15th century, the European mathematicians had started their independent study. Exploring formula for roots of the cubic equations was an initial research topic. In the 16th century,an Italian legend Taritalia(1500-1557) the earliest claimed he could solve the cubic equations. Another Italian mathematician G.Cardano(1501—1576) wanted studying the method of solving the cubic equations after learning that. When G.Cardano made confidentiality solemn oath, Taritalia gave him the solution manuscript. However, he did not keep his promise. In 1545, G.Cardano published a Latin book about algebra named “ars magna”,which detailed introduced the solution of the cubic equations from G.Cardano.

The formula for roots of cubic equation could not be solved by general deductive thought. We should change the standard type $ax^3+bx^2+cx+d=0$ into the special type $x^3+px+q=0$.

Three roots of $x^3+px+q=0$ are: y_1+z_1 , $\omega y_1+\omega^2 z_1$, $\omega^2 y_1+\omega z_1$.

Here

$$y_1^{\,3}=-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}\ ,\ z_1^{\,3}=-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}\ ,\ y_1z_1=-\frac{p}{3}\ .$$

Looking into extract formula above, we can find that

$$y_1^{\,3}=-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}\ ,\quad z_1^{\,3}=-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}\ ,$$

have the greatly similar form with

$$x_1=-\frac{b}{2a}+\frac{\sqrt{b^2-4ac}}{2a},\quad x_2=-\frac{b}{2a}-\frac{\sqrt{b^2-4ac}}{2a},$$

just as in quadratic equation B^2-4AC which is the root discriminator. Actually, in cubic equation, we can discuss characters of roots from the value of $\frac{q^2}{4}+\frac{p^3}{27}$ whether is positive, zero, or negative.

If $\frac{q^2}{4}+\frac{p^3}{27}>0$, y^3 and z^3 are real numbers, besides $y^3\neq z^3$ .Suppose their real number cubic roots are y_1 and z_1, well then roots are y_1+z_1 ,

$$\omega y_1 + \omega^2 z_1 = -\frac{y_1 + z_1}{2} + \frac{y_1 - z_1}{2}\sqrt{3}i\,,\ \omega^2 y_1 + \omega z_1 = -\frac{y_1 + z_1}{2} - \frac{y_1 - z_1}{2}\sqrt{3}i\,.$$

equation has a real number and two conjugate imaginary numbers.

If $\frac{q^2}{4} + \frac{p^3}{27} = 0$, y^3 and z^3 are real numbers, besides $y^3 = z^3$.

Suppose their real number cubic roots are y_1, well then roots are

$$y_1 + y_1 = 2y_1\,,\ \omega y_1 + \omega^2 y_1 = -y_1\,,\ \omega^2 y_1 + \omega y_1 = -y_1\,.$$

equation has three real numbers , therein two roots are equal.

If $\frac{q^2}{4} + \frac{p^3}{27} < 0$, y^3 and z^3 conjugate imaginary numbers. Suppose they are $r(\cos\theta + i\sin\theta)$ and $r(\cos\theta - i\sin\theta)$, then

$$y_1 = \sqrt[3]{r}\left(\cos\frac{\theta}{3} + i\sin\frac{\theta}{3}\right)$$

and

$$z_1 = \sqrt[3]{r}\left(\cos\frac{\theta}{3} - i\sin\frac{\theta}{3}\right),\quad \text{satisfy } yz = -\frac{p}{3},$$

so the roots are:

$$\begin{aligned}
x_1 &= \sqrt[3]{r}\left(\cos\frac{\theta}{3} + i\sin\frac{\theta}{3}\right) + \sqrt[3]{r}\left(\cos\frac{\theta}{3} - i\sin\frac{\theta}{3}\right)\\
&= 2\sqrt[3]{r}\cos\frac{\theta}{3},\\
x_2 &= \sqrt[3]{r}\left(\cos\frac{\theta}{3} + i\sin\frac{\theta}{3}\right)\omega + \sqrt[3]{r}\left(\cos\frac{\theta}{3} - i\sin\frac{\theta}{3}\right)\omega^2\\
&= -\sqrt[3]{r}\left(\cos\frac{\theta}{3} + \sqrt{3}\sin\frac{\theta}{3}\right),\\
x_3 &= \sqrt[3]{r}\left(\cos\frac{\theta}{3} + i\sin\frac{\theta}{3}\right)\omega^2 + \sqrt[3]{r}\left(\cos\frac{\theta}{3} - i\sin\frac{\theta}{3}\right)\omega\\
&= -\sqrt[3]{r}\left(\cos\frac{\theta}{3} - \sqrt{3}\sin\frac{\theta}{3}\right)。
\end{aligned}$$

In such complexion, equation has different three roots.

4 The derivation of roots formula for quartic equation

L. Ferrari(1522-1565) ,who is the student of Cardano, an Italy mathematician, began to conduct the research to the quartic equation under his teacher's direction.,after the cubic equation be solved. He thought: Taritalia and Cardano worked out the solution of cubic equations which changed the cubic equation into a quadratic equation, and then obtained the result of the equations. Therefore, he associated that he could try to make the quartic equation melt into a cubic equation. Finally he proved that his mentality is completely correct. In the process of composing "ars magna", the roots formula for quartic equation was solved successfully by Ferrari, which be write into "ars magna" by Cardano.

The feature of Ferrari Method is making the left of equation melt into square form, which transforms the original equations into two variant quadratic equation.

Quartic equation

$$x^4 + ax^3 + bx^2 + cx + d = 0 \qquad (1)$$

Transposing, get

$$x^4 + ax^3 = -bx^2 - cx - d$$

coordinate, make equation (1) melt into

$$\left(x^2+\frac{ax}{2}\right)^2=\left(\frac{a^2}{4}-b\right)x^2-cx-d \tag{2}$$

The right of polynomials in equation (2) is not a complete square in the general situation, so we need to add a polynomial with undetermined coefficients t in both sides of equation, which equivalents to add $2\left(x^2+\frac{ax}{2}\right)\frac{t}{2}+\left(\frac{t}{2}\right)^2$, then get

$$\left(x^2+\frac{ax}{2}+\frac{t}{2}\right)^2=\left(\frac{a^2}{4}-b+t\right)x^2+\left(\frac{at}{2}-c\right)x+\frac{t^2}{4}-d \tag{3}$$

equation（3）is a polynomial about x,and the condition of its complete square is that the radical discriminent is zero.

Namely

$$\left(\frac{at}{2}-c\right)^2-4\left(\frac{a^2}{4}-b+t\right)\left(\frac{t^2}{4}-d\right)=0$$

Namely

$$t^3-bt^2+(ac-4d)t-a^2d+4bd-c^2=0 \tag{4}$$

suppose t_0 is one root, that equation (3) turn into

$$\left(x^2+\frac{ax}{2}+\frac{t_0}{2}\right)^2=\left(\sqrt{\frac{a^2}{4}-b+t_0}x+\sqrt{\frac{t_0^2}{4}-d}\right)^2$$

this equation decomposes two equations below:

$$x^2+\left(\frac{a}{2}+\sqrt{\frac{a^2}{4}-b+t_0}\right)x+\left(\frac{t_0}{2}+\sqrt{\frac{t_0^2}{4}-d}\right)=0 \tag{5}$$

$$x^2+\left(\frac{a}{2}-\sqrt{\frac{a^2}{4}-b+t_0}\right)x+\left(\frac{t_0}{2}-\sqrt{\frac{t_0^2}{4}-d}\right)=0 \tag{6}$$

Solve quadratic equation (5) and (6), and get four roots of quartic equation (1).

Thus, Ferrari Method tries to transform a quartic equation into a cubic equation （4）and two quadratic equation（5）,（6）. Theses two questions are also known in mathematics research, which commonly is used the reduction mathematics thought. The general quartic equation might solve, and these expressions of root specific formula was extremely complex, so this paper has not written them.

We have another method to solve common quartic equation, which called Descartes Method.

Descartes Method and Ferrari Method may successfully solve a quartic equation. They both use the same mathematics thought, which falls the high order equation to low order simplification.

5 Solving the quintic and higher degree equation

What can be seen from above is that the roots of quadratic equation, cubic equation or quartic equation all can be represented by algebraic of the equation's coefficients. Moreover, seeking formulas for roots of the general quadratic equation relies on solving a related cubic equation, and seeking formulas for roots of the general cubic equation relies on solving a related quartic equation, which urges the people to ponder whether make the higher order equation induce into the low order equation to solve.

Mathematicians had made efforts to explore the solution of quintic and higher degree equation for more than two hundred without any results. In 1824, a 24-year-old Norway mathematician Niels Henrik Abel (1802～1829) solved the mystery of quintic equation problem successfully. He concluded:" The general quintic and higher degree equation can not be solved with radical expression."

6 The contribution of equation development course to Algebra domain

In 1831, French mathematician Variste Galois(1811-1832) told us which kind of equation could be solved with radical expression. Abel inspired Galois to find the "Galois Basic Theorem" which represents that radical solution is available for Algebraic equations and prove that the general algebra equations with more than 4 orders could not be solved with radical expression. Furthermore, Galois used the "Group" theory and method creativity, opened a new field of mathematics "Group Theory", which over an age of algebra with a center of Equation Theory and turned the research to algebraic structure.

As time goes on, Group Theory penetrates into other branch in mathematics gradually. In the 20th century, Group Theory have occupied an important position in mathematic field and been as one of the foundations in modern mathematic.

References

[1].Howard Eves, OuYangjiang, ZhaoWeijiang translation, Great moments in mathematics, Beijing Science and Technology Press(1990).

[2].Richard Mankiewicz, FengSu, MaJing, FengDingni, The Story of Maths, Hainan Press(2002).

[3].Paul Benacerraf, Hilary Putnam. ZhuShuilin YingZhiyi, ZhangYugang, Philosophy of mathematics, The Commercial Press(2003).

[4].MO Zong jian, LAN Ying zhong, ZHAO Chun lai, The Algebra(I). Peking University Press(1986).

[5].MENG Hu,ZHAO Yun,A discussion on the phase of development of mathematics from the view of epistemology and the partition of mathematical history,Journal of Northwest University(Natural Science Edition)(2003).

[6].GUO Shu chun,The status of Chinese traditional mathematics in the history of mathematics, Studies In College Mathematics(2003).

[7]. QU Yuan hai, Quintic root formula of one variable and Galois group, High School Mathematics Teaching Reference[J], 2005,Vol.9, pp.60-61(2005).

[8]. LIU Xian hai, With a natural formula for finding roots of quadratic equation question the meaning of Results, Mathematical education studies[J], 2007,Vol.5, pp.22-23(2007).

[9].ZENG Chun yan, Comparison of Abel and Galois, Mathematics Communication, 2008.7,Vol.296, pp.42-44(2008).

[10]Gaohua Liu, Shenglin Zhu,On Galois Extension of Hopf Algebras, Chinese Annals of Mathematics [J], 2008, Vol.1(2008).

[11] SHI Yu fa, Accurate calculation of Bingham axial laminar flow drop tube using Quartic formula Roots[J], 2009, Vol.7(2009).

[12] LI Jiang-hua, YAO Zhong-lei, ZHENG Jin-liang, The Factorization of Primes in a Quartic Field, Journal of Hefei University :Natural Sciences [J], 2009, Vol.2(2009).

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.423

Solution to China's GDP Prediction Problem by BP Neural Network

Juan Li[1], Yafeng Ya[2], Haiming Wu[1]

[1] Hebei United University Hebei Tangshan 063009,China

[2] College of Light Industry Hebei United University Hebei Tangshan 063009,China

43698059@qq.com

Keywords: GDP, BP neural network, sigmoid function and learning rate.

Abstract. Because the choice and important of learning rate η, the higher of η and the faster convergence it will be, but it may cause instability or function vibration if η is too high; if η is lower, although it may avoid instability, the speed of function convergence will reduce. In order to solve the contradiction, we introduce a variable of $\Delta W_{ij}(n_0)=\eta\sum_{m=0}^{n_0}\alpha^{n_0-1}\delta_i(m)y_j(m)=-\eta\sum_{m=0}^{n_0}\alpha^{n_0-1}\sum_m\frac{\partial F(m)}{\partial W_{ij}(m)}$, and if the $\sum_m\frac{\partial F(m)}{\partial W_{ij}(m)}$ this time is the same as that of the previous time, the weighted summation value will increase and it results in the regulation speed of right value W at the stable regulation; and if the $\sum_m\frac{\partial F(m)}{\partial W_{ij}(m)}$ this time is contrary to that of the previous time, it indicates that a certain vibration and now the result of summation will make the value of $\Delta W_{ij}(n_0)$ decrease to play a role in stability and increase the speed of function convergence.

1 Brief Introduction of BP Model

BP network was put forward by the scientists group led by Rumelhard and McCelland in 1986 and is a multilayer feed-forward network trained according to error reversion propagation algorithm and is also one of the neural networks most widely applied at present. BP network may learn and store a great amount of input-output model mapping relations and need not to disclose the mathematical equation indicating the mapping relation in advance. Its learning rule is to use the steepest descent method and regulate the network weights and thresholds through counter propagation to minimize the error sum squares of network[1-2].

Topological structure of BP neural network model includes input layer, hidden layer and output layer[3]. It may be proved that the three layer network model with one hidden layer may approximate any nonlinear function and its topological structure is shown as follows:

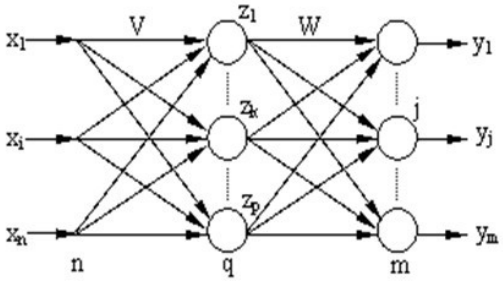


Fig. 1 Topological structure of three layer neural network

Suppose the input vector $X\in R^m$, $X=(x_1,x_2\cdots x_m)$,and among which the hidden layers of $x_1,x_2\cdots x_m$ has E_k neurons and the input vector stands for the input from neurons $1,2,3\cdots m$; and the hidden layer has n neurons and $W=(W_{1m},W_{2m}\cdots W_{nm})$ indicates the connection strength between

the neurons 1,2,3$\cdots m$ and No. n neuron, which is the weight value; and b_n is the threshold value and f is the transfer function and $Y=(y_1,y_2\cdots y_n)$ is the output of No. n neuron[4].

S_n of the net input value of No. n neuron is:

$$S_n=\sum_{m=1}^{m}W_{nm}x_m \tag{1}$$

The application matrix shall be expressed as

$$S_n=\sum_{m=1}^{m}W_{nm}x_m=W_nX \tag{2}$$

The net input S_n, via the transfer function f, will get the output h_n:

$$h_n=f(S_n)=f(\sum_{m=1}^{m}W_{nm}x_m) \tag{3}$$

Then the output of output layer

$$F_n=f(h_n)=f(\sum_{m=1}^{m}W_{nm}h_n) \tag{4}$$

Where: f is a monotone increasing function and must be a bounded function, because that the cell transfer signal can not be increased infinitely and must have a max value.

Therefore, f is a transfer function and in general adopts the sigmoid function: $f(x)=\dfrac{1}{1+e^{-x}}$

Because that we research on GDP which involves in the large data and the change in output value after incorporating into the transfer function is not high, we adopt the function of $y=\dfrac{1}{1+e^{-\frac{x}{10000}}}$ other than the one mentioned above[5].

BP model is a kind of learning algorithm with teacher and belongs to δ learning rate. Suppose that there are Q learning paired samples (X_1,T_1), (X_2,T_2), (X_q,T_q) and (X_Q,T_Q), where $X_q=(x_1^q,x_2^q,\cdots x_n^q)$ is the input information and $T_q=()$ is corresponding training teachers. If we take $X_1,X_2\cdots X_q\cdots X_Q$ as the input of BP, the corresponding output is $Y_1,Y_2,\cdots Y_q\cdots Y_Q$ and $Y_q=(y_1{}^q,y_2{}^q\cdots y_n{}^q)$. The learning algorithm is to modify the weight and threshold connected by it by means of the error between $Y_1,Y_2,\cdots Y_q\cdots Y_Q$ and $T_1,T_2\cdots T_q\cdots T_Q$ to make Y_q approach to T_q as near as possible.

The process of network learning is divided into two steps:

(1) From bottom to top, transfer the input information to the output layer via the action of weight and threshold and then calculate the output values of various layers;

(2) From top to bottom, calculate the errors of various layers and then modify the weight and threshold based on the algorithm result before transferring the error upwards. The two processes will alternate repeatedly until the convergence terminates.

The concrete algorithm steps are shown as follows:

1) Suppose that iteration times n_0 is 1 and the random initialization weight is W_{n_0} and the threshold value is threshold value and learning rate η_{n_0} is a positive initial value lower than 1.

2) Suppose that $n_0=n_0+1$ and then input the data in sample Q into the model and then perform calculation respectively.

Hidden layer output $h_n=f(S_n)=f(\sum_{m=1}^{m}W_{nm}x_m)$ and input layer output $F_n=f(h_n)=f(\sum_{m=1}^{m}W_{nm}h_n)$

3) Calculate the total error $E(n_0) \le \varepsilon$ (specified positive decimal) of the No. n_0 iteration according to the following formula and then terminate the iteration; otherwise, continue to execute the iteration of the fourth step.

4 Calculate the errors of various layers:

Input layer:

$$\delta_{ym} = \sum_{m=1}^{m} (t_m^p - y_m^p) f_2'(S_m) \tag{5}$$

Hidden layer:

$$\delta_{hm} = \sum_{m=1}^{m} (t_m^p - y_m^p) f_2'(S_m) W_{mn} f_1'(S_n) \tag{6}$$

$$W_{ij}(n_0+1) = W_{ij}(n_0) - \eta \frac{\partial F}{\partial W_{ij}(n_0)} = W_{ij}(n_0) + \Delta W_{ij}(n_0) \tag{7}$$

5) Counter propagation method. Because the input layer performs reverse calculation according to "gradient descent method" and regulate the weight value layer by layer. Neuron i cascade right n_0+1 regulation shall be calculated according to the equation below:

$$W_{ij}(n_0+1) = W_{ij}(n_0) - \eta \frac{\partial F}{\partial W_{ij}(n_0)} = W_{ij}(n_0) + \Delta W_{ij}(n_0) \tag{8}$$

6) Return to step 2) to continue to perform iteration.

The corresponding actual values and simulative values and the errors are shown in the table below:

Table1

Actual value	Simulative value	Error	Actual value	Simulative value	Error
0.406258	0.4153	-2.23%	4.819786	4.8831	-1.31%
0.454562	0.4851	-6.72%	6.079373	6.0552	0.40%
0.489156	0.5661	-15.73%	7.117659	7.0972	0.29%
0.532335	0.5995	-12.62%	7.897303	7.8387	0.74%
0.596265	0.6336	-6.26%	8.440228	8.4684	-0.33%
0.720805	0.7022	2.58%	8.967705	9.0439	-0.85%
0.901604	0.8254	8.45%	9.921455	9.8711	0.51%
1.027518	1.0146	1.26%	10.96552	10.9743	-0.08%
1.205862	1.1599	3.81%	12.03327	12.0386	-0.04%
1.504282	1.366	9.19%	13.58228	13.5766	0.04%
1.699232	1.7444	-2.66%	15.98783	15.9955	-0.05%
1.866782	1.9587	-4.92%	18.32174	18.3088	0.07%
2.17815	2.2216	-1.99%	21.19235	19.32	8.84%
2.692348	2.6298	2.32%	24.95299	19.87	20.37%
3.533392	3.4704	1.78%			

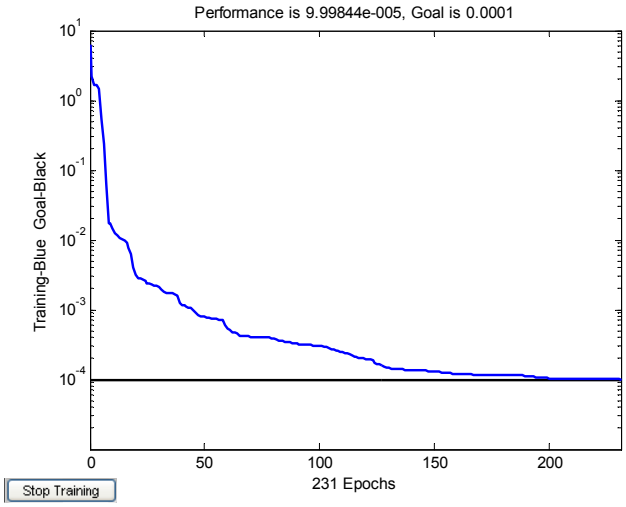


Fig.2

The number of training steps is determined to be 200 according to the diagram of relation between the number training steps and error.

The simulative diagram based on it is shown as follows:

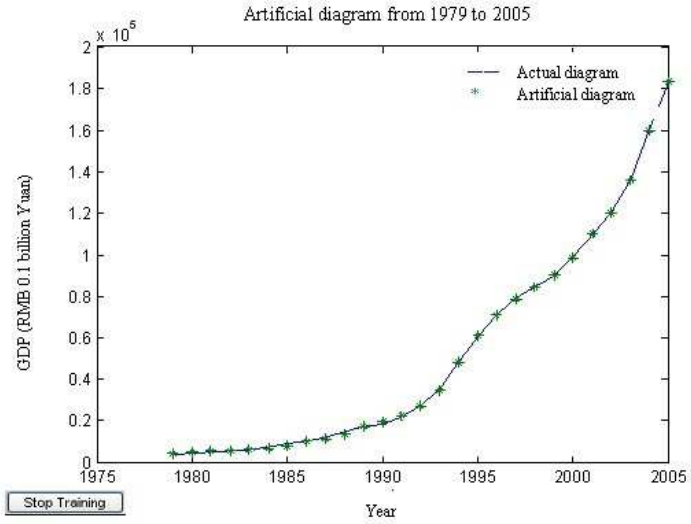


Fig. 3

The corresponding actual values and simulative values and the errors are shown in the table below:

Table2

Actual GDP	Simulative GDP	Error	Actual GDP	Simulative GDP	Error
4062.6	3790	6.71%	48197.9	48140	0.12%
4545.6	4630	-1.86%	60793.7	60730	0.10%
4891.6	5220	-6.71%	71176.6	71200	-0.03%
5323.4	5740	-7.83%	78973.0	78300	0.85%
5962.7	6370	-6.83%	84402.3	84940	-0.64%
7208.1	7120	1.22%	89677.1	90500	-0.92%
9016.0	8340	7.50%	99214.6	98390	0.83%
10275.2	10130	1.41%	109655.2	109620	0.03%
12058.6	11680	3.14%	120332.7	120720	-0.32%
15042.8	13830	8.06%	135822.8	135620	0.15%
16992.3	17560	-3.34%	159878.3	159830	0.03%
18667.8	20180	-8.10%	183217.4	183300	-0.05%
21781.5	21690	0.42%	211923.5	205920	2.83%
26923.5	26720	0.76%	249529.9	236670	5.15%
35333.9	35190	0.41%			

The number of training steps and errors of the latter are far less than that of the former, so the prediction model of the latter is better than that of the former.

References

[1] Cheng Xin, He Yue. Research on application of GMDH-ARCH combined model in prediction of GDP, Statistics and Consulting(2007).

[2] Zhou Xuan. National GDP prediction model based on time-series analysis, Consume Guide·Finance and Accounting Latitude(2009).

[3] Tan Guolan and Li Jian. Application of grey prediction in GDP prediction, Technology & Economy in Areas of Communications(2007).

[4] Ding Webin. Discussion and research on prediction method for total amount of GDP, Statistics and Prediction, Vol, 6: 57-59(2003).

[5] Zhou Zhihua and Cao Cungen. Neural network and it application, Beijing: Tsinghua University Press, 2004.

doi:10.4028/www.scientific.net/AMM.50-51.428

Research of Genetic Algorithm for Parabolic Equation Inverse Problem

Yamian Peng[1, 2, a], Guanchen Zhou[2, a] and Huijuan Zhao[2, a]

[1]College of Sciences, He'Bei United University, Tang'shan, He'bei, china

[2]College of Light Industry, He'Bei United University, Tang'shan, He'bei, china

[a]pym806@163.com

Keywords: GA; parabolic equation; inverse problem; ill-posed.

Abstract. It is difficult to solve the inverse problem because it always ill-posed. This paper introduced a new approach based on Genetic Algorithms (GA) for solve the parabolic equation inverse problem. The GA transforms the inverse problem into an optimization problem. The results of numerical simulation show that the method has high accuracy and quick convergent speed. And it is easy to program and calculate. It is worth of practical application.

1 Introduction

The research of inverse problems is a relatively new area. It had its origins in the sixties the basic thesis. In the history of science there is a well-known "blind listening to drum", which is a Danish physicist, proposed a mathematical problem in 1910. Drum in the shape of known circumstances, to determine the law of drum sound, which is the study of mathematical physics is a mature as early as their own issue. From the experience of people's lives, we can see the "ears on behalf of Head" has some possibilities. For example, in our selection of watermelon, through the voice by beatting, you can know what in it. This is because when the objects to determine the future of the material, its shape and its tone is closely related to the level of experienced people is not difficult to find some link between them. After nearly a century mathematician who explored in depth, so that the people of the reality of similar problems are prevalent in a more thorough understanding.

The most roughly to say, the inverse question is how to get the input from the output. A simple example is that we have a group of input value g by a converter f, after a group of measured output o. The inverse problem is the output o and conversion functions (transfer function) f to get the input g.

Everyone who research is more or less of inverse problem, but many people don't pay attention to this one word. Many people solve the inverse problem solution all the time more broadly. The inverse problem in reality is a purpose, such as physical each branch (involving collision or scattering, radiation, lattice vibrations), geology, sea, air, remote sensing, medical imaging and nondestructive testing of all kinds of graphics and image processing. Therefore, the inverse problem of different names, such as the inverse problem, inverse problems, the scattering problem, projection, etc. In the mathematics method often referring to the scattering of partial differential equation solution to a kind of method, but these equations can become integral equation.

2 Parabolic Equation Inverse Problem

If the partial differential equations of some original conditions into the unknown, and the original condition of the unknown function equation may still be unknown, or just know and the unknown function of some relevant information, we must through the equation, some additional condition and conditions to determine if these unknown quantity, such problems as the inverse problem of partial differential equations. Differential equation is studied how to describe the problem is with the physical process characterization, system status, and society, and a biological phenomenon, such as the differential equation established according to the specific conditions and process to solve certain solution, thus obtained the mathematical description of the process and the state. The ordinary differential equation by "refers to the" because ", namely known or partial

differential equation known unknown part of reverse equation. In recent years, various fields of science development of the inverse problem of these methods often have certain limitations, such as PST (pulse doppler technique) method, Tikhonov regularization, but these methods have shortcomings. The biggest advantage of genetic algorithm (GA) is directly to target function as a search for information, without the gradient of the objective function values and other ancillary information, and use multiple search point search information, has been successfully applied to machine learning and pattern recognition and economic forecasts and optimizing control and various complex data analysis and calculation, etc.

The inverse problem of differential equation by partial differential equations and boundary conditions and the initial conditions and partial differential equations of the three component plus an additional conditions. The general form of written for is as follows:

Differential equation: $Lu(x,t) = f(x,t) \quad x \in \Omega \quad t \in (0,\infty)$

The initial conditions: $Iu(x,t) = \varphi(x) \quad x \in \Omega \quad t = 0$

Boundary conditions: $Bu(x,t) = \phi(x,t) \quad x \in \partial\Omega$

Additional conditions: $Au(x,t) = k(x,t) \quad x \in \partial\Omega$

Now the ordinary differential equation used in many fields, such as resource exploration, atmospheric measurement, economic forecasts, biological organs morphology analysis of disease diagnosis and quantum mechanics and other natural science and engineering field. The inverse problem of significance is hadamard is usually ill-posed. The concept is studied for mathematician's hadamard first partial differential equations. If a partial differential equations and satisfy three conditions as follows:

(1) The problem solution exists (existence);

(2) Its solution at only one (uniqueness);

(3) Its solution depends on continuous data (stability) says this solution is well-posed.

If one does not meet the conditions will say the problem is ill-posed. In a long period of time, people always think from actual down out of the mathematics problems, if it is correct, always well-posed, this concept to people's attention firmly limitations in the circle of well-posed. Until the mid in the 1950s, the interpretation geophysical data cause people to focus on it. People have found wave equation; heat conduction equations and elliptic equations of classical ill-posed problems are not practical significance. This would encourage people to study, well-posed problems from the concept, theory and methods on the development of a new fields. The Soviet Union experts Tikhonov "ill-posed problem" is the first book in this field. Introduction of "conditions Tikhonov well-posed" concept for the appropriate solution (limited conditions, problems to solve well-posed) is the foundation of ill-posed. However, should admit ordinary differential equation, a new mathematical direction, say, is just in the form, in the vast majority of disciplines infancy.

3 Genetic algorithms

Genetic algorithm is a natural biological evolution by natural selection and biological and genetic mechanism and development of problem solving strategy and stochastic computation model. It uses simple encoding technology to say all kinds of complicated structure, and through a group of simple genetic encoding said operation and natural selection to guide the direction of study and determine the search. Because of the evolutionary computation itself has, since learning and adaptive intelligent characteristics such as the essence of parallelism and easy operation, strong commonality etc, has been successfully applied to machine learning, pattern recognition and economic forecasts, optimizing control and various complex data analysis and calculation, but the genetic algorithm for solving the problem of literature.

Genetic Algorithm (GA) as Genetic done, in essence is a kind of specific issues not rely on direct searching method. Genetic algorithm applied in pattern recognition and neural network and image processing and machine learning and industrial optimization control and adaptive control and biological sciences and social sciences etc. Genetic algorithm is based on the basic thoughts of

Darwin's theory of evolution, the genetic and Mendel. Darwin by natural Selection (NAT. Ral theory to explain the species are) the origin and evolution of the theory of natural Selection, including the following three aspects:

(l) Genetic (heredity) is the biological characteristics of "All things are difficult reap", the biological information, to have offspring, in accordance with the information and offspring, differentiation, thus progeny always and have with the same or similar characters. Biological characteristics and species can stability.

(2) Variation between offspring hue and the different between the individual and the progeny total some differences, this kind of phenomenon, called the variation. Variation is random variable selection and accumulation is the root of diversity of life.

(3) The survival of the fittest survives from the surplus of natural selection and the struggle of existence. Because of the struggle of existence constantly, the result is the survival of the fittest, adaptive mutation of individuals are preserved, does not have to be eliminated, the variation of the individual through generations living environment, the choice is sure to variation of species for a direction, and gradually accumulated characters and the ancestors of different, for new species. This kind of natural selection process is a long, slow, continuous process.

4 Solving Process Description

Based on differential equation with several parameters to adjust (except), as long as the initial forecast system input parameters, and the numerical model of computer can output the forecast future, at d_n certain moment in vector axel says some observations on $t = T$ and $h = h(x,t)$, d_n forecast system can realize the mapping as follows:

$$A(q_n(x_i)) = h(x_i, T) = d_n$$

Usually A expression cannot know. Requirements in this condition, namely using parameter inversion of observation, the parameters in d_n within its area K_{ad} allows to find the optimal approximate q_{op}, make model output and the corresponding observations of the minimum:

$$\left\| A(q_{op}) - d_n \right\| = \inf_{q \in k_{ad}} \left\| A(q) - d_n \right\|$$

If $\left\| A(q) - d_n \right\|$ genetic algorithm of fitness functions, the parameter inversion as function optimization problems, according to the latest observation data provided by the evolutionary computation techniques, the model parameters to adjust the minimum, and obtained $\left\| A(q) - d_n \right\|$ to get the optimal parameters estimation. This method can in evolutionary parameters of function space by trigonometric function and polynomial function and spline functions on the undetermined parameter parameterized by genetic algorithm to the evolution of combination coefficients are obtained by the optimal parameters estimation.

Using genetic algorithm is presented to solve the above problems, according to the first question in the region to allow K_{ad} parameter by certain parameters of the operator, o coding strategies for the solution space (operator parameters for the search space). Don't assume operator only contain an unknown parameter $q(x)$ (parameters can also is the source, boundary condition, the initial conditions, etc.), assuming that given a set of functions, the parameters of the representation of the basic functions $\varphi_1(x), \varphi_2(x), \cdots \varphi_n(x)$ expresses as follows:

$$q(x) = \sum_{i=1}^{n} k_i \varphi_i(x)$$

Make $q(x) \in K_{ad}$, the real coded parameters of $q(x)$ can be expressed as the vector$(k_1, k_2, \cdots k_n)$. The question now is transformed into these coefficients of differential equation, the parameter inversion. If these undetermined coefficients of the genetic algorithm, using real number coding

chromosome into $(k_1, k_2, \cdots k_n)$, $\|F(\mu,q)-d_{ob})\|$ expressed as a function of the adaptive genetic algorithm, genetic algorithm can use to search for the solution space for operator parameters.

5 Numerical Examples

The homogeneous parabolic equation parameter inverse problem.

$$\begin{cases} \dfrac{\partial U(x,t)}{\partial t} = \dfrac{\partial^2 U(x,t)}{\partial x^2} - q(x)U(x,t), (x,t) \in \Omega_t \\ U(x,0) = f(x), x \in [0,1] \\ U(0,t) = g_1(t), \dfrac{\partial}{\partial_x} U(1,t) = g_2(t), t \in (0,0.1) \end{cases}$$

Among them: $f(x) = \sin(\pi x^2)$, $g_1(t) = 0$, $g_2(t) = -2\pi e^t$.

Additional conditions for: $U(x_i, 0.02)(i = 1,2,...n)$, $x_i = ih, h = \dfrac{1}{n}, n = 10$. Undetermined parameter is: $q(x) = 2\pi ctg(\pi x^2) - 4\pi^2 x^2 - 1$, True solution for the problem is $U(x,t) = \sin(\pi x^2)e^t$.

Select function groups $\Phi(x) = \{1, x, x^2, ...\}$ approximation $q(x)$ take $q_0(x) = k_1 + k_2 x + k_3 x^2$ as the initial value. As a global variables of k_1, k_2, k_3. Search in the interval $[-20,10]$ with true solution inside the near.

Take nvars number of variables, population size = 3 and prize = 20, the biggest maxgen = 10, evolutionary crossover probability and mutation probability P_C = 0.8 and P_M = 0.2. Undetermined coefficients of unknown quantity calculation results are as follows:

Figure1 The parameter inversion of initial value of search results

Maximum evolutionary algebra	Adaptive value	Search range	$k_{j+1} = (k_0^{j+1}, k_1^{j+1}, k_2^{j+1})$
1000	4.3183×10^{-5}	[-20， 10]	(5.19946277 -19.999999 -19.999999)

6 Summary

We studied using genetic algorithm to solve the parabolic inverse problem in this paper, and simulated of four kinds of Parabolic Equations inverse problems. The results show that we obtained with a small error between the approximation by using genetic algorithms and the real solution. It is feasible in practice and will have an important significance for inverse problem research.

Acknowledgement

Fund item: Hebei United University science study imbrues item(z201019).

References

[1] Victor Isakov. Stefan Kindermann.. *Identification of the diffusion coefficient in one-dimensional parabolic equation*. Inverse Problem. 2000, 16:665-680.

[2] Ting-yan Xiao. Shen-gen Yu. Yan-fei Wang. *Numerical Method for Inverse Problem*. Science publishing company.2003.

[3] Jin ZQ, Zhou ZF. *Inverse problem of engineer hydraulics* [M], Hehai university publisher, 1997

[4] Li W. *Headway of environmental hydraulics* [M], WuHan university publisher, 1999

[5] Min T, Zhou XD. *A Numerical Solution to Initial Condition Inverse Problem of One-dimensional Unstable Furbulent Diffusion*. Journal of Xi’ an University of Technology (2001) Vol.17No.2

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.432

Convergence theorems for a finite family of strictly asymptotically pseudocontractive mappings in q-uniformly smooth Banach spaces

Huancheng Zhang[a] , Aimin Yang, Yamian Peng, Jingguo Qu

Department of Basic Courses, College of Light Industry, Hebei Polytechnic University, Tangshan,063000, P.R. China

[a]zhanghuancheng0@yahoo.com.cn

Keywords: Accretive operator; Resolvent; Composite iterative; Uniformly Gâteaux differentiable; Common zeros

Abstract. Let E be a real q-uniformly smooth and uniformly convex Banach space and K a nonempty closed convex subset of E. Let $T_i : K \to K$, $i = 1, 2, \ldots, N$ be k_i-strictly asymptotically pseudocontractive mappings with $\cap_{i=1}^{N} F\ (T_i) \neq \emptyset$, where $F(T_i) = \{x \in K : T_i x = x\}$. Let $\{x_n\}$ be the sequence generated by

$$x_{n+1} = (1 - \alpha_n)x_n + \alpha_n T_{[n]}^n x_n,$$

where $\{\alpha_n\}$ is a sequence in [0,1] satisfying certain conditions and $T_{[n]}^n = T_i^n, i = n(modN)$. Weak and strong convergence theorems for the iterative approximation of common fixed points of the family $\{T_i\}_{i=1}^N$ are proved.

Introduction

Let E be a real Banach space and $J_q(q > 1)$ denote the generalized duality mapping from E into 2^{E^*} given by $J_q(x) = \{f \in E^* : \langle x, f\rangle = \|x\|^q, \|f\| = \|x\|^{q-1}\}$, where E^* denotes the dual space of E and $\langle ., .\rangle$ denotes the generalized duality pairing. In particular, J_2 is called the normalized duality mapping and it is usually denoted by J. It is well known (see [1]) that $J_q(x) = \|x\|^{q-2}J(x)$ if $x \neq 0$, and that if E^* is strictly convex then J_q is single valued. Let K be a nonempty subset of E. A mapping $T : K \to K$ is called k-strictly asymptotically pseudocontractive if there exists a constant $k \in [0, 1)$ and a sequence $\{k_n\} \subseteq [1, \infty)$ with $k_n \to 1$ such that for all $x, y \in K$ and $j(x - y) \in J(x - y)$,

$$\langle T^n x - T^n y, j(x - y)\rangle \leq \frac{1}{2}(1 + k_n)\|x - y\|^2 - \frac{1}{2}(1 - k)\|(I - T^n)x - (I - T^n)y\|^2$$

for all $n \in N$. T is called uniformly L-Lipschitzian, if there exists $L > 0$ such that $\|T^n x - T^n y\| \leq L\|x - y\|$, for all $x, y \in K$ and $n \in N$, and is said to be demiclosed at a point p if whenever $\{x_n\}$ is a sequence in $D(T)$ such that $\{x_n\}$ converges weakly to $x \in D(T)$ and $\{Tx_n\}$ converges strongly to p, then $Tx = p$.

In [2], Acedo and Xu introduce an explicit iteration called the cyclic algorithm. The cyclic algorithm is the sequence $\{x_n\}$ generated from an arbitrary $x_1 \in E$ as follows:

$x_2 = \alpha_1 x_1 + (1 - \alpha_1)T_1 x_1$
$x_3 = \alpha_2 x_2 + (1 - \alpha_2)T_2 x_2$
$\vdots$
$x_{N+1} = \alpha_N x_N + (1 - \alpha_N)T_N x_N$
$x_{N+2} = \alpha_{N+1} x_{N+1} + (1 - \alpha_{N+1})T_1 x_{N+1}.$

In general, $\{x_n\}$ is generated from arbitrary $x_1 \in E$ by $x_{n+1} = \alpha_n x_n + (1-\alpha_n)T_{[n]}x_n$, where $T_{[n]} = T_i, i = n(modN), 1 \leq i \leq N$.

In 2009, H.Zhang and Y.F.Su [4-5] modified the normal Mann iteration of strictly pseudocontractive map by the cyclic algorithm and proved weak and strong convergence theorems in a q-uniformly smooth Banach space which is also uniformly convex or satisfies Opial's condition.

Recently, M.O.Osilike and Y.Shehu [3,6] extended the cyclic algorithm to asymptotically strictly pseudocontractive mapping and proved strong convergence theorem but in a 2-uniformly smooth Banach space.

In this paper, motivated by [2-6], we extend the cyclic algorithm to asymptotically strictly pseudocontractive mapping and at the same time extend the space to q-uniformly smooth Banach space.

Preliminaries

Let E be a real Banach space. The modulus of smoothness of E is the function $\rho_E : [0,\infty) \to [0,\infty)$ defined by $\rho_E(\tau) = \sup\{\frac{1}{2}(\|x+y\| + \|x-y\|) - 1 : \|x\| \leq 1, \|y\| \leq \tau\}$. E is uniformly smooth if and only if $\lim_{\tau\to 0} \frac{\rho_E(\tau)}{\tau} = 0$. E is called q-uniformly smooth ($q > 1$) if there exists $c > 0$ such that $\rho_E(\tau) \leq c\tau^q$. Hilbert spaces, L_p, l_p spaces ($1 < p < \infty$), and the Sobolev spaces, ${W_m}^p$, ($1 < p < \infty$), are q-uniformly smooth. Hilbert spaces are 2-uniformly smooth while $L_p, l_p, {W_m}^p$ *is* $p-uniformly\ smooth,\ if\ 1 < p \leq 2;\ 2-uniformly\ smooth\ ,\ if\ \ p \geq 2$. It is shown in [7] that there is no Banach space which is q-uniformly smooth with $q > 2$.

We need the following Lemmas for the proof of our main results.

Lemma 2.1 ([9]). *Let E be a real q-uniformly smooth Banach space which is also uniformly convex. Let K be a nonempty closed convex subset of E and $T : K \to K$ be a k-strictly asymptotically pseudocontractive mapping with a nonempty fixed point set. Then $(I - T)$ is demiclosed at zero.*

Lemma 2.2 ([8]). *Let $\{a_n\}, \{b_n\}$ and $\{\delta_n\}$ be sequences of nonnegative real numbers satisfying the inequality*

$$a_{n+1} \leq (1+\delta_n)a_n + b_n, \quad n \geq 1.$$

If $\sum_{n=1}^{\infty} \delta_n < \infty$ and $\sum_{n=1}^{\infty} b_n < \infty$, then $\lim_{n\to\infty} a_n$ exists. If in addition $\{a_n\}$ has a subsequence which converges strongly to zero, then $\lim_{n\to\infty} a_n = 0$.

Main Results

In the following we denote $T_{[n]}$ by T_n.

Lemma 3.1. *Let E be a real q-uniformly smooth and uniformly convex Banach space. Let K be a nonempty convex subset of E and $T_i : K \to K$, $i = 1, 2, \ldots, N$ be a k_i-strictly asymptotically pseudocontractive mapping with a sequence $\{k_n^i\}$ such that $\sum_{n=1}^{\infty}(k_n^i - 1) < \infty$. Let $k = \max\{k_i : 1 \leq i \leq N\}$ and $k_n = \max\{k_n^i : 1 \leq i \leq N\}$ such that $\sum_{n=1}^{\infty}(k_n - 1) < \infty$ and $F = \cap_{i=1}^{N} F(T_i) \neq \emptyset$. Let $\{\alpha_n\}$ be a real sequence in [0,1] satisfying the condition:*

$$0 < a \leq \alpha_n^{q-1} \leq b < \frac{q(1-k)}{2c_q}(1+L)^{-(q-2)}, \quad n \geq 1$$

where a, b are constants in (0,1), and c_q is the constant appearing in inequality (2.1), and L is the maximal Lipschitz constant of $\{T_i\}_{i=1}^N$. Let $\{x_n\}$ be the sequence generated from an arbitrary $x_1 \in K$ by

$$x_{n+1} = (1-\alpha_n)x_n + \alpha_n T_n^n x_n.$$

Then $\forall\, p_1, p_2 \in F$, the limit $\lim_{n\to\infty} \|tx_n + (1-t)p_1 - p_2\|$ exists for all $t \in [0,1]$.

Lemma 3.2. *Let E be a real q-uniformly smooth and uniformly convex Banach space. Let K be a nonempty convex subset of E and $T_i : K \to K$, $i = 1, 2, \ldots, N$ be a k_i-strictly asymptotically pseudocontractive mapping with a sequence $\{k_n^i\}$ such that $\sum_{n=1}^{\infty}(k_n^i - 1) < \infty$. Let $k = \max\{k_i : 1 \le i \le N\}$ and $k_n = \max\{k_n^i : 1 \le i \le N\}$ such that $\sum_{n=1}^{\infty}(k_n - 1) < \infty$ and $F = \cap_{i=1}^N F(T_i) \neq \emptyset$. Let $\{\alpha_n\}, \{x_n\}$ be as in Lemma 3.1. Then $\forall\, p_1, p_2 \in F$, $\lim_{n\to\infty}\langle x_n, J(p_1 - p_2)\rangle$ exists. Furthermore, if $w_w(x_n)$ denotes the set of weak subsequential limits of $\{x_n\}$, then $\langle p - q, J(p_1 - p_2)\rangle = 0$, $\forall\, p_1, p_2 \in F$ and $\forall\, p, q \in w_w(x_n)$.*

Theorem 3.3. *Let E be a real q-uniformly smooth and uniformly convex Banach space. Let K be a nonempty closed convex subset of E and $T_i : K \to K$, $i = 1, 2, \ldots, N$ be a k_i-strictly asymptotically pseudocontractive mapping with a sequence $\{k_n^i\}$ such that $\sum_{n=1}^{\infty}(k_n^i - 1) < \infty$. Let $k = \max\{k_i : 1 \le i \le N\}$ and $k_n = \max\{k_n^i : 1 \le i \le N\}$ such that $\sum_{n=1}^{\infty}(k_n - 1) < \infty$ and $F = \cap_{i=1}^N F(T_i) \neq \emptyset$. Let $\{\alpha_n\}, \{x_n\}$ be as in Lemma 3.1. Then $\{x_n\}$ converges weakly to a common fixed point of $\{T_i\}_{i=1}^N$.*

Proof. Let $p \in F$. From Lemma 3.1, $\lim_{n\to\infty} \|x_n - p\|$ exists. So $\{\|x_n - p\|\}$ is bounded. Let $\|x_n - p\| \le M$, $\forall n \ge 1$. It follows from (3.2) that

$$\begin{aligned} M_n\|x_n - T_n x_n\|^q &\le \|x_n - p\|^q - \|x_{n+1} - p\|^q + \delta_n\|x_n - p\|^q \\ &\le \|x_n - p\|^q - \|x_{n+1} - p\|^q + \delta_n M^q. \end{aligned}$$

Since M_n is bounded, we have $\|x_n - T_n^n x_n\| \to 0$, as $n \to \infty$. Observe that $\|x_{n+1} - x_n\| = \alpha_n\|x_n - T_n^n x_n\| \to 0$, as $n \to \infty$. Hence $\|x_{n+i} - x_n\| \to 0$, as $n \to \infty$, for all $i = 1, 2, \ldots, N$. Since T_i is uniformly L-Lipschitzian, then for all $i = 1, 2, \ldots, N$ we have

$$\begin{aligned} \|x_n - T_{n+i}^n x_n\| &\le \|x_n - x_{n+i}\| + \|x_{n+i} - T_{n+i}^n x_{n+i}\| + \|T_{n+i}^n x_{n+i} - T_{n+i}^n x_n\| \\ &\le \|x_n - x_{n+i}\| + \|x_{n+i} - T_{n+i}^n x_{n+i}\| + L\|x_n - x_{n+i}\|. \end{aligned}$$

Then $\|x_n - T_{n+i}^n x_n\| \to 0$, as $n \to \infty$, for all $i = 1, 2, \ldots, N$. So $\lim_{n\to\infty} \|x_n - T_i^n x_n\| = 0$, for all $i = 1, 2, \ldots, N$. We have

$$\begin{aligned} \|x_n - T_i x_n\| &\le \|x_n - T_i^n x_n\| + \|T_i^n x_n - T_i x_n\| \\ &\le \|x_n - T_i^n x_n\| + L\|T_i^{n-1} x_n - x_n\| \\ &\le \|x_n - T_i^n x_n\| + L\|T_i^{n-1} x_n - T_i^{n-1} x_{n-1}\| \\ &\quad + L\|T_i^{n-1} x_{n-1} - x_n\| \\ &\le \|x_n - T_i^n x_n\| + L^2\|x_n - x_{n-1}\| \\ &\quad + L\|T_i^{n-1} x_{n-1} - x_{n-1}\| + L\|x_n - x_{n-1}\|. \end{aligned}$$

Then $\|x_n - T_i x_n\| \to 0$, as $n \to \infty$, for all $i = 1, 2, \ldots, N$. Since $\{x_n\}$ is bounded, it has a weakly convergent subsequence $\{x_{n_j}\}$. Suppose $\{x_{n_j}\}$ converges weakly to p. Since K is convex and closed, K is weakly closed, then $p \in K$. Since $\lim_{n\to\infty} \|x_n - T_i x_n\| = 0$, $i = 1, 2, \ldots, N$, and by Lemma 2.1, $(I - T_i)$ is demiclosed at zero for all $i = 1, 2, \ldots, N$, we have $p \in F$. If $\{x_{m_k}\}$ is another subsequence of $\{x_n\}$ such that $\{x_{m_k}\}$ converges weakly to q , then we have $q \in F$. So $p, q \in F$ and $p, q \in w_w(x_n)$. From Lemma 3.2 we have $p = q$. Hence $w_w(x_n)$ is singleton, so that $\{x_n\}$ converges weakly to a point in F. This completes the proof.

Theorem 3.4. *Let E be a real q-uniformly smooth and uniformly convex Banach space. Let K be a nonempty closed convex subset of E and $T_i : K \to K$, $i = 1, 2, \ldots, N$ be a k_i-strictly asymptotically pseudocontractive mapping with a sequence $\{k_n^i\}$ such that $\sum_{n=1}^{\infty}(k_n^i - 1) < \infty$.*

Let $k = \max\{k_i : 1 \le i \le N\}$ *and* $k_n = \max\{k_n^i : 1 \le i \le N\}$ *such that* $\sum_{n=1}^{\infty}(k_n - 1) < \infty$ *and* $F = \cap_{i=1}^{N} F(T_i) \neq \emptyset$. *Let* $\{\alpha_n\}, \{x_n\}$ *be as in Lemma 3.1. Then* $\{x_n\}$ *converges strongly to a common fixed point of* $\{T_i\}_{i=1}^{N}$ *if and only if* $\lim_{n\to\infty} d(x_n, F) = 0$, *where*

$$d(x_n, F) = \inf_{p\in F} \|x_n - p\|.$$

Proof. By Lemma 3.1, we have $\lim_{n\to\infty} \|x_n - p\|$ exists for all $p \in F$. Also, by (3.3) we have

$$d(x_{n+1}, F) \le (1 + \delta_n) d(x_n, F), \quad n \ge 1.$$

From Lemma 2.2, $\lim_{n\to\infty} d(x_n, F)$ exists. Now if $\{x_n\}$ converges strongly to a common fixed point of $\{T_i\}_{i=1}^{N}$, then $\lim_{n\to\infty} \|x_n - p\| = 0$. Since

$$0 \le d(x_n, F) \le \|x_n - p\|,$$

we have $\lim_{n\to\infty} d(x_n, F) = 0$.

Conversely, suppose $\lim_{n\to\infty} d(x_n, F) = 0$. Thus, for arbitrary $\varepsilon > 0$, there exists a positive integer N_0 such that

$$d(x_n, F) < \frac{\varepsilon}{3}, \quad \forall n \ge N_0.$$

Furthermore, $\lim_{n\to\infty} \prod_{j=n}^{\infty} a_j = 1$ implies that there exists a positive integer N_1 such that $\prod_{j=n}^{\infty} a_j < \frac{3}{2}$, $\forall n \ge N_1$. Choose $N = \max\{N_0, N_1\}$. Since $\|x_{n+1} - p\| \le a_n \|x_n - p\|$, for all $n, m \ge N$ and $p \in F$, we have

$$\begin{aligned} \|x_n - x_m\| &\le \|x_n - p\| + \|x_m - p\| \\ &\le (\textstyle\prod_{j=N}^{n-1} a_j + \prod_{j=N}^{m-1} a_j)\|x_N - p\| \\ &\le 2 \textstyle\prod_{j=N}^{\infty} a_j \|x_N - p\|. \end{aligned}$$

Taking infimum over all $p \in F$, we obtain

$$\|x_n - x_m\| \le 2 \prod_{j=N}^{\infty} a_j d(x_N, F) < \varepsilon.$$

Thus $\{x_n\}$ is Cauchy sequence. Suppose $x_n \to p$, as $n \to \infty$. From Theorem 3.3, $\lim_{n\to\infty} \|x_n - T_i x_n\| = 0$, $i = 1, 2, \ldots, N$. So $p \in F$. This completes the proof.

Acknowledgments.

This project is supported by the National Natural Science Foundation of China under grant (10771050).

References

[1] H.K. Xu: Inequalities in Banach spaces with applications, Nonlinear Anal., 1127-1138 (1991), TMA16(2)

[2] G.L. Acedo, H.K. Xu: Iterative methods for strict pseudocontractions in Hilbert spaces, Nonlinear Anal., 2258-2271 (2007), TMA67(7)

[3] M.O. Osilike, Y. Shehu: Cyclic algorithm for common fixed points of finite family of strictly pseudocontractive mappings of Browder-Petryshyn type, Nonlinear Analysis, 3575-3583 (2009), 70

[4] H. Zhang, Y.F. Su: Strong convergence theorems for strict pseudo-contractions in q-uniformly smooth Banach spaces, Nonlinear Anal., 3236-C3242 (2009), 70

[5] H. Zhang, Y.F. Su: Convergence theorems for strict pseudo-contractions in q-uniformly smooth Banach spaces, Nonlinear Anal., 4572-4580 (2009), 71

[6] M.O. Osilike, Y. Shehu: Explicit averaging cyclic algorithm for common fixed points of a finite family of asymptotically strictly pseudocontractive maps in Banach spaces, Nonlinear Anal., 1502-1510 (2009), 57

[7] Z.B. Xu, G.F. Roach: Characteristic inequalities of uniformly smooth Banach spaces, J. Math. Anal. Appl., 189-210 (1991), 157

[8] M.O. Osilike, S.C. Aniagbosor and B.G. Akuchu: Fixed points of asymptotically demicontractive mappings in arbitrary Banach spaces, Panamer. Math. J., 77-88 (2002), 12(2)

[9] M.O. Osilike, A. Udomene, D.I. Igbokwe and B.G. Akuchu: Demiclosedness principle and convergence theorems for k-strictly asymptotically pseudocontractive maps, J. Math. Anal. Appl., 1334-1345 (2007), 326

[10] R.E. Bruck: A smple proof of the mean ergodic theorem for nonlinear contraction in Banach spaces, Isreal J. Math., 107-116 (1979), 32(2-3)

doi:10.4028/www.scientific.net/AMM.50-51.437

The study of the Brake test bench model

Yuhuan Cui[1, a], Baofeng L i[2], Jingguo Qu[1,b] and Zhen Wang[1]

[1] College of Light industry, Hebei Polytechnic University, 063000 ,China

[2] Department of Mathematics and Information Science, Tangshan Normal University, Tangshan, China

[a]yuhuan,cui@yahoo.com.cn, [b]qujingguo@163.com

Keywords: Brakes; Inertia; Drive Current; Torque; Gray Statistics

Abstract. This paper will focus on the control method of the brake test- bench, we have used physics, theoretical mechanics, numerical methods method and use Matlab software, Excel software to calculated data and drawing. Eventually we establish a mathematical model to control electric current.

$$I(i+1)=1.5*[(10.83333+60.1984/0.0007)*(e^{0.007i}-e^{0.007(i-1)})+(-49.3937*0.9987-144.0344)*(e^{0.9987i}-e^{0.9987(i-1)})].$$

The model will help us testing the merits of the brake, and assist us in the design process.

Introduction

There are driving brakes and parking brake. In the driving process, we generally use driving brakes to facilitate the process of deceleration in the forward stop; when the car have completely stopped, use the parking brake to prevent the vehicle front and rear slip slide. Therefore, during operation of the vehicle, the brake is the important automotive components for driver. a car, the brake design directly relates to the life and property. Test a car's brakes and improving qualified brake brakes braking effect is the important task of the designer[1].

Bench refers to the stage in a special brake test on the road test designed to simulate the test, make up the shortage that the road can not test. The principle is that let the brake test on bench wiht the vehicle brakes on road test during braking as consistent as possible[2].

Brake test rig installed the flywheel generally spindle, the drive spindle motor, base, auxiliary devices and brake applied measurement and control system, etc., as shown in Fig 1.

(1)the measured brake device, (2) twist moment sensor, (3) flywheel plate, (4) safety brake, (5) speed sensor, (6) tachometer generator

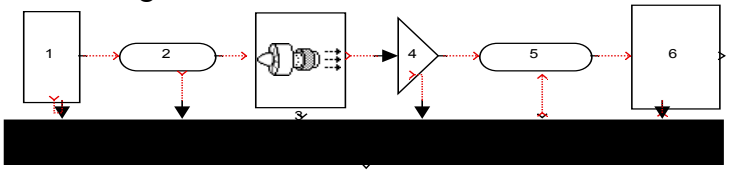


Figure (1) brake test rig the general structure

The model assumptions

1. Assuming that the friction between tire and the ground is infinite in the road test, so the tire and the ground without sliding.
2. The spindle angular velocity and angular velocity of the wheel is always consistent.
3. Spindle and the flywheel and etc. are the rigid body on the experimental bench
4. The equivalent in energy conversion process has ignored the wheels turning their energy.
5. Spindle acceleration is certain.
6. We do not consider the error of energy measurement error, random error and continuous problems arising from the discretization error.

Model and Solution

By the kinetic energy theorem we know that

$$\begin{cases} E_p = \frac{1}{2}mv^2 \\ E_d = \frac{1}{2}J_d\omega^2 \\ E_p = E_d \\ r = \frac{v}{w}, m = \frac{G}{g} \end{cases}$$

So we can get the equation $J_d = mr^2 = \frac{G}{g}r^2$. when $r = 0.286m, G = 6230N$, $J_d = 52kg.m^2$, Therefore, the equivalent moment of inertia. is $52kg.m^2$.

The mass flywheel is $m = \rho v = \rho \int_{\text{Inside diameter}}^{\text{Inside diameter}} 2d\pi r dr$, and then from the definition type of inertia $J_d = mr^2$, Obtained flywheel inertia,

$$J_f = r^2\rho \int_{\text{Inside diameter}}^{\text{Inside diameter}} 2d\pi r dr$$

Among them, $d = [0.0392m, 0.0784m, 0.1568m]$, $\rho = 7810$ kg/m3, $\pi = 3.14$, $r_{\text{outside diameter}} = 1m$, $r_{\text{inside diameter}} = 0.2m$

We can obtain the moment of inertia of the three flywheels,

$$J_f = [29.9931\text{kg.m}^2, 59.9862\text{kg.m}^2, 119.9724\text{kg.m}^2],$$

Approximate to the flywheel moment of inertia,

$$J_f = 30\text{kg.m}^2, 30\text{kg.m}^2, 30\text{kg.m}^2$$

Combine with the basis of inertia, $J_{basis} = 10kg$ and the three moment of inertia of the flywheel, we can get the following result.

The basis of inertia $J_{basic}(kg.m^2)$	J_{j1}	J_{j2}	J_{j3}	J_{j4}	J_{j5}	J_{j6}	J_{j7}	J_{j8}
10	40	70	130	100	160	190	220	10

$$J_{j8} = J_{basis}$$
$$J_{j1} = J_{basis} + J_{j1},$$
$$J_{j2} = J_{basis} + J_{f2},$$
$$J_{j3} = J_{basis} + J_{f3},$$
$$J_{j4} = J_{basis} + J_{f1} + J_{f2},$$
$$J_{j5} = J_{basis} + J_{f1} + J_{f3},$$
$$J_{j6} = J_{basis} + J_{f2} + J_{f3},$$
$$J_{j7} = J_{basis} + J_{f1} + J_{f2} + J_{f3},$$

Because the equivalent moment of inertia $J_d = 52kg.m^2$, the motor can compensate for the inertia of the energy range corresponding to [-30, 30] kg • m2. According $J_d - J_j$, we can get the compensation inertia to meet the conditions. $J_b = 12kg.m^2, J_b = -18kg.m^2$.

According to the theoretical knowledge of mechanics $M_b = aJ_b$, and then meat the meaning of the questions $I = 1.5M_b$,

Because $a = \frac{\alpha}{r}$, therefore

$$I = 1.5\frac{J_b\alpha}{r}$$

According to physics

$$\alpha = \frac{v_t - v_0}{t}$$

Therefore,

$$I = 1.5\frac{J_b(v_t - v_0)}{rt}$$

(Motor drive current depends on the observable mathematical model)

Among them,

$$J_b = [12kg.m^2, -18kg.m^2], \quad r = 0.286\,\text{m}, \quad \begin{matrix} v_t = 0\text{km}/\text{h}, \\ v_0 = 50\ \text{km/h}, \end{matrix}$$

So, I = [-174.8266A, 262.2399A]

According to kinetic energy theorem, the total energy consumption when the road test,

$$\Delta E = \frac{1}{2}J_d w_t^2 - \frac{1}{2}J_d w_0^2$$

Among them

$$w_t = 2\pi n_t,\ w_0 = 2\pi n_0,\ J_d = 48\,\text{kg.m}^2,\ n_t = 514rpm/60,\ n_0 = 257rpm/60$$

So $\Delta E = 52150J$.

If find braking energy consumption on the brake test bench, we need to find the points within a certain time period of the energy loss, so we find the energy of a single time formula,

$$E_i = M_i\theta_i$$

Among them,

$$\theta_i = w_i\Delta t;$$
$$w_i = 2\pi n_i$$

got $\Delta E_1 = \sum_3^{470} M_i\theta_i = \sum_3^{470} 2M_i\pi n_i\Delta t = 49291.94\text{J}$.

According to the definition of relative error (the ratio of absolute error and the true measured value),

$$\delta = \left|\frac{\Delta E - \Delta E_1}{\Delta E} * 100\%\right| = 5.6008455\%$$

Through access to information, the relative error of evaluation of control method should fluctuations in the 5%, so that the control method can be used, however, the control method is not consistent with control requirements, when the torque due to a certain value , computer controlled current fluctuations cause torque unstable, thereby affecting the value of the observed torque fluctuations around this constant value, leading the torque to the observed value is discontinuous change, when get loss of energy by the sub-periods time, resulting in inaccurate data obtained.

According mathematical model $I = 1.5\frac{J_b\alpha}{r}$

We get a equation about M_b,

$$M_b = \frac{J_b}{JM},$$

According to this mathematical model and data in the table to get a new set of data on M_b, and then the application of this set of data fitted a new curve[3]

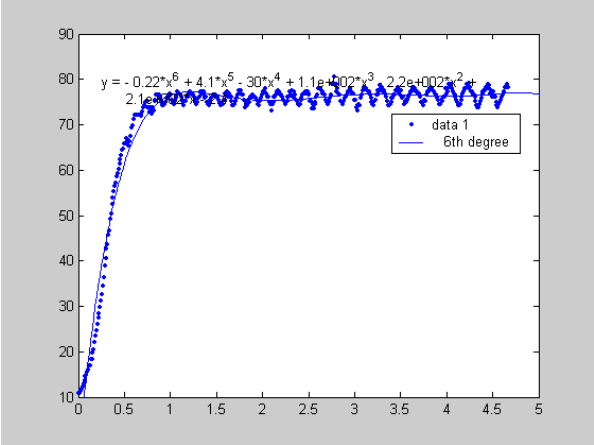


Fig.2

The equation of this curve,

$$Mb = -0.2196*t^6 + 4.0656*t^5 - 29.9807*t^4 + 111.7456*t^3 - 219.5596*t^2 + 212.8460*t - 2.4164 = I/1.5;$$

So

$$I = -1647/5000*t^6 + 7623/1250*t^5 - 6329112625701873/140737488355328*t^4 + 209523/1250*t^3 - 11587599993112677/35184372088832*t^2 + 1404159911553663/4398046511104*t - 18123/5000$$

The equation is the computer control equation for the current value, then evaluate this method, Assuming the brakes's consumption of energy when the road test,

$$\Delta E = \frac{1}{2}J_b w_t^2 - \frac{1}{2}J_b w_0^2,$$

Substituting the obtained data,

$$\Delta E = 14141.97092\text{J}$$

Braking energy consumption during brake test bench,

$$\Delta E_1 = \sum_3^{470} M_i \theta_i = \sum_3^{470} 2M_i \pi n_i \Delta t = 13352.49\text{J}$$

According to the concept of relative error,

$$\delta = \left| \frac{\Delta E - \Delta E_1}{\Delta E} * 100\% \right| = 5.5825\%$$

The model needs to be improved, we can assume uniform changes based acceleration, and then improve the model[4].

Model Evaluation and Promotion

Influence by the previous stage Value of the model is relatively large, relatively slow, so the model should be optimized[5].

We that the nervous control system is very efficient, in recent years many scientists study it, there are mentioned on the mathematical neural network and its application in pattern recognition, so if its introduction into the model can accelerate the speed to fast and efficient control methods, and this is where we should try next[6].

References

[1] Haotao Li, Ying Deng, "MATLAB Programming Guide", Higher Education Press, 2002.8.
[2] Kemeng Zhang, Shengliang Han, "Theoretical Mechanics", Xi'an Jiaotong University Press, 2010.7.1.
[3] Dongyang Shi, "numerical method", Zhengzhou University Press, 2007.5.
[4] Qinshan Fan, "Mechanics of Materials", Higher Education Press, 2008.2.1.
[5] Frank, G., M. Weir, W. Fox. Mathematical Modeling. 3rd edition. Beijing: China Machine Press (2006).
[6] Jing Zhao, Qi Dan. Mathematical Modeling and experiment. 2nd edition. Beijing: Higher Education Press(2003).

©(2014) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.442

The study of the competition model of portal and search engine

Jingguo Qu [1, a], Donghua Wang[2], Yuhuan Cui[1,b] and Yamian Peng[3]

[1] College of Light industry, Hebei Polytechnic University, 063000 ,China

[2] Department of Mathematics and Information Science, Tangshan Normal University, Tangshan, China

3 College of Science, Hebei Polytechnic University, 063000 ,China

[a]qujingguo@163.com, [b]yuhuan, cui@yahoo.com.cn

Keywords: the global stability; the least squares method; search engine; portal

Abstract. This paper is to discuss the competition and cooperation of portal and search engine mainly by building stability model. We can use Lotka-Volterra mutualism model to describe and build cooperation model of portal and search engine, and we can use Lotka-Volterra predator – prey model with Logistic items to describe and build competition model of portal and search engine.

The basic assumption

(1) There is only a search engine in the Internet search market or other search engines click-through rate is very low, they can be ignored
(2) The considered portal in the same market occupies the absolute position, and it is the main source of search engine users and searching content, so we ignore the impact of other portals
(3) Search engines and portals are competing, it is that the number of users M flows to search engine from the portal is less than the number N from search engine to the portal in Unit time
(4) The Web site users click rate (simply click-through rate) means that CTR is the ratio of number of users visiting the website among all the Internet users in the unit time (usually in days)
(5) Ignoring the impact of unexpected events, Such as web to be attacked.

Model

From assumption 3, We known $M < N$.

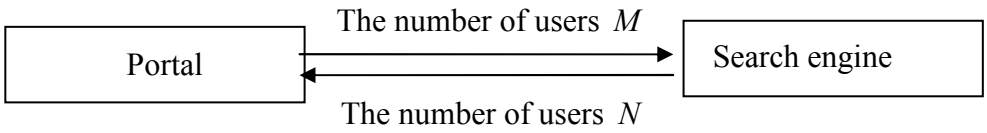


Fig. 1

According to background analysis and Figure 1[1], we have established the following models:

$$\begin{cases} x_1'(t) = x_1(t)(b_1 - a_{11}x_1(t) - a_{12}x_2(t)) \\ x_2'(t) = x_2(t)(-b_2 + a_{21}x_1(t) - a_{22}x_2(t)) \end{cases} \quad (1)$$

Among them: $x_i(t)$ is user click-through rate of website i at the moment t, $x_i'(t) = dx_i(t)/dt$ is the rate of change of $x_i(t)$ $(i = 1,2)$. For convenience, let's remember it as $x_i = x_i(t)$, $i = 1,2$. Since x_1 and x_2 are all positive, we only discussed at $\text{int}R_+^2 \triangleq \{(x_1, x_2) : x_1 > 0, x_2 > 0\}$.

Parameter $b_1(b_1 > 0)$ is the intrinsic growth rate when there is no other portal site's influence; parameters $a_{11}(a_{11} > 0)$ is Block interaction coefficients that portals of their own hits retard the growth factor. Parameter $a_{12}(a_{12} > 0)$ is the competition rate of search engine with portal site;

parameter $-b_2(b_2>0)$ is the internal growth rate of search engine if there is no portal site. Parameter $a_{21}(a_{21}>0)$ is the Popularity rating that the portal site supports search engine, the parameter $a_{22}(a_{22}>0)$ is Block interaction coefficients that the search engine hits on their own retard the growth factor. We have established a portal and search engine competition model, the following that we will qualitative analysis the model.

Mathematical analysis of Model (qualitative and stability analysis)

We will prove that there is no closed trajectory in system (1) , and Analyze the localized states and the overall state of singular point[2].

1 closed trajectory

Theorem 3.1 there is no no closed trajectory in system (1) inside $\text{int}\ R_+^2$.

Proof: Take the Dulac function $B(x_1,x_2)=x_1^{-1}x_2^{-1}$ so $\dfrac{\partial(BX_1)}{\partial x_1}+\dfrac{\partial(BX_2)}{\partial x_2}$

$$=\frac{\partial}{\partial x_1}(x_2^{-1}(b_1-a_{11}x_1-a_{12}x_2))+\frac{\partial}{\partial x_2}(x_1^{-1}(-b_2+a_{21}x_1-a_{22}x_2))=-(\frac{a_{11}}{x_2}+\frac{a_{22}}{x_1})<0$$

When $(x_1,x_2)\in \text{int}\ R_+^2$, there is no no closed trajectory in system (1) inside $\text{int}\ R_+^2$.

2 the stability of critical point

Let $X_1(x_1,x_2)=x_1(b_1-a_{11}x_1-a_{12}x_2)$, $X_2(x_1,x_2)=x_2(-b_2+a_{21}x_1-a_{22}x_2)$,then the singular point of the system (13) is the solution of the following equations:

$$\begin{cases}X_1(x_1,x_2)=0\\X_2(x_1,x_2)=0\end{cases} \qquad (2)$$

Therefore, non-negative singular point of the system (1) is: $O_1(0,0)$， $E_1(b_1/a_{11},0)$， $E_2(x_1^*,x_2^*)$。

and

$$\begin{cases}b_1-a_{11}x_1^*-a_{12}x_2^*=0\\-b_2+a_{21}x_1^*-a_{22}x_2^*=0\end{cases} \qquad (3)$$

So

$$\begin{cases}x_1^*=\dfrac{a_{22}b_1+a_{12}b_2}{a_{11}a_{22}+a_{12}a_{21}}\\x_2^*=\dfrac{a_{21}b_1-a_{11}b_2}{a_{11}a_{22}+a_{12}a_{21}}\end{cases} \qquad (4)$$

When $a_{21}b_1>a_{11}b_2$, E_2 is positive singular point[3]; when $a_{21}b_1=a_{11}b_2$, E_2 and E_1 coincide.

For the stability of singular points in system (1), we obtain the following theorem[4]:

Theorem 3.2 The singular points of system (1) are O_1,E_1,E_2.

(*i*) O_1 is a saddle point;

(*ii*) When $b_1a_{21} < b_2a_{11}$, E_1 is local asymptotically stable point;

When $b_1a_{21} > b_2a_{11}$, E_1 is a saddle point;

When $b_1a_{21} = b_2a_{11}$, E_1 is the saddle node, the part in the upper half plane is locally asymptotically stable node;

(*iii*) When $b_1a_{21} > b_2a_{11}$, E_2 is locally asymptotically stable node or focus;

When $b_1a_{21} \le b_2a_{11}$, E_2 is not in $\text{int}\, R_+^2$.

Proof:

We have to calculate the Jacobi matrix eigenvalues of the system (13), which Jacobi matrices are defined as follows:

$$J \overset{\Delta}{=} \begin{bmatrix} \frac{\partial X_1}{\partial x_1} & \frac{\partial X_1}{\partial x_2} \\ \frac{\partial X_2}{\partial x_1} & \frac{\partial X_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} b_1 - 2a_{11}x_1 - a_{12}x_2 & -a_{12}x_1 \\ a_{21}x_2 & -b_2 + a_{21}x_1 - 2a_{22}x_2 \end{bmatrix} \tag{5}$$

First, the eigenvalue of J at O_1 is $b_1 > 0, (-b_2) < 0$: So O_1 is unstable saddle point[5].

Further, the eigenvalue of J at E_1 is $(-b_1) < 0$, $a_{21}(\frac{b_1}{a_{11}} - \frac{b_2}{a_{21}})$, then, when $b_1a_{21} < b_2a_{11}$, E_1 is locally asymptotically stable node; When $b_1a_{21} > b_2a_{11}$, E_1 is a saddle point. When $b_1a_{21} = b_2a_{11}$, there is a zero eigenvalue. Do coordinate transformation, so that E_1 As the starting point, let $y_1 = x_1 - b_1/a_{11}$, $y_2 = x_2$, $\tau = -t$, the system (1) becomes

$$\begin{cases} \frac{dy_1}{d\tau} = (a_{11}y_1 + a_{12}y_2)(y_1 + \frac{b_1}{a_{11}}) \\ \frac{dy_2}{d\tau} = (-a_{21}y_1 + a_{22}y_2)y_2 \end{cases} \tag{6}$$

The zero solution $E_1^{'}(0,0)$ of system (6) corresponds to E_1, from the first formula of (6), we get

$$y_1 = -\frac{a_{12}}{a_{11}}y_2 + o(y_2) \tag{7}$$

Substituting to the second formula of (6), we obtained

$$\psi(y_2) = \frac{a_{12}a_{21} + a_{22}a_{11}}{a_{11}}y_2^2 + o(y_2^2) \tag{8}$$

Since $(a_{12}a_{21} + a_{22}a_{11})/a_{11} > 0$, therefore we have known that $E_1^{'}(0,0)$ is saddle node, which in the last half-plane is the unstable node (Parabolic sector). Therefore E_1 is a saddle node too, but the trajectory direction is the opposite with $E_1^{'}(0,0)$, its part in the upper half plane is locally asymptotically stable node.

The last: the value of J at E_2 is:

$$J\big|_{E_2} = \begin{bmatrix} -a_{11}x_1^* & -a_{12}x_1^* \\ a_{21}x_2^* & -a_{22}x_2^* \end{bmatrix} \tag{9}$$

x_1^* x_2^* satisfy (4). Therefore the eigenvalues λ of (9) satisfy the equation:

$$\lambda^2 + (a_{11}x_1^* + a_{22}x_2^*)\lambda + (a_{11}a_{22} + a_{12}a_{21})x_1^*x_2^* = 0 \tag{10}$$

When $b_1a_{21} > b_2a_{11}$, $x_1^* > 0$, $x_2^* > 0$, therefore the two roots of (10) have negative real part, E_2 is locally asymptotically stable node or focus (possibly degenerate). Certificate completion.

3 the stability of singular points at infinity

In order to analyze the global stability of system (1) , we also need to consider the singular point of system at infinity. For Poincaré transformation[9], let $x_1 = 1/x$, $x_2 = u/x$, the system (1) becomes

$$\begin{cases} \dfrac{dx}{d\tau} = -x[-a_{11} + b_1x - a_{12}u] \\ \dfrac{du}{d\tau} = u[(a_{11} + a_{12}) - (b_1 + b_2)x - (a_{22} - a_{12})u] \end{cases} \tag{11}$$

Among them $d\tau = dt/x$. Let $x = 0$, we got singularity point $O_2(0,0)$ and $E_3(0,(a_{11} + a_{21})/(a_{22} - a_{12}))$. Here, O_2 and E_3 respectively corresponding singular points of system (1) at infinity which are along the x_1 - axis and the line $x_2 = x_1(a_{11} + a_{21})/(a_{22} - a_{12})$.

Consider the singular points of system (1) at infinity along the x_2- axis, for Poincaré transformation[9], let $x_2 = 1/y$, $x_1 = v/y$, system (1) becomes

$$\begin{cases} \dfrac{dy}{d\xi} = -y[-a_{22} - b_2y + a_{21}y] \\ \dfrac{dv}{d\xi} = v[(a_{22} - a_{12}) + (b_1 + b_2)y - (a_{11} + a_{21})v] \end{cases} \tag{12}$$

Among them $d\xi = dt/y$. We got the singular points $O_3(0,0)$ of system (1) at infinity along the x_2- axis. All types as follows:

Theorem 3.3 All types of singular points at infinity as follows:

(*i*) O_2 is the unstable node (or degraded node);

(*ii*) When $a_{22} > a_{12}$, E_3 is a saddle point;

When $a_{22} < a_{12}$, E_3 is stable node, but not in the first quadrant;

When $a_{22} = a_{12}$, E_3 does not exist;

(*iii*) When $a_{22} > a_{12}$, O_3 is the unstable node (or degraded node);

When $a_{22} < a_{12}$, O_3 is a saddle point;

When $a_{22} = a_{12}$, O_3 is the saddle node, the saddle point is in $\text{int}\, R_+^2$.

Proof:

(*i*) The Jacobi matrix eigenvalue of system (11) at O_2 is: $a_{11} > 0$, $a_{11} + a_{21} > 0$, so O_2 is the unstable node (or degraded node).

(*ii*) The Jacobi matrix eigenvalue of system (11) at E_3 is $\text{t} - (a_{11} + a_{21}) < 0$, $\frac{a_{11}a_{22} + a_{21}a_{21}}{a_{22} - a_{12}}$, so it can be evidence.

(*iii*) The Jacobi matrix eigenvalue of system (11) at O_3 is $a_{22} > 0$, $a_{22} - a_{12}$, so when $a_{22} > a_{12}$, O_3 is the unstable node (or degraded node); when $a_{22} < a_{12}$, O_3 is a saddle point; when $a_{22} = a_{12}$, (12) is the standard form, by the first formula of (12), we got $y = 0$, substituted into the second formula of (12), we got $\psi(v) = -(a_{11} + a_{21})v^2$, and $-(a_{11} + a_{21}) < 0$, so O_3 is a saddle node, the part in $\text{int}\, R_+^2$ is a saddle point.

Conclusions

Because many literature has discussed the cooperation model of portal and search engine[6],this paper will not be addressed on this model in detail.Then for the competition model of portal and search engine, we mathematically analyze the qualitative and stability of the model and prove the global stability of the model. Then we give a series of competitive strategy for portals and search engines.

Finally we test net hits of Tencent and Baidu net from Alexa Web, we use the least squares method to estimate parameters, we get the result: the model we established not only fit with the actual data quite well, and the short-term prediction capabilities are also very good.

References

[1] L.A. Adamic, B.A.Huberman, Power - law distribution of the world wide web, Science, 2000, 287(8), 2115.

[2] L.A.Adamic,B.A.Huberman, The nature of markets on the world wide web, Q.J.Electron. Commerce, 2000, 34(1), 5-12.

[3] Sebastian M. Maurer, Bernardo A. Huberman, Competitive dynamics of web sites, Journal of Economic Dynamics & Control, 2003, 27, 2195-2206.

[4] L.Lopez, M.A.F.Sanjuan, Defining strategies to win in the Internet market, Physica A, 2001 301, 512-34.

[5] Hong Wu, Yuanshi Wang, site competition model of qualitative analysis, Control Theory and Applications, 2005,22 (2) ,218-228.

[6] Zhifen Zhang, etc. Qualitative Theory of Differential Equations, Science Press, 1985

©2010 Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.447

Improved Tikhonov regularization Research and Applied in Inverse Problem

Yamian Peng[1, 2, a], Kaili Wang [1, a] and Huancheng Zhang[2, a]

[1]College of Sciences, He'Bei United University, Tang'shan, He'bei, china

[2]College of Light Industry, He'Bei United University, Tang'shan, He'bei, china

[a]pym806@163.com

Keywords: Partial differential equation, Inverse problem, Tikhonov regularization; ill-posed.

Abstract. The regularization which constructs with the first filter function is precisely the Tikhonov regularization. This article has proven the Tikhonov functional minimization problem is decides suitably, namely satisfies the solution the existence, the solution unique reconciliation to rely on continuously the data stability; and this minimization problem in solves the first class equation equally the normal equation. The numerical simulation experiment's result indicated that distinguishes the inverse with the regular reduction solution parameter to have the numerical precision to be high and the stability is good and convergence rate quick characteristic.

1 Introduction

If the two problems in a statement or handling problems involved or contains another problem of all or part of the knowledge. Then says one problem and another is for the inverse problem according to J.B.K eller formulation. As far as the partial differential equations be concerned, as for the inverse problem definition: The inverse problem of partial differential equations is refers to from some functional of the partial differential equation's solution to determine the coefficient or right margin the item of the partial differential equation." For example a sheet stretch process, what needs to know is that the sheet nature also has the sheet when the withstanding stretch receives drawing force. Then may obtain the corresponding response by these known parameters through the computation. So namely this circumstance and the deformation for problems. But sometimes we are unable to get material parameters or by the load is unknown. Therefore we can measure sheet deformation and the deformation to get material parameters or load process for the inverse problem solving. With the following form to illustrate:

Set space for certain significance and operaters U and F are the metric spaces, and called the solution space and data space. The existing one operator K makes F allude to U, namely $K:F\to U$ and expressed as follows:

$$K(f)=u \qquad f\in F, u\in U$$

The relationship between a system described the working principle. The system as K is mathematical model. The input parameters of f, u for output variable. The so-called direct problem is by the system model and the input parameter inspection system's output, so namely obtains u by known mathematical model K and parameter f through the above relations. But the inverse known outputs to the extrapolation model or the input. The value of u is known and operator K is known and obtaining parameter f through the above relations the process. The operator K is part unknown and also for an undetermined parameter part.

If on above defined available popular saying that was the direct problem for obtaining the output response by the model and the input parameter, but the inverse obtains the input parameter by the output response and the system model. Perhaps obtains system's partial model parameter by the output response and the input parameter.

As a result of inverse's ill-posed ness and the misalignment causes its theory and the solution are more difficult than direct problem. Moreover involves the aspect of knowledge to be broad and

involves the partial differential equation and the ordinary differential equation and optimized method transitive volumetric heat mechanics and so on a series of knowledge. Taking the simple member axial tension question as the example and assigning restraint which and load pole's material parameter and receives. Then obtaining pole's various nodes displacement through the finite element equation. But by obtaining the displacement responds in turn the solution material parameter or the restraint and the load is named the inverse problem.

2 The regularization method

Regularization method is an effective and widely used for solving the ill-posed ness problem of stability of approximate solutions of the numerical method. The concept of the problem posed, such as Hadamard definition and qualitative characteristics of the problem will not produce the numerical solution of the nature of the impact of the problem can not be characterized as a numerical approximation of the solution of its disturbance data. Therefore, approximate solution directly to the equation the standard method for the qualitative equations will not produce meaningful numerical results. The extent of the problem ill-conditioned is determined the size of the condition number and the greater the condition number of the morbid nature of the problem is more serious.

Having the following form of Fredholm integral equation:

$$\int_b^a K(x,s)f(s)ds = u(x) \qquad x \in [c,d]$$

The function $K(x,s)$ is a integral kernel function and $f(s)$ is the integrand and a,b are the integral upper and lower limits. Respectively, c,d for the variable x in the upper and lower bounds. It is clear from the known integral kernel function $K(x,s)$ and the decision of integrated function $f(s)$ the right hand side $u(x)$ is positive problem solving process. We consider the inverse problem is that the right hand side of $u(x)$ and the integral equation obtained integrand $f(s)$.

If a given $t(s) = p\ \sin(ms)$, where p,m is a given constant and there are points on the lower limit of $a = 0, p = \pi$ .For any of the integral kernel function $K(x,s)$ and in m are sufficiently large as follows:

$$\int_0^\pi K(x,s)t(s)ds \to 0$$

Therefore

$$\int_0^\pi K(x,s)(f(s)+t(s))ds \to u(x)$$

If so

$$u^\delta = u(x) + p\int_0^\pi K(x,s)t(s)ds$$

From the functions $u^\delta(x)$ can see the approximate to $u(x)$. Its raw data to add a small perturbation $u(x)$, then $u^\delta(x)$ corresponds to the solution for the disturbance data as follows:

$$f' = f(s)+t(s) = f(s)+p\sin(ms)$$

Function $f'(x)$can be described as the approximate solution of $f(x)$, when the p value will be for different problems of varying degrees of approximation and p get great value. The solution accuracy is clearly very poor, that is the problem of data is very sensitive to disturbance response and solutions the uniqueness and stability are not met. So regularization method has emerged to solve such problems as an effective method.

3 The proposed Tikhonov regularization method

Many scholars have done a great deal and made a lot for this kind of problem solving techniques and computation because the sick or ill-posed ness for inverse problem solving. The regularization

method is widely used to attenuate the pathological nature of the problem of a calculation in many proposed inversion method.

To consider any kind of problem as a model, we using an operator equation description as follows:

$$K(f)=u \qquad f\in F,u\in U$$

The type of a mapping operator K, f and u belong to the corresponding parameter space. The inverse problem is investigated by the known data u to reverse parameters f for the above form. Typically, the definition of generalized inverse operators and the generalized inverse method to express the parameters of the inversion results. Thus the unknown parameters f of the least squares norm solution $f_{ls}=K^{+}u$ of which, K^{+} is the Moore-Penrose generalized inverse and the definition of generalized inverse problem of pathological characteristics according to different differentiated. The regularization method is to consider the known data and calculated data and least squares and based on known information provided by the root of the problem to limit the solution to construct a functional on the determined parameters. The application scale factor to coordinate the relationship between the two, as described singular value of the function of the amendment in order to get the stable solution.

It is based on the operator's singular value decomposition for simple linear problem, and then constructed on the singular value of the correction function. The Tikhonov regularization method for the correction function for such problems as follows:

$$\omega_{\alpha}(s^2)=\frac{s^2}{s^2+\alpha}$$

The parameter s is the system of singular value and α is the regularization parameter. For large-scale pathological system, singular value decomposition is difficult or impossible to achieve, then there is the expression used variation Tikhonov regularization method as follows:

$$M^{\alpha}[f]=\rho_U{}^2(K(f),u)+\alpha\Omega[f]$$

The functional Tikhonov regularization method is known as the flattening of functional (smoothing functional).Where $\alpha>0$ is the regularization parameter. This parameter weighed the request of the solution accuracy and stability, regulating the relationship between the two in order to reverse the error of solutions by small $\rho_U{}^2(K(f),u)$ for the corresponding norm of space considerations. It is usually in the second norm. The operater $\Omega[f]$ or the penalty for the stability of functional and its definition of information based on the number of unknown parameters to understand the difference is usually norm of the parameters or their derivatives to define the norm. Reverse the first question in the investigation into the space has an extreme value as follows:

$$\min_{f\in F} M^{\alpha}[f]$$

Parameters belong into reverse to find an optimal parameter space makes the target to reach the minimum and the optimal parameter shall approximate solution of inverse problems. The Tikhonov regularization for the construction of the problem operator, through a functional extremism treatment to achieve the purpose to solve the problem.

4 Numerical Simulation

We use the iterative method and select the basis function of $\{1,x,x^2,\cdots\}$. The first three $(n=3)$ force $c(x)$ close to solving the following parameter identification problem (I) and (I') as follows:

$$\begin{cases}\dfrac{\partial u}{\partial t}=\dfrac{\partial}{\partial x}\left(c(x)\dfrac{\partial u}{\partial x}\right) & 0<x<1,t<0\\ u(x,0)=x^3 & 0\le x\le 1\\ u(0,t)=0,u(1,t)=e^t & t\ge 0\end{cases} \qquad (\mathrm{I})$$

Additional information

$$u(x,1)=e\cdot x^3 \qquad (I')$$

Exact solution of this problem is: $c^*(x)=\frac{1}{12}x^2, u^*(x,t)=x^3e^t$ .If taking the initial value as follows:

$$c_0(x)=0.1-0.1x+0.1x^2$$

The regularization parameter $\alpha=0.06, \tau=0.1$ and the variable x where the interval [0,1] was divided into 11 nodes $[0:0.1:1]$. The time variable t where the interval [0,1] was divided into5 nodes $[0:0.2:1]$. The numerical results the following table:

iterative times j	$\delta K_j=\left(\delta k_1{}^j,\delta k_2{}^j,\delta k_3{}^j\right)$	$K_{j+1}=\left(k_1{}^{j+1},k_2{}^{j+1},k_3{}^{j+1}\right)$
1	(-0.0983,0.0295,-0.0999)	(0.0017,-0.1295,0.1999)
10	(-0.0003,0.0013,-0.0010)	(0.0010,-0.0049,0.0869)
30	$(-0.2864,0.9625,-0.6904)\times10^{-5}$	(-0.0004,-0.0002,0.0835)
50	$(-0.1882,0.6087,-0.4536)\times10^{-7}$	(-0.0004,-0.0002,0.0834)
70	$(-0.1239,0.3999,-0.2966)\times10^{-9}$	(-0.0004,-0.0002,0.834)
90	$(-0.2101,0.2323,0.1689)\times10^{-11}$	(-0.0004,-0.0002,0.834)

Exact solution is $k^*=(0,0,0.8333)$ and the numerical solution is k_{91}=(-0.0004,-0.0002,0.0834).

5 Summary

The use of regularization method in solving the time and the selection of the regularization parameter α is greater impact on the effect of the algorithm. Since it is mainly the role of regulatory factors. If α is larger, then the stability of functional role of D on the large and the stability of the solution process the better. But the deviation increased with the original problem. If α is minor, then the stability and functional role of D is also small and the solution may not achieve stability in the role. There are several real calculation shows that the test of the regularization parameter was best to take $\alpha=10^{-5}$ .From the numerical simulation results show that the proposed numerical iterative algorithm has high accuracy and good stability and fast convergence characteristics.

Acknowledgement

Fund item: Hebei United University science study imbrues item(z201019).

References

[1] Victor Isakov. Stefan Kindermann.. Identification of the diffusion coefficient in one-dimensional parabolic equation. Inverse Problem. 2000, 16:665-680.

[2] Lan Li. Inverse problem of longitudinal dispersion coefficient for river water quality. Water conservancy journal, 1998, (8):67-81.

[3] Victor Isakov, Stefan Kindermann. Identification of the diffusion coefficient in one-dimensional parabolic equation. Inverse Problem. 2000,16:665-680.

[4] Barbarra Kaltenbacher. Regularization by projection with a posterior discretization level choice for level choice for linear and nonlinear ill-posed problem. Inverse problem. 2000,16:1523-1539.

[5] Andreas Kirsch. An Introduction to the Mathematical Theory of Inverse Problem. New York: Springer-Verlag. 1996.

©2014 Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.451

Solution to China's GDP Prediction Problem by BP Neural Network

Guanchen Zhou[1], Kaili Wang[2], Tong Cui[1]

[1] College of Light Industry Hebei United University Hebei Tangshan 063009,China

[2] College of Science Hebei United University Hebei Tangshan 063009,China

43698059@qq.com

Keywords: GDP, BP neural network, sigmoid function and learning rate.

Abstract. GDP algorithm has become the important means for China's economic management departments to acquaint with the economic operation conditions and the important basis on which the economic development strategies and planning and various macroeconomic policies are worked out by them. Therefore, research and establishment of GDP model have an important and practical significance. In this paper, the BP neural network is adopted to perform the research and prediction of China's GDP. Firstly, the original data of larger ones are observed and introduced into the commonly applied function of sigmoid function, but the effect is not ideal and the normalization point concentrates on $y=1$, which may affect the predicted values, so a simple improvement is made for the sigmoid function to make the function value distribute between 0 and 1 and then the normalized function after improvement is better.

0 Introduction

GDP is short for gross domestic product and refers to the sum of market value of all final goods and services produced within a country in a given period of time of a nation or region. It is commonly regarded as the optimum index to balance a national economy condition and the core index of modern national economy accounting system and it reflects not only economic development and people living standard of a country or a district but also strength and wealth of a country. GDP may provide a more complete image of economic conditions and it may judge whether the economy of a country is in the state of contraction or prosperity or not, determine whether the economy is required to be stimulated or restricted and whether it is in the serious recession period or under the threat of inflation by means of GDP. GDP algorithm has become the important means for China's economic management departments to acquaint with the economic operation conditions and the important basis on which the economic development strategies and planning and various macroeconomic policies are worked out by them. GDP is of great significance to economic research and management and is closely related to each of us. In the business circles, GDP is always the important basis on which the management decisions are made. Whether the economy grows or not demonstrates the change in enterprise's management environment in the macroscopic view and the entrepreneurs pay attention to the basic situation of GDP growth. Governments at all levels are always pursuit of GDP healthy development, and in case of unbalanced development or low development speed of GDP it indicates that the problems exist in the economic development and the governments shall adopt appropriate measures to carry out macro-control to guarantee the economic sustained and stable development. Therefore, a timely and accurate prediction on GDP has an important and practical significance.

At present, many methods may be adopted to predict China's GDP and common methods used by the scholars include ARMA model prediction, grey prediction model, DMDH-ARCH combined model, time-series analysis model and data fitting model and so on. The method of time-series analysis model among which is widely applied in the prediction and Zhou Xuan made an analysis of actual GDP time series data from 1987 to 2006 by means of statistical software and established ARMA model and predicted the CDP of the period from 2009 to 2011 and achieved a good prediction result[1]. Tan Guolan and Li Jian and other persons adopted the grey prediction method and amended the model by means of residual value prediction to promote the prediction accuracy and drawn a

conclusion that the kind of prediction model has a good prospect of application of it in GDP prediction[2]. Ding Wenbin compared the regression model with the time-series model and the former is popular in practical application, but some requirements and assumption conditions during establishment of model leads to great restriction in practical operation. Dai Yu and other persons applied another method of GM (1,1) to predict the total amount of GDP and achieved a high precision[4].

The neural network may perform the parallel processing for the information on a large scale and has a good robustness and fault tolerance and learns a great amount of information and abstracts the objective laws implied in data and is applicable to classification, identification and prediction and other works. The neural network models are of various types and the most representative and widely applied among which is BP model because of its simple structure, reliable basis, rigorous derivation process, high precision, better universality and ability to simulate the non-linear input and out relations, but the standard BP algorithm has such disadvantages as follows: slow convergence speed and easy to fall in local minimum value; however, there are many improvement methods at present. In this paper, it attempts to apply the BP neural network to make an analysis of China's GDP from 1978 to 2007 and predict the GDP in the coming five years.

1 Model Improvement

Because the choice and important of learning rate η, the higher of η and the faster convergence it will be, but it may cause instability or function vibration if η is too high; if η is lower, although it may avoid instability, the speed of function convergence will reduce. In order to solve the contradiction, we introduce a variable as follows:

$$W_{ij}(n_0+1)=W_{ij}(n_0)-\eta\frac{\partial F}{\partial W_{ij}(n_0)}+\alpha(W_{ij}(n_0)-W_{ij}(n_0))=W_{ij}(n_0)+\Delta W_{ij}(n_0)$$

We can obtain that

$$\Delta W_{ij}(n_0)=-\eta\frac{\partial F}{\partial W_{ij}(n_0)}+\alpha\Delta W_{ij}(n_0-1)$$

Where $\alpha\Delta W_{ij}(n_0-1)$ is a variable, $0<\alpha<1$

It action is that: when the training samples are added orderly, the equation above becomes the time series with n_0 as a variable and m is from 0 to n_0; therefore, the equation above may be regarded as the first order difference equation of $\Delta W_{ij}(n_0)$ and $\Delta W_{ij}(n_0)$ may be solved according to the equation below:

$$\Delta W_{ij}(n_0)=\eta\sum_{m=0}^{n_0}\alpha^{n_0-1}\delta_i(m)y_j(m)=-\eta\sum_{m=0}^{n_0}\alpha^{n_0-1}\sum_m\frac{\partial F(m)}{\partial W_{ij}(m)}$$

If the $\sum_m\frac{\partial F(m)}{\partial W_{ij}(m)}$ this time is the same as that of the previous time, the weighted summation value will increase and make the $\Delta W_{ij}(n_0)$ increase and it results in the regulation speed of right value W at the stable regulation; and if the $\sum_m\frac{\partial F(m)}{\partial W_{ij}(m)}$ this time is contrary to that of the previous time, it indicates that a certain vibration and now the result of summation will make the value of $\Delta W_{ij}(n_0)$ decrease to play a role in stability.

Therefore, the error function surface of the function may become smoother through change in momentum term and we may take the large η values to accelerate the convergence speed without generating the vibration.

2 Application of BP Neural Network in China's GDP Prediction

Because that the units of data we obtained are different, it needs to perform [-1,1] normalization processing for the data to facilitate the data processing and at the same time accelerate the convergence during program execution so as to improve the disadvantage of slow convergence speed of the standard BP neural network. Because that the common conversion relationship function is sigmoid function, the scatter diagram conversed by the function for $y = \frac{1}{1+e^{-x}}$ is shown as follows:

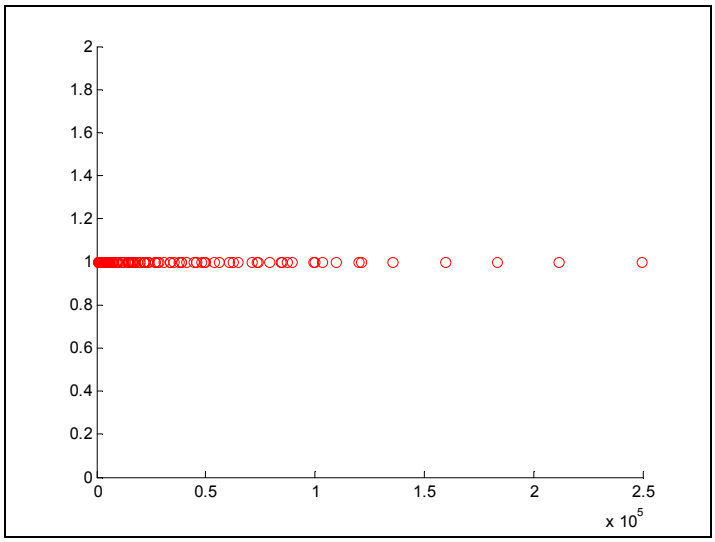


Fig. 1

The scatter diagram shows that the values of y are around 1 and have small variation, so it may affect the predicted value; therefore, we make a small adjustment for the function to make the values distribute between 1 and 1, and the function is $y = \frac{1}{1+e^{-\frac{x}{10000}}}$ and the scatter diagram is shown as follows:

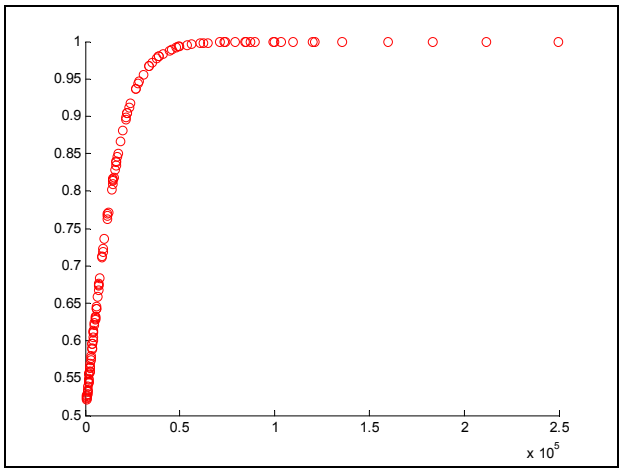


Fig. 2

It shows obviously that the function after improvement contributes to promotion of prediction precision.

Determine the number of training steps respectively:Work out the diagram of relation between training steps and errors respectively.Primary normalized function

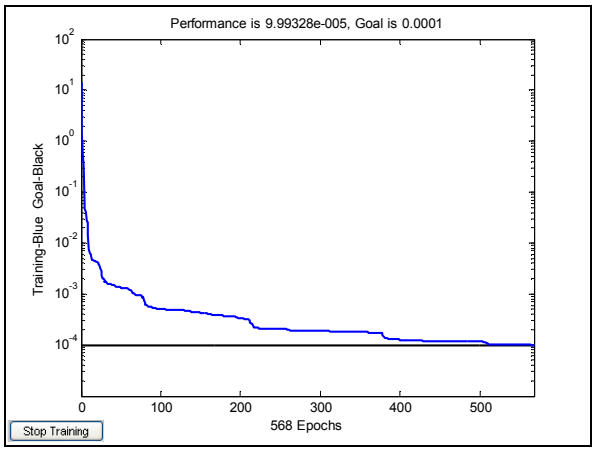


Fig. 3

The number of training steps is determined to be 600 according to the diagram of relation between the number training steps and error.The simulative diagram based on it is shown as follows:

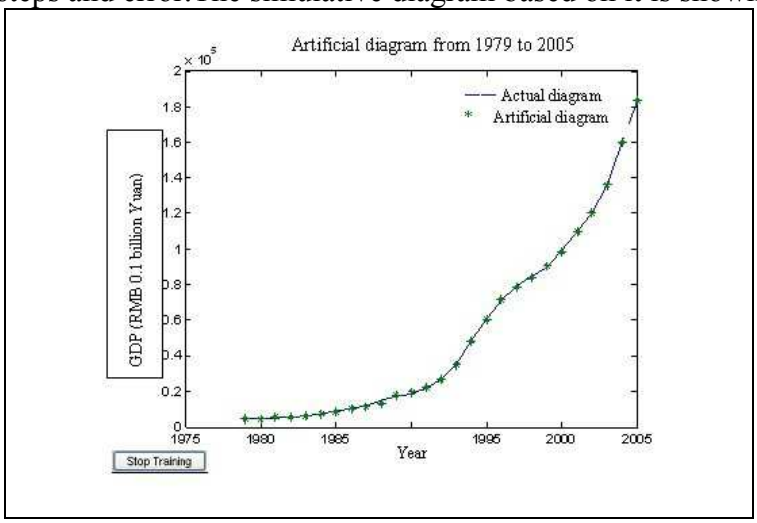


Fig. 4

References

[1] Dai Yu, Wang Yuanyuan and Wang Lunfu. Prediction on total quantity of GDP of Anhui Province based on grey GM(1,1) model, Journal of Chongqing Institute of Technology: Nature Publishing, Vol. 22 No. 2: 74-74(2008).

[2] Cheng Xin and He Yue, research on application of GMDH-ARCH combined model in prediction of GDP, Statistics and Consulting, Vol. 5(2007).

[3] Han Zhongxiang, Dong Shufu, Zhang Rui and Tao Xiaoyan. Intrusion detection algorithm based on counter propagation neural network, Journal of Air Force Engineering University(Natural Science Edition), vol.10(2009).

[4] Zhu Shan. Establishment and analysis of gross domestic product model, Modern Economic Information.

[5] Dong Chanhong. Application of Matlab neural network. Beijing: National Defence Industrial Press, 2005.

doi:10.4028/www.scientific.net/AMM.50-51.455

Numerical Method Research of Partial Differential Equations Inverse Problem

Yali He[1, a], Yamian Peng[1,2, a] and Lichao Feng[1, a]

[1]College of Sciences, He'Bei United University, Tang'shan, He'bei, china

[2]College of Light Industry, He'Bei United University, Tang'shan, He'bei, china

[a]pym806@163.com

Keywords: Inverse Problem; Genetic Algorithm; PDE; Tikhonov regularization

Abstract. It is feasible for the inverse problem of research in the very vital significance between in practical application. Genetic algorithm is applied in many aspects, but we are more concerned with the application in mathematics. From the start of genetic algorithm, the collection to search for comprehensive coverage of preferred. Due to genetic algorithm is used to search the information, and does not need such problems with the problem is directly related to the derivative of the information. Finally, the results of numerical simulation show that the GA method has high accuracy and quick convergent speed. And it is easy to program and calculate. It is worth of practical application.

1 Introduction

In the natural sciences and the engineering field that has many problems of partial differential equation can be used to describe. The partial numerical solution of differential equations is a powerful tool for solving these problems. Partial differential equations of the inverse problem due to its actual background and study from the vast field and multi-disciplinary characteristics and the theoretical research and practical application has become a specialized subject with the important significance. There are many scholars at home and abroad to study in this field, and use various numerical methods, and the latest research results to solve various problems of partial differential equation.

The basic structure and the basic process of research based on genetic algorithm theory and application of some research work on the traditional genetic algorithm and based on analysis of traditional genetic algorithm, the improvement of the traditional genetic algorithm extends the concept of community. According to the biological "multiply, the principle of" survive competition, subdivision of the traditional genetic algorithm was proposed.

2 Partial Differential Equations Inverse Problem

From the perspective of mathematics, the differential equation of the inverse problem of research in the three aspects. Firstly, to solve the problems in the theory, the correctness of the formulation. Usually the inverse problem of complete description should contain differential equation, definite (they are composed of differential equation and the additional problem). Additional conditions and general refers to the solution of the parameters and the boundary condition. The second is the inverse problem of solving methods. The problem for differential equation is the development of a series of numerical analysis method, etc. But these methods are usually not directly used to solve inverse problem. Therefore for different types of inverse problem must through the proper operation and transform the pattern can convert.

At present, the development of fully nonlinear inverse methods, but represents the direction of research, the theory of inversion is inversion of theoretical research frontiers. There is widely used in the practical choice fitting method, the method, and the approximate formula of the methods and replace fitting process.

For equations as follows:

$$Az = u \tag{1}$$

In the formula $u \in U$ and $z \in F$. U and F is metric Spaces. Operator to reflect A F U .Again, there are inverse operator hypothesis A^{-1}， but in general A is tight operator, it is not a continuous equation and its solution is ill-posed.

In practice, a method of calculating formula is approximate solution of (1) is widely used. This is a possible solution can be given in advance of the subclasses $M(M \in F)$ element and calculate Az . Which namely solution is a problem. An element in the collection of M such as approximate z_0 as the approximate solutions which in the gap of integrating M make $P_u(Az,u)$ minimum as follows:

$$0 \le \rho_u (Az_0,u) = \inf_{z \in M} \rho_u (Az,u)$$

But because of the original material is error, and may not belong to the set elements $u \in AM$, in this case, the equations (1) is not exist solution (classical solution), thus introducing fitting. The right solution for fitting, approximate $u \notin AM$. Therefore, according to the formula we can get $z = A^{-1}u$: constituent equation (1) the approximate solution. This is the u projection and have u to AM again, That is, the solution of infinite dimension to approximate solution to turn the finite dimensional space. But there is only one in a given set of sufficient conditions of the solution is M to set is convex tight and space within the sphere is strictly convex.

The inversion method is to substitute for an operator to heat it "close" operator. Approximate solution method and with its approximate formula of the type and the method of solution process may require such is set. However, many problems of the application of such tight, and not with the approximate formula right associated change, can make it out of AM .

Regularization method is to establish the approximate solution of the equation $Az = u$ in the original material, it changed circumstances mentioned above can't solve the problem is not top of stability.

Tikhonov regularization method is to ask questions of the solution of equation $\alpha z_{\alpha,\delta} + A^{\bullet} A z_{\alpha,\delta} = A^{\bullet} u_{\delta}$ for Tikhonov functional transformation of the minimum value as follows:

$$J_{\alpha}(z) := \|Az - u\|^2 + \alpha \|z\|^2 \quad x \in X$$

For Tikhonov functional J_{α} only α is minimum, and the minimum value is the only solution.

$$\alpha z_{\alpha,\delta} + A^{\bullet} A z_{\alpha,\delta} = A^{\bullet} u_{\delta}$$

The shaped equation of regularization parameter is the inverse problem solving, so the equation for the transformation of the solution, but this is α must in advance, if not given α proper. The regularized solutions is obtained.

3 Genetic algorithms solve the initial value inverse problem

The following with parabolic equation parameter identification counter-question titled example is given to illustrate the application of genetic algorithm.

Considering the rectangular domain R inside the parabolic partial differential equations of the initial-boundary value problem:

$$\text{(I)} \begin{cases} p(x)\dfrac{\partial^2 h}{\partial t} = \dfrac{\partial}{\partial x}(q(x)\dfrac{\partial h}{\partial x}) + c(x,t)\dfrac{\partial h}{\partial x} + d(x,t)h + f(x,t) \\ h(\alpha,t) = g_1(t), h(\beta,t) = g_2(t); \qquad 0 \le x \le a, 0 \le t \le b \\ h(x,0) = h_0(x);\ h_t(x,0) = h_1(x); \end{cases}$$

In the equation, $p(x), q(x), c(x,t), d(x,t), f(x,t), g_1(t), g_2(t), h_0(x), h_1(x)$ is given. The inverse problem is according to the equations (I) and other known additional information to determine the $q(x)$. In order to identify the parameters, often based on certain observation. It is nonlinear optimization problems. The different ways of observation we often have the following kinds:

(1) When we get a little observation at different moments of $h(x_0,t_i), (i = 1,2\cdots n)$ value.

(2) When we get the same time difference observation on $h(x_i,t_0),(i=1,2\cdots n)$.

(3) When we get the observation on the differences in different times $h(x_i,t_j)$, $(i=1,2\cdots n, j=1,2\cdots m)$

If the fixed functions $p(x),c(x,t),d(x,t),f(x,t),g_1(t),g_2(t),h_0(x),h_1(x)$ is knowm, it is said the corresponding solution is different when equation (I) on the $q=q(x)$ to say the compliance. The equations (I) for the solution of the different $h=h(q(x),x,t)$ observed, according to the parameter $h(x,t)$ recognition problem $q=q(x)$ can be transformed into below the corresponding optimization problem.

(4) $\min\limits_{k\in M}\sum\limits_{i=1}^{n}\left[h(q(x),x_0,t_i)-h(x_0,t_i)\right]^2$

(If you get A bit in different time when $h(x_0,t_i)$, $i=1,2\cdots n$)

(5) $\lim\limits_{k\in M}\sum\limits_{i=1}^{n}\left[h(q(x),x_i,t_0)-h(x_i,t_0)\right]^2$

(If you get a temporary differences on the $h(x_i,t_0),i=1,2\cdots n$)

(6) $\min\limits_{k\in M}\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m}\left[h(q(x),x_i,x_j)-h(x_i,t_j)\right]^2$

(If have differences in different time $h(x_i,t_j),i=1,2\cdots n$, $j=1,2\cdots m$)

For these problems, we have given in chapter 4 of the inversion of unknown variables of the best perturbation iterative method, during iteration algorithm, due to the unknown variables before to determine the $\Phi(x)$ group based on the function as follows:

$$\Phi(x)=(\varphi_1(x),\varphi_2(x),.....,\varphi_n(x))^T$$

To function as a form of group as follows:

$$c^*(x)=\sum_{i=1}^{\infty}k_i^*\varphi_i(x)$$

To simulate the iteration of the unknown variables and functions and to get the coefficient to approximate solution is unknown variables of the form. That is as follows:

$$k^T=(k_1,k_2,...,k_n)\in R^n$$

Question is how to determine the initial value iteration coefficient is close to the initial value can satisfy real solutions, and the precision of the best perturbation iterative solution. This chapter genetic algorithm is used to determine the initial value coefficient.

4 Numerical Simulations

Test kit for two boundary problem of diffusion equation as follows:

$$\begin{cases}\dfrac{\partial u}{\partial t}=\dfrac{\partial^2 u}{\partial x^2}+(1+x^2+t^2)\dfrac{\partial u}{\partial x}+(x+1)e^{-x-1} & 0<x<1,0<t<1\\ u\big|_{x=0}=\varphi(t),u\big|_{x=1}+1+e^t & 0<t<1\\ u\big|_{t=0}=1+x^2 & 0<x<1\end{cases}$$

Additional information is $u(x,1)=x^2+1$.

The true value of $\varphi(t)$ test is $1+t+t^2$. Calculate by additional information on [0, 1] and $t=1$ time points. In the genetic algorithm, the group of 20 individuals and each individual contains 3 to estimate the parameters of the scope and the function $u(x,1)$ parameters in interval [0,1].

After 1000 generation of genetic algorithm, the estimated parameter (a_0, a_1, a_2) for:

$$1.08228429148493+0.81781096910138\,t+1.09998498478719\,t^2$$

Its mean square error is 12904849836121. We sign the approximate solution to "o" asfollows:

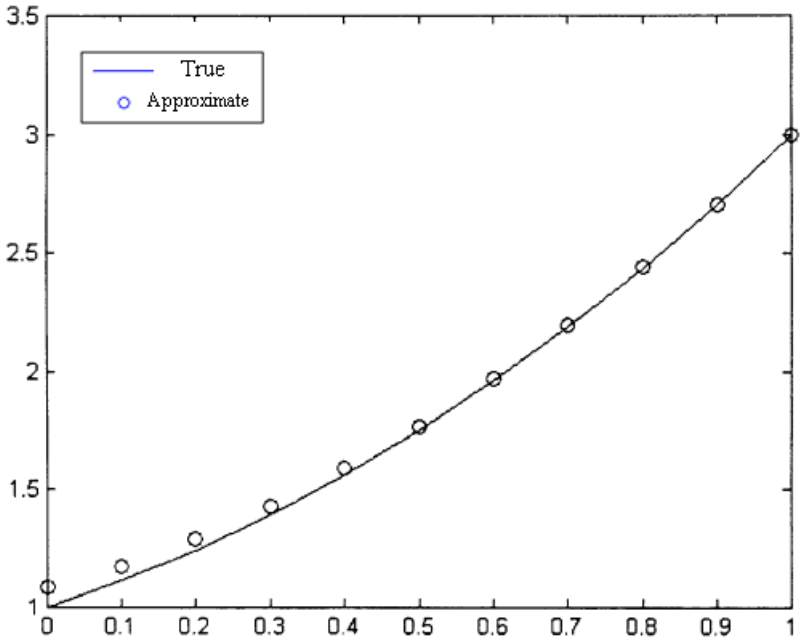


Fig. 1 The approximation and the real value

5 Introduction

The genetic algorithm is adopted random search method of optimal solution, choose to embody the optimal solution is approaching, cross reflects the optimal solution, mutation reflects the global optimal solution of the cover.

Acknowledgement

Fund item: Hebei United University science study imbrues item(z201019).

References

[1] M. S. Pilant, W. Rundell. *An inverse problem for a nonlinear parabolic equation*. Commum Partial Differ Equations. 1986, 11:445~457.

[2] Victor Isakov. Stefan Kindermann. *Identification of the diffusion coefficient in one-dimensional parabolic equation*. Inverse Problem. 2000, 16:665-680.

[3] Jin ZQ, Zhou ZF. *Inverse problem of engineer hydraulics* [M], Hehai university publisher, 1997

[4] Li W. *Headway of environmental hydraulics* [M], WuHan university publisher, 1999

[5] Wang Xiaoping,Cao Liming. *Genetic Algorithms:Theroy and Application and Software.*Xi'an: Xi'an Jiao Tong university publisher,2002.

[6] Wei Peijun. *Numerical Solution of Integral Equations,* Metallurgical Industry Press,2007

[7] Qi Zhongtao,Qian Zhaohua.Ill-posed problem and Tikhonov regularization Methods. Armored Force Engineering,1996.

doi:10.4028/www.scientific.net/AMM.50-51.459

Tikhonov regularization Solve Partial Differential Equations Inverse Problem

Yali He[1, a], Shiqiu Zheng[1, a] and Yanmei Yang[1, a]

[1]College of Sciences, He'Bei United University, Tang'shan, He'bei, china

[a]pym806@163.com

Keywords: Partial differential equation, Inverse, Tikhonov regularization

Abstract. This article discussed the partial differential equation inverse problem. Because the partial differential equation inverse is the misalignment improperly posed problem, therefore analyzed has had the improperly posed reason. To process the partial differential equation inverse correctly not well-posed ness this difficulty, obtains relies on continuously the data stable approximate solution, has drawn support from the regularization related concept and the regularization general theory.

Introduction

The inverse problems exist in many natural science and engineering technology fields, particularly in engineering practice such inverse problem (parameter inverse problems)of undetermined known parameters is the most common. For the practical problems of differential equations, initial conditions, boundary conditions and additional conditions described, namely:

Differential equation: $L(u(x,t), p(x)) = f(x,t) \qquad x \in \Omega, t > 0$

Initial conditions: $I[u(x,t)] = \varphi(x) \quad x \in \Omega, t = 0$

Boundary conditions: $B[u(x,t)] = \psi(x,t) \qquad x \in \Gamma$

Additional conditions: $A[u(x,t)] \equiv h(x,t) = h_0(x,t)$

It is assumed that $u(x,t)$ is the solution of the differential equations, $f(x,t)$ for the right end of the source term, $\varphi(x), \psi(x,t)$ and $h_0(x,t)$ are the initial, boundary and additional conditions, L, I, B, A respectively differential operator, the initial operator, boundary operator and the additional operator son; $p(x)$ as a model system in the physical parameters, Ω for the region of space where the model parameters, Γ for the Ω region boundary.

For the above problem, if the conditions are known, the direct problem is the process of solving the problem $u(x,t)$, these conditions are indispensable. When at least one of the above conditions is unknown, by responding to get this process is unknown conditions for the inverse problem of partial differential equations.

Regularization Theory

Regularization method is Tikhonov and Phillips in the 20th century, made 60 years separately, regularization method through a long development process, the theory more mature. Usually regularization method introduced, first introduce the concept of the problem posed, Hadamard made in the early 20th century. Considering the following formal definition of the equation

$$K(f) = u \qquad f \in F, u \in U \tag{*}$$

F, U belong respectively to the corresponding parameter space, $K : F \to U$ is a mapping operator.

Definition: If the above equation as posed, the need to meet the following conditions:

(1) Existence: For any $u \in U$, there $f \in F$; (Existence)

(2) Uniqueness: Any $u_1, u_2 \in U$, F exist in space, with f_1, f_2 one-one correspondence, and when $u_1 \neq u_2$, then $f_1 \neq f_2$; (unique solution)

(3) relative to the solution space of stability (the stability of solution), for any $\varepsilon > 0$, there is

$\delta(\varepsilon)>0$, as long as follows:

$$\rho_U(u_1,u_2)<\delta(\varepsilon) \qquad (u_1,u_2\in U)$$

There is

$$\rho_F(f_1,f_2)<\varepsilon \qquad (K(f_1)=u_1,K(f_2)=u_2)$$

Where $\rho_U(u_1,u_2)$ means the parameter u_1 and u_2 in the metric space U. The function $\rho_F(f_1,f_2)$ have two parameters f_1 and f_2 in their respective metric space. If these three conditions are met, the problem is posed, or, if there is one or more unsatisfied, it is called improperly posed problem, that malaise.

Target, such as (*)-type operator equations considered, usually the real problem model to consider are:

$$u^\delta = K(f_{true})+\eta$$

Where, u^δ is the data entry with error, η is the error term, f_{true} is defined as the true solution, the operator K, the singular value decomposition:

$$K = Tdiag(s_i)V^T$$

s_i is the singular value, T and V are the singular vector matrix. t_i Remember for the T column vectors, called the left singular vector, v_i is the column vector V, called the right singular vector, then the following relationship:

$$t_i t_j=\delta_{ij} \qquad v_i v_j=\delta_{ij} \qquad Kv_i=s_i v_i \qquad K^T t_i=s_i v_i$$

Where δ_{ij} is directed function, that is as follows:

$$\delta_{ij}=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}$$

And there is $T^T=T^{-1},V^T=V^{-1}$.Then

$$f'=K^{-1}u=Vdiag(s_i^{-1})T^Tu=f_{true}+\sum_{i=1}^{n}s_i^{-1}(t_i^T\eta)v_i$$

f' is the study and the approximate solution.it is an error for the true solution and the role of and. Singular value is too small will result in the error term η is magnified, resulting in the solution can not meet the problem get the required accuracy. One way to solve this problem are singular value m for the amendment, namely singular value in the original by adding a correction function based on the item $\omega_a(s_i^2)$, and the meet when the singular value $s\to 0$, there are $\omega_\alpha(s^2)s^{-1}\to 0$.Have:

$$f_\alpha=Vdiag(\omega_\alpha(s_i^2)s_i^{-1})T^Tu=\sum_{i=1}^{n}\omega_\alpha(s_i^2)s_i^{-1}(t_i^Tu)v_i$$

Where $\omega_\alpha(s_i^2)$ is called a filtering function, $\omega_\alpha(s_i^2)$ the device originates from different, will be different forms of regularization operator.

For example:

$$\omega_\alpha(s^2)=\begin{cases}1 & s^2>\alpha\\ 0 & s^2<\alpha\end{cases} \tag{1}$$

$$\omega_\alpha(s^2)=\frac{s^2}{s^2+\alpha} \tag{2}$$

In this method, the filter function if access to type (1), then known as the truncated singular value decomposition (TSVD) regularization method, principle of this method is to filter out the small singular values; if taken as the type (2), Tikhonov regularization method was the method by adding a correction term to correct the situation of singular value is too small.

Regularization Method to Implement

For the residual $\phi = \left\| K\left(f_\alpha^{\ \delta}\right) - u^\delta \right\|$, from the perspective of approximation to the regularization parameter α should be smaller the better, however, numerical stability point of view, the parameters of α getting bigger is better. On the regular solution $f_\alpha^{\ \delta}$ and the true solution f_{true} of the error estimates of the following:

$$\left\| f_\alpha^{\ \delta} - f_{true} \right\| \le \left\| f_\alpha^{\ \delta} - f_\alpha \right\| + \left\| f_\alpha - f_{true} \right\| \le \left\| R_\alpha \right\| \left\| u_\delta - u_{true} \right\| + \left\| f_\alpha - f_{true} \right\|$$

on the type of data items u_{true} is the non-error, f_α is the grounds of non-error data obtained by regular solution, R_α is said the regularization operator. Type in the true solution in the exact right hand side of f_{true} and u_{true} are unknown, but there is $\lim_{\alpha>0} \left\| f_\alpha - f_{true} \right\| = 0$, so the relationship must exist for a constant $C > 0$, so $\left\| f_\alpha - f_{true} \right\| \le C$, while there are $\left\| R_\alpha \right\| \le 1/\sqrt{\alpha}$, if so $u_T \approx K\left(f_\alpha^{\ \delta}\right)$, get regular solution of the residual value of the upper bound of the approximate estimate

$$\left\| f_\alpha - f_{true} \right\| \le \frac{\left\| K\left(f_\alpha^{\ \delta}\right) - u_\delta \right\|}{\sqrt{\alpha}} + C$$

As the regular solution $f_\alpha^{\ \delta}$ for any regularization parameter $\alpha > 0$ are numerically stable, it should be minimized so that the corresponding residuals, from numerical stability point of view, the parameter α have to get bigger we go as far as possible, so, selected so that the error bound for the small parameter, which determine such a parameter α, so $\varphi(\alpha) = \frac{\left\| K\left(f_\alpha^{\ \delta}\right) - u^\delta \right\|}{\sqrt{\alpha}}$ for the small can. This is a very small error criterion, which accounts for both the approximation and numerical stability requirements.

Achieving of Tikhonov regularization method

First of all, given the initial regularization parameter $\alpha_0 > 0$, the initial guess $f_{\alpha_0}^{\ 0} \in F$, $\varepsilon > 0$ and the observation data u_δ.

(1) initial application of the regularization parameter on the objective function (3.4) type is solved, get extreme point $f_{\alpha_i}^{\ i}$.

(2) if $\left| K\left(f_{\alpha_i}^{\ i}\right) - u_\delta \right| \le \varepsilon$ meet, that is, the solution obtained to meet certain precision, then $f^* = f_{\alpha_i}^{\ i}$ for the approximate solution, turn (4); otherwise go to (3).

(3) According to the calculated approximate solution and the regularization parameter, using principles of re-selection bias regularization parameter, get the appropriate regularization parameter α_{k+1}, by solving the approximate solution $f_{\alpha_{i+1}}^{\ i+1}$, go to (2).

(4)Initial regularization parameter is the man given, for their problems to be large enough to ensure that the revised question is posed. It detects the approximate solution for solving the obtained problem is what we seek the stability of the approximation.

Numerical Simulation

Using iterative method, select the basis function family $\{1, \sin x. \cos x, \sin 2x, \cos 2x, \cdots\}$ of the first 5 items $(n = 5)$ approximation of $c(x)$, the following parameters identification problem solving (I) and (I')

$$\begin{cases} \dfrac{\partial u}{\partial t}=\dfrac{\partial}{\partial x}\left(c(x)\dfrac{\partial u}{\partial x}\right)+e^{-t}\left(\sin^2 x-\cos^2 x\right) & 0<\mathrm{x}<2\pi,\mathrm{t}>0 \\ u(x,0)=\sin x & 0\le \mathrm{x}\le 2\pi \\ u(0,t)=0,u(\pi,t)=0 & t\ge 0 \end{cases} \quad (\mathrm{I})$$

Additional information

$$u(x,1)=e^{-1}\sin x \quad (\mathrm{I}')$$

The exact solution is $c^*(x)=1+\sin x, u^*(x,t)=e^{-t}\sin x$ for the inverse problem.

If the initial value of taking $c_0(x)=0.5+0.4\sin x+0.3\cos x+0.2\sin 2x+0.1\cos 2x$, are the parameters of $\alpha=0.0003,\tau=0.1$, the variable x where the interval $[0,\pi]$ was divided into 31 nodes $[0:\pi/10:\pi]$, the time t where the interval [0,1] was divided into 11 nodes $[0:0.1:1]$, the numerical results the following table:

iterative times j	$\delta K_1{}^j=\left(\delta k_1{}^j,\delta k_2{}^j,\delta k_3{}^j,\delta k_4{}^j,\delta k_5{}^j\right)$	$K_{j+1}=\left(k_1{}^{j+1},k_2{}^{j+1},k_3{}^{j+1},k_4{}^{j+1},k_5{}^{j+1}\right)$
5	(-0.0406,0.0610,0.0002,-0.1274,0.0325)	(1.1162,0.8305,0.0002,-0.0007,-0.906)
30	$(0.1845,0.2718,-0.0003,0.0013,0.1483)\times10^{-4}$	(1.0012,0.9967,0.0000,-0.0000,-0.0000)
50	$(-0.3716,0.5451,-0.0006,0.0026,0.2974)\times10^{-7}$	(1.0012,0.9968,-0.0000,0.0000,0.0000)
80	$(-0.3536,0.5240,-0.0563,0.0905,0.4362)\times10^{-11}$	(1.0012,0.9968,-0.0000,0.0000,0.0000)

Exact solution is K^*=(1,1,0,0,0),Arithmetic solution is K_{81}=(1.0012,0.9968,-0.000,0.000,0.000).

The genetic algorithm is adopted random search method of optimal solution, choose to embody the optimal solution is approaching, cross reflects the optimal solution, mutation reflects the global optimal solution of the cover.

Summary

The use of regularization method in solving the time, the selection of the regularization parameter α is greater impact on the effect of the algorithm. When the data have small errors, caused about a very big difference, but this difference was not caused by the algorithm (in fact two kinds of solutions are exact solutions of the corresponding equation). This situation in practical engineering calculations frequently encountered as the coefficient matrix elements and the right end of the constant is measured in practice.

Acknowledgement

Fund item: Hebei United University science study imbrues item(z201019).

References

[1] Victor Isakov. Stefan Kindermann.. *Identification of the diffusion coefficient in one-dimensional parabolic equation*. Inverse Problem. 2000, 16:665-680.

[2] Ting-yan Xiao. Shen-gen Yu. Yan-fei Wang. *Numerical Method for Inverse Problem*. Science publishing company.2003.

[3] Ting-yan Xiao. Shen-gen Yu. Yan-fei Wang. *Numerical Method for Inverse Problem*. Science publishing company.2003.

[4] Chao-wei Su. *Numerical Method for Inverse Problem of Partial Differential Equation and Application* . Northwestern Polytechnical University publishing company . 1995.10.

©(2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.463

Research on The Model of Positioning System and It's Application

Jin Dianchuan, Meng Xiaoli, Wu Haiming

Hebei United University College of Sciences Hebei Tangshan 063009,China

jdchuan@heut.edu.cn

Keywords: Canny edge detection method, distortion error，Multi-Vision Inspecting Technique scale factor

Abstract. Digital Cameras positioning has a wide range of application in the aspect of traffic monitoring (e-police). The introduction of space Cartesian coordinate system together with the principle of image formation of digital camera and the physical relations among image, camera, and the original picture is where the general linear space model, with the Gaussian formula, Physical optics imaging as the constraint condition, is built, putting forward the theoretical formula of pinpointed coordinates of the center of the circle in the target plane on the plane of camera. Then we can safely draw the conclusion that the track of the target plane formed on the plane of an oval track. Second, the use of graphic tools extracting the track of the edge can help to fit out of elliptic equations, showing the coordinates of the five centers.

Introduction

Digital Cameras positioning has wide range of applications in traffic monitoring (e-police) and the most commonly used method is binocular positioning method, which uses two cameras to locate.

The images of the characteristics point in the objects taken by two cameras fixed in different positions can be used to determine the coordinates in the plane of the cameras. Once the precise relative position of the two cameras is determined, the coordinates of the characteristics point in certain camera can be known geometrically, which means the determination of location of the characteristics point. Thus the key to binocular positioning method is to accurately determine the relative position of two cameras, which is known as system calibration[1].

Take the photos of several points randomly chosen in the flat panel by the two cameras and get the image points of them in the image plane. And the relative position of the cameras can be known through the geometric relationship of the two groups of image point. While the points without geometric size cannot be got in the object plane or the image plane[2]. The practical approach is to draw several circles in the object plane and regard the centers as the geometric points (known as targets). Therefore, once the centers of the circles could be precisely found in the target, the calibration can be achieved.

According to the imaging principle of a digital camera, you can know all light from all the directions will suffer some distortions when traveling through the lens, causing the change of the direction of its propagation direction.

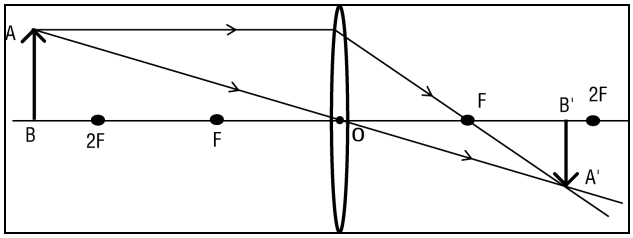


Fig.1 Convex (simple camera) of the imaging light path map

If there is one point on the lens, through which the light in any direction will not suffer any distortion of its propagation direction, that is, the direction of the exit direction and the incident direction are parallel to each other, then the point is called optical lens center, regarded as point O. In order to simplify the process of the solution, when the object distance and image distance are certain, the line is the only light to be studied, and establish three-dimensional space Cartesian coordinate system, in which the optical center is taken as origin[3]. And the image of the independent casual point formed in the image plane is the intersection of the image plane and the line connecting the casual point and the origin point, as shown in the Fig. 1.

Establishment of the model

Establish three-dimensional space Cartesian coordinate system, in which the optical center is taken as origin and z-axis as the main light axis and the xoy plane as the plane in which the image plane is. $Q(X,Y,Z)$: an independent casual point; f :focal length; u： object distance; v： image distance, xoy plane parallel to the image plane.According to the principles of imaging, the optical configuration is established in the coordinate system as shown in Fig.2

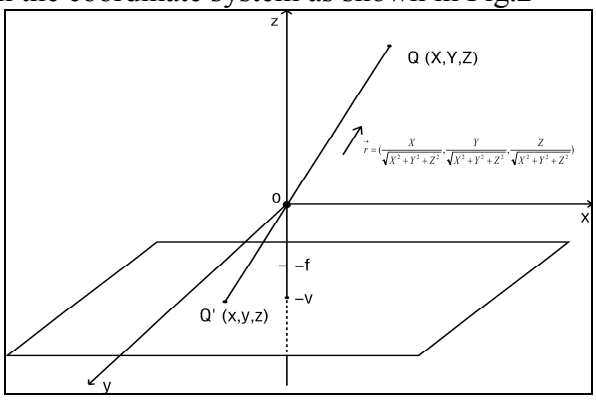


Fig.2

Because the light will not suffer any distortion of its propagation direction, therefore, the direction vector of the straight line traveling through point $(0,0,0)$ and point (X,Y,Z) is

$$\vec{r}=(\frac{X}{\sqrt{X2+Y2+Z2}},\frac{Y}{\sqrt{X2+Y2+Z2}},\frac{Z}{\sqrt{X2+Y2+Z2}})$$

And the parametric equation of straight line QQ' is

$$\frac{x}{X}=\frac{y}{Y}=\frac{z}{Z}=t$$

For the $z=-v$ plane parallel to the xoy plane which the lens plane is in, if $z=-v$ is introduced, we can get:

$$\begin{cases} x=Xt=-X\dfrac{v}{Z} \\ y=Yt=-Y\dfrac{v}{Z} \\ z=-v \end{cases}$$

In accordance to Gaussian imaging formula, we can get $Z=u$, for $\frac{1}{f}=\frac{1}{u}+\frac{1}{v}$, then it can be simplified and we can deduce the image coordinates: $Q'=(-X\frac{v}{u},-Y\frac{v}{u},-v)$

The moment when the camera takes photographs, the distance between objects and camera remains unchanged within exposure time. The distance between the lens and the film is called image distance whose value is constant. Clear pictures can be achieved through automatically adjusting focal length (The camera lens is not a single convex lens but a group lens, which can

make the camera automatically adjust the focal length). Digital camera can even automatically adjust the focal length. And the five points from the circle corresponds to a different object distance[4].

Establish three-dimensional space Cartesian coordinate system, in which z-axis is the main light axis and the xoy plane as the plane in which the image plane is. f :focal length; u : object distance; v : image distance, xoy plane parallel to the image plane. In order to demonstrate the imaging process, the three-dimensional space Cartesian coordinate system is under the condition.

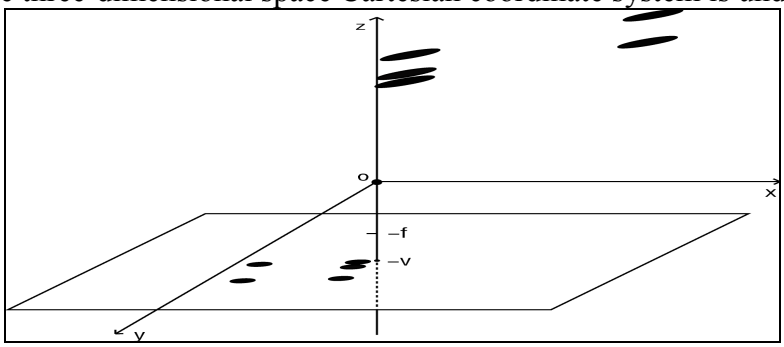


Fig.5 Three-dimensional Space Cartesian coordinate system

Fig.5 demonstrates the5 target original images and their corresponding images in the image plane. Set the corresponding coordinates of the 5 centers (in the target) are:

$O_A(x_a,y_a,z_a),O_B(x_b,y_b,z_b),O_C(x_c,y_c,z_c),O_D(x_d,y_d,z_d),O_E(x_e,y_e,z_e)$

Based on the position relations of the 5 targets in the original images, the center coordinates can be deduced as:

$$O_O(x_O,y_O,z_O)=O_O(\frac{x_a+x_c+x_d+x_e}{4},\frac{y_a+y_c+y_d+y_e}{4},\frac{z_a+z_c+z_d+z_e}{4})$$

Therefore the normal vector of the image is:

$$\vec{n}=(\alpha,\ \beta,\ \gamma)=\begin{vmatrix}\vec{i} & \vec{j} & \vec{k}\\ x_a-x_e & y_a-y_e & z_a-z_e\\ x_c-x_d & y_c-y_d & z_c-z_d\end{vmatrix}$$

The equations in space of the 5 targets in the original image are:

$$\alpha(x-x_O)+\beta(y-y_O)+\gamma(z-z_0)=0$$

And the cosine of angle between the xoy plane and the original image in which the 5 targets are can be deduced as follows:

$$cos\,\theta=\frac{\gamma}{\sqrt{\alpha^2+\beta^2+\gamma^2}}。$$

The image can change with the change of the angle. When the angle is θ, it can match what has been discussed about image situation; therefore, the image coordinates can be figured out.

Set object distance u in the general model ① as (u_A,u_B,u_C,u_D,u_E), then

$$\begin{cases}x_i=-X\dfrac{v}{u_i}\\ y_i=-Y\dfrac{v}{u_i}\\ z_i=-v\end{cases}\qquad (i=A,B,C,D,E,o)$$

Calculation of the image coordinates of the center of the circle in the target plane on the image plane

In order to make the relation between the target and the image consistent with the position relation in which the images in the picture are, the first thing to do is to identify the corresponding image motion equation of motion equation of the circle in the target, and then make image track and its

track in the pictures match with each other. Then the angle θ, target coordinates (x_i, y_i, z_i), and image coordinates $(x'_i, y'_i, -v)$ can be derived.

Identify the curvilinear equation of the circle

Determine the equation of the target in accordance with Title 1: put the target into the space Cartesian coordinate system, and set normal vector: $\vec{n} = (\alpha, \beta, \gamma)$ For the shape of the target is round, the edge trajectory equation is as follow (intercepted by the plane in which the target is, with the sphere):

For the normal vector is $\vec{n} = (\alpha, \beta, \gamma)$, the equation of the plane in which the target locates is

$\alpha(x - x_i) + \beta(y - y_i) + \gamma(z - z_i) = 0$

In accordance to the locus equation parameters of the target circle, it can be known that every point can have its image in the image plane through differentiation, which can help to identify the trajectory equation of the image. The parameter equation of the circle is as follow:

$$\begin{cases} x = R\cos\theta\cos\varphi + x_i \\ y = R\cos\theta\sin\varphi + y_i \left(-\frac{\pi}{2} \le \theta \le -\frac{\pi}{2}\right) \\ z = R\sin\theta + z_i (-\pi \le \varphi \le \pi) \\ \alpha\cos\theta\cos\varphi + \beta\cos\theta\sin\varphi + \gamma\sin\theta = 0 \end{cases}$$

The direction vector of the light is $(R\cos\theta\cos\varphi + x_i,\ R\cos\theta\sin\varphi + y_i,\ R\sin\theta + z_i)$, then

$$\begin{cases} \alpha\cos\theta\cos\varphi + \beta\cos\theta\sin\varphi + \gamma\sin\theta = 0 \\ \dfrac{x}{R\cos\theta\cos\varphi + x_i} = \dfrac{y}{R\cos\theta\sin\varphi + y_i} = \dfrac{z}{R\sin\theta + z_i} = t \\ z = -v \end{cases} \quad \begin{matrix} \left(-\frac{\pi}{2} \le \theta \le -\frac{\pi}{2}\right) \\ (-\pi \le \varphi \le \pi) \end{matrix}$$

Therefore, the parameter equation of the image is as follow:

$$\begin{cases} x = t(R\cos\theta\cos\varphi + x_i) \\ y = t(R\cos\theta\sin\varphi + y_i) \\ z = t(R\sin\theta + z_i) \\ t = -\dfrac{v}{R\sin\theta + z_i} \\ \alpha\cos\theta\cos\varphi + \beta\cos\theta\sin\varphi + \gamma\sin\theta = 0 \end{cases}$$

After simplification, what the parameter equation can confirm is that the track it forms in the image plane is an oval.

Further solution and conclusion

Solving the above we can see through the validation: the graphic plane as elliptical, but also because of the digital photo is comprised of pixels, using canny edge detection method to find graphics on the edge of the discrete trajectory points (call MATLAB Toolbox) (see procedure Annex 2) as shown in Figure 6 results:

From the validation and solution above, it is very clear that the graphic on the image plane is elliptical. And the digital photo is comprised of pixels, so by using the method of canny edge detection, the discrete trajectory points on the edge of the circle can be found (launching the toolbar MATLAB)(Procedure: see Annex 2). In order to make it clear, the author studies the 5 images in the picture as 5 parts separately (see Annex 3). Fit out the relative information of 5 ellipses by the method of least square based on the discrete points on the edge of the circle(see Annex 4).

Table 1: Information Table of Elliptical coordinates (unit: pixels)

ellipse	x	y	a major axis	b stub axle	θ (included angle with the horizontal plane)
A	188.7418	323.0611	44.1380	41.4442	-1.1261
B	196.2783	421.6797	42.9862	41.1858	-1.2216
C	213.2580	639.8139	41.4217	37.0768	-0.5934
D	502.3286	285.0095	41.6357	36.5573	-1.2802
E	502.8296	582.9677	38.2625	32.2437	-0.9970

As the image plane is parallel to xoy plane, and the origin point is the optical center, therefore, the axis which goes through the optical center and is normal to xoy plane is the primary optical axis. The intersection of the primary optical axis and the image plane should be located in the central place （512，384） of camera's sensitive film. Since the camera's image distance(distance between the optical center and the image plane) is 1577 pixel units (1 mm is about 3.78 pixel units), and camera's resolution is 1024 × 768, so when measured in pixels, the image coordinates of the center in space should be converted to $(x_O - 512, y_O - 384, -v)$. In order to unify the units in calculation, the pixels are converted into mm, and the coordinates are $(\frac{x_O - 512}{3.78}, \frac{y_O - 384}{3.78}, \frac{-v}{3.78})$.

Set the pinpointed coordinates of the 5 centers as O_i, and the corresponding results of $(i = A, B, C, D, E)$ are as follow:

O_A(-51.6538,-50.0201,-417.196), O_B(-49.6883,-23.8187,-417.196), O_C(-45.1993,33.7191,-417.196)

O_D(31.31153,-60.0023,-417.196), O_E(31.3636,18.73698,-417.196)

Acknowledgment

This work is supported by the Natural Scientific Fund Project in Hebei Province (No.A2009000735), Scientific Technology Research and Development Plan Project of Tangshan(No.09130202C).The authors are grateful for the anonymous reviewers who made constructive comments.

References

[1] Zhou Fuqiang,Zhang Guangjun,Wei Zhenzhong,Jiang Jie, The targetlocation of two eyes visual sensor of one-mansion based on unknown motion,Mechanical Engineering Journal,Volume 42, Journal 9: 3-4,2006.

[2] Zhou Fuqiang,Hu Kun, Zhang Guangjun, Correction based on linear feature points of Video Camera detortion Mechanical Engineering Journal,2006,Volume 42, Journal 9:174-177

[3] Song Keou，Wang Fengti, *Search for the quick decrease of the target Centroid of two values*,Volume 7,Journal 2：56-62，1994.

[4] SONKA M，HLAVAC V，et a1．Image Processing，Analy—sis and Machine Vision(2nd Edition)[M]，New York：Thommn Press，2002．460—469．

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.468

Research on a Single Camera Location Model and it's Application

Liu Chunfeng, Kong Shanshan, Wu Haiming

Hebei United University College of Sciences Hebei Tangshan 063009,China

Liucf403@163.com

Key words: Transition matrix, Centroid Algorithm, Least-squares Method ,Binocular Calibration

Abstract: Digital cameras have been widely used in the areas of road transportation, railway transportation as well as security system. To address the position of digital camera in these fields this paper proposed a geometry calibration method based on feature point extraction of arbitrary target. Under the meaning of the questions, this paper first defines four kinds of coordinate system, that is the world coordinate system. The camera's optical center of the coordinate system is the camera coordinate system, using the same point in different coordinate system of the coordinate transformation to determine the relationship between world coordinate system and camera coordinate. And thus determine the camera's internal parameters and external parameters, available transformation matrix and translation vector indicated by the camera's internal parameters of the external parameters and the establishment of a single camera location model. According to the model, using the camera's external parameters to be on the target circle center point in the image plane coordinates.

Introdaction

Digital Camera positioning technology has a wide range of applications in traffic monitoring and satellite remote sensing. Digital camera positioning refers to the digital camera shooting pictures of objects to determine some characteristics of a surface location of the point. A feature point on the object, using the photos taken by two cameras fixed at different locations respectively to obtain the point in the two cameras on the image plane coordinates. As long as we know precisely the relative position of two cameras, we can get the geometric method of feature points, a camera at a fixed coordinate system of coordinates, that is characterized by the location of the point. Thus positioning of the eyes to accurately determine the relative position of two cameras is the key, a process known as system calibration.

One of calibration approaches is paint a number of points on a plate at the same time with these two cameras to take photograph, the image points of these points were obtained in their plate. Using these two sets of geometric relations, the relative positions of two cameras can be got. However, whether on the object plane or image plane, we can not directly get any "points" of no geometry size. Practical approach is to draw a number of circles on the object plate (called the target), and the center of the circle is a geometric point. And their images are usually deformed, as shown in Figure 1, we must find the image of the circle center from those images of the cicles in the target accurately, thus calibration can be achieved.

To address the digital camera binocular positioning, we need to include the four main steps:

The first step, give a single digital camera calibration, establish mathematical models to determine the image plane coordinates of the circle center in the cameras under normal circumstances; this key issue is how to select the coordinate system, selection of coordinate system

and establishment of the model and the degree of difficulty of the solution is closely related to each other.

The second step, by the given image and data in the question to derive and find the camera's internal parameters and external parameters, Using the established mathematical model to calculate the coordinates of the image of the circle center in the imaging plane target. The key issue is how to determine the camera's internal parameters and external parameters.

The third step, give a centroid algorithms of image processing, find the image plane coordinates of the center of a circle of targets, and thereby test the model and discuss the algorithm accuracy and stability of this model.

The fourth step, apply the method of least squares to establish a mathematical model with three-dimensional coordinates, using this target to give a mathematical model and methods of a fixed relative position of two cameras. This issue focuses on how to discuss the relationship between the nature of matrix and external function, you can obtain the mathematical model with relative position of the fixed two digital cameras.

The creation of a single camera calibration model

The world coordinate system: $O_w - X_w Y_w Z_w$ in the figure, available point p_w at any point in the space can be shown by (x_w, y_w, z_w). The world coordinate system origin O_w can be seen as the center of the square target surface..

Camera coordinate system: $O_c - X_c Y_c Z_c$ in the diagram, in which O_c is the camera's optical center, z axis and the optical axis are collinear.

The image plane coordinate system: $o - xy$ in the diagram, parallel to the camera coordinate system, its origin o is overlapped with pixel coordinates of the center.

Pixel Coordinates: $o^{'} - uv$ in the diagram.

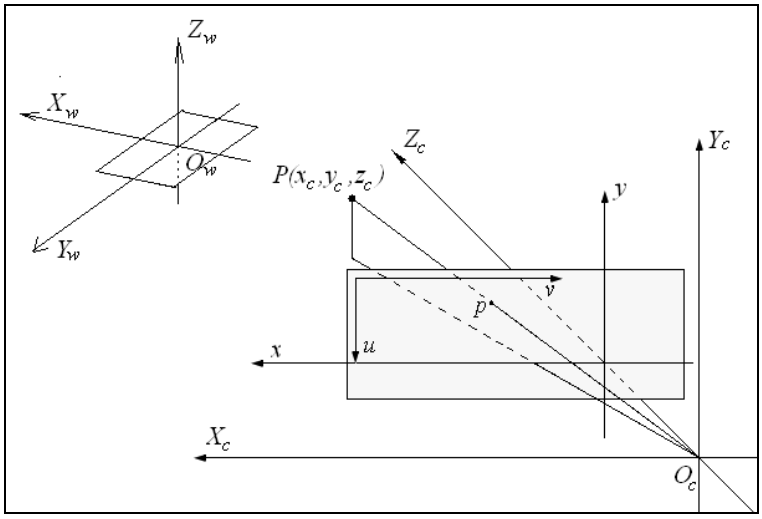


Chart 1: Four kinds of coordinate system-building programs

R is transformation matrix of the world coordinate system and camera coordinate system, T is the translation vector of the world coordinate system and camera coordinate system ,

then $P_c = P_w R + T$. In which $R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix}$, $T = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$,we can get $P_c = (x_c, y_c, z_c)$. In which,

$$\begin{cases} x_c = r_1 x_w + r_2 y_w + r_3 z_w + t_1 \\ y_c = r_4 x_w + r_5 y_w + r_6 z_w + t_2 \\ z_c = r_7 x_w + r_8 y_w + r_9 z_w + t_3 \end{cases}.$$

Assume the image plane coordinates of the ideal image point is (X, Y), the distance from the optical center to the plane is v .

From knowledge of optical phenomena, we can see the object point, image point, optical center are in the same straight line, then the object point and image point is proportional to the corresponding coordinate values (Figure 2), namely, that is, $\frac{x_c}{z_c} = \frac{X}{v}$, and then $X = v\frac{x_c}{z_c}$ same token available, so $Y = v\frac{y_c}{z_c}$, so

$$\begin{cases} X = v\dfrac{r_1 x_w + r_2 y_w + r_3 z_w + t_1}{r_7 x_w + r_8 y_w + r_9 z_w + t_3} \\ Y = v\dfrac{r_4 x_w + r_5 y_w + r_6 z_w + t_2}{r_7 x_w + r_8 y_w + r_9 z_w + t_3} \end{cases}$$

Suppose, the coordinates of image point p in pixel coordinates is (i, j) , that is, the position of image point p in the image is the i row and the j column then. Assume the identifying rate of the camera is $m \times n$, then $X = \frac{1}{3.78}(i - \frac{m}{2})$, $Y = \frac{1}{3.78}(\frac{n}{2} - j)$. As long as to define the camera image distance v, transformation matrix R, translation vector T, we can obtain the image point coordinates of the object point in the image plane of. by the world coordinate system the coordinates of object points.

Solving target on the circle's center in the image on the image plane coordinates

In this problem, from Figure 5 we can easily get target coordinates on the vertical circle $z_w = 0$, so model (1) reduces to

$$\begin{cases} X = v\dfrac{r_1 x_w + r_2 y_w + t_1}{r_7 x_w + r_8 y_w + t_3} \\ Y = v\dfrac{r_4 x_w + r_5 y_w + t_2}{r_7 x_w + r_8 y_w + t_3} \end{cases}$$

The transition matrix R in the column vector is an arbitrary value. given the camera image distance in the problem (ie, the distance between the focus to the imaging plane) is a 1577-pixel units (1 mm is approximately 3.78 pixel units), that is $v = \frac{1577}{3.78} = 417.1958mm$.

The world coordinate system at any point in the pixel coordinate system corresponding pixel coordinates from the pixel coordinate system and the image plane coordinate system transformation between the relationship, to determine pixel coordinates of object points with the relationship between the transformation:

$$\begin{cases} \frac{1}{3.78}(i - \frac{m}{2}) = 417.1958 \frac{r_1 x_w + r_2 y_w + t_1}{r_7 x_w + r_8 y_w + t_3} \\ \frac{1}{3.78}(\frac{n}{2} - j) = 417.1958 \frac{r_4 x_w + r_5 y_w + t_2}{r_7 x_w + r_8 y_w + t_3} \end{cases}$$

In this problem, the substitution of (3) and simplify as follows:

$$\begin{cases} i = 1577 \times \frac{r_1 x_w + r_2 y_w + t_1}{r_7 x_w + r_8 y_w + t_3} + 512 \\ j = 387 - 1577 \times \frac{r_4 x_w + r_5 y_w + t_2}{r_7 x_w + r_8 y_w + t_3} \end{cases}$$

Mapping method we use to determine targets for each circle on the tangent point in pixel coordinates on the pixel coordinates of object points and be known as the point of the pixel coordinates into (4) type, solution obtained transformation matrix and translation vector

$$R = \begin{bmatrix} 1.023165 & 0.1913855 & r_3 \\ 0 & 1.015238 & r_6 \\ 0.6682491 & 0 & r_9 \end{bmatrix}, \quad T = \begin{bmatrix} -16.5975 & 8.968 & 515.0878 \end{bmatrix}^T$$

Target coordinates on the center of each round are as follows:

$$A(-50,50), \quad B(-20,50), C(50,50), D(50,-50), E(-50,-50)$$

Substitution of (3) centered on the various obtained the image of the image plane coordinates:

$$A' = (-50.3983,51.73513), \quad B' = (-22.8603,49.66789), \quad C' = (33.56638,45.43202)$$

Conclusion

Through the interpretation of the subject, it is not difficult to find this is a kind of system calibration problems. Let us set up a single camera calibration model, and then solving them and verify the last set up a model of binocular positioning. Careful analysis of the model we have established is not difficult to find: These models not only for positioning a digital camera, its

calibration of binocular stereo vision problems also played a guiding role.

This model was created to solve the digital camera binocular positions. Through the establishment of a single camera calibration model, we have leads to binocular positioning model.

The solution of this problem is a typical system calibration, we model the use of a very broad, can be applied to the integrity of the workpiece, surface roughness measurement; microelectronic devices (IC chips, PC boards, BGA) and other auto-detection; soft, fragile parts of the testing; a variety of mold detection of three-dimensional shapes; robot vision guidance and so on.

Acknowledgment

This work is supported by the Natural Scientific Fund Project in Hebei Province (No.A2009000735), Scientific Technology Research and Development Plan Project of Tangshan (No.09130202C).The authors are grateful for the anonymous reviewers who made constructive comments.

Referneces

[1] Song Keou，Wang Fengti, Search for the quick decrease of the target Centroid of two values,Volume 7,Journal 2：56-62，1994.

[2] Zhou Fuqiang，Zhang Guangjun，Wei Zhenzhong，Jiang Jie, The targetlocation of two eyes visual sensor of one-mansion based on unknown motion,Mechanical Engineering Journal，Volume 42, Journal 9: 3-42006.

[3] SONKA M，HLAVAC V，et a1．Image Processing，Analy—sis and Machine Vision(2nd Edition)[M]，New York：Thommn Press，2002．460—469．

[4] Zhou Fuqiang，Hu Kun， Zhang Guangjun, Correction based on linear feature points of Video Camera detortion Mechanical Engineering Journal，Volume 42, Journal 9：174-177，2006.

doi:10.4028/www.scientific.net/AMM.50-51.473

Research of Binocular Positioning Mathematical Model Based on the Least Square Method

Liu Chunfeng, Meng Xiaoli, Zhang Huancheng

Hebei United University College of Science Hebei Tangshan 063009,China

Liucf403@163.com

Key words: Transition matrix, Centroid Algorithm, Least-squares Method ,Binocular Calibration

Abstract: By giving a centroid algorithm to extract pixel coordinates of each circle in the center of a circle under the coordinates, through the coordinate transformation the image coordinates can be obtained. Based on the data obtained from above to verify the model, from the specific data of the relative error, absolute error and error propagation theory to discuss the algorithm accuracy and stability. Finally using the least square method, according to the mean square error criteria for the establishment of a minimum of three-dimensional coordinate system fixed relative position of two cameras, binocular positioning mathematical models and by discussing the nature of matrix and the relationship between the relative position of the camera to calibrate the binocular digital camera.

Introduction of Centroid Algorithm

By application of centroid algorithm, the image coordinates of various circles of the target on the plane can be found by computer image processing methods, its algorithm is as follows. First, for the image of the target[1-2].

Set $m \times n$ the value of two-dimensional images (i, j), in which the objective part is A, the background portion of B. Namely:

$$I(i,j) = \begin{cases} 0, (i,j) \in B \\ 1, (i,j) \in A \end{cases}$$

After binarization of the image signal processing, specifications into a 1024 × 768 of the region, will enable target centroid coordinates (K, L).It has been known: m_i is a continuous image pixel points of targets , a_{ij} is starting position of the image pixel row, b_{ij} is endting position of the image pixel row. Of which: $i = 1,2,\cdots,1024$, $j = 1,2,\cdots,768$, $m_{i1} + m_{i2} + \cdots + m_{i\,768} = \sum_{j=1}^{768} m_{ij} = n_i$

let

$$x_1 + x_2 + \cdots + x_{1024} = \sum_{i=1}^{1024} x_i$$

in which

$$\sum_{i=1}^{1024} x_i = \sum (a_{ij} + b_{ij}) m_{ij}$$

Therefore, the column number

$$K = \frac{\sum_{i=1}^{1024} x_i}{2\sum_{i=1}^{1024} n_i} = \frac{\sum_{j=1}^{768} (a_{ij} + b_{ij}) m_{ij}}{2\sum_{i=1}^{1024} (\sum_{j=1}^{768} m_{ij})} = \frac{\sum_{j=1}^{768} (a_{ij} + b_{ij}) m_{ij}}{2\sum_{i=1}^{1024}\sum_{j=1}^{768} m_{ij}} = \frac{\sum_{j=1}^{768} (2a_{ij} + m_{ij}) m_{ij}}{2\sum_{i=1}^{1024} (\sum_{j=1}^{768} m_{ij})}$$

in which $b_{ij} = a_{ij} + m_{ij}$

Line Number

$$L = \frac{n_1 \times 1 + n_2 \times 2 + \cdots + n_{1024} \times 1024}{n_1 + n_2 + \cdots + n_{1024}} = \frac{\sum_{i=1}^{1024} n_i i}{\sum_{i-1}^{1024} n_i}$$

Using the centroid algorithm to test the above model

By the centroid algorithm in the image plane coordinates, image coordinates in the plate of respective circle center can be obtained. The model can be tested, an then calculate the corresponding center point coordinate error[3]. The results show that the absolute error is less than 0.685687, the relative error is less than 0.029121 under both methods given in the coordinates of the center of a circle as the horizontal and vertical coordinates.

Table 1: Model solution value and the value of the centroid algorithm

X-coordinate	model for solving the value	centroid algorithm for solving the value	relative error	absolute error
A'	-50.39828562	-50.02783069	0.007405	0.370455
B'	-22.86034516	-23.54603175	0.029121	0.685687
C'	33.56637735	33.83582011	0.007963	0.269443
D'	19.0091277	18.71240741	0.015857	0.29672
E'	-59.97516382	-60.14134921	0.002763	0.166185
Y-coordinate	model for solving the value	centroid algorithm for solving the value	relative error	absolute error
A'	51.73512504	51.45674603	0.00541	0.278379
B'	49.6678933	49.48616402	0.003672	0.181729
C'	45.43202418	45.19783069	0.005182	0.234193
D'	-31.78946349	-31.4768254	0.009932	0.312638
E'	-31.19983698	-31.15690476	0.001377	0.04293

The centroid algorithm is a simple addition, multiplication, and division operations, in which there is no involution operation, the index operation. From the perspective of the error analysis, the error will not be enlarged, therefore, this algorithm is more stable.

The creation of binocular positioning model

The mathematical model of binocular digital camera model composed by two sets digital cameras is as shown in the figure. The measurement coordinate system $Oxyz$ is based on left digital camera and right digital camera coordinate system is $O_2x_2y_2z_2$. The projection of the left camera coordinate system is $O_1'x_1'y_1'$ and the right digital camera for the projection coordinate system is $O_2'x_2'y_2'$. Assume the space point Q_i is located in the projections of the left and right digital cameras are respectively q_{1i} and q_{2i}. Q_i in the homogeneous coordinates $Oxyz$ is $\tilde{q}_i = [x_i \quad y_i \quad z_i \quad 1]^T$, and in the homogeneous coordinates $O_2x_2y_2z_2$. is $\tilde{q}_{1i} = [x_{1i}' \quad y_{1i}' \quad 1]^T$. The homogeneous coordinate of q_{1i} is $\tilde{q}_{1i} = [x_{1i}' \quad y_{1i}' \quad 1]^T$ and q_{2i}'s is $\tilde{q}_{2i} = [x_{2i}' \quad y_{2i}' \quad 1]^T$.

From the digital camera model, we can get

$$\begin{cases} \tilde{q}_{1i}' = \left[\dfrac{x_i}{z_i} \quad \dfrac{y_i}{z_i} \quad 1 \right]^T \\ \tilde{q}_{2i}' = \left[\dfrac{x_{2i}}{z_{2i}} \quad \dfrac{y_{2i}}{z_{2i}} \quad 1 \right]^T \end{cases}$$

Euclidean transformation between $Oxyz$ and $O_2x_2y_2z_2$, it can be expressed as

$$\lambda \tilde{q}_{2i} = [R \mid t] \tilde{q}_i \qquad \lambda \neq 0$$

Where - 3 × 3 is orthogonal rotation matrix , - 3 × l is translation vector.

By equation (10) and (11) we can see that the relationship Q_{wi} between q_{1i} and q_{2i} are respectively:

$$\lambda_1 \tilde{q}_{1i}' = [I \mid 0] \tilde{q}_i = P_1 \tilde{q}_i \qquad \lambda \neq 0 \,, \lambda_2 \tilde{q}_{2i}' = [R \mid t] \tilde{q}_i = P_2 \tilde{q}_i \qquad \lambda \neq 0$$

Where - 3 × 4 is camera projection matrix of the left 3 × 4 is projection matrix the right digital camera,- 3 × 3 is unit matrix . As shown in Figure 3, we can see Q_i, O and O_2 define a plane, space point in the two projection plane of the corresponding points satisfy the epipolar constraint relationship. Plane Q_iOO_2 is the very flat, l_{1i} is the right projection points, corresponding to the left of the polar projection plane, l_{2i} is for the left-right projection points, corresponding to the

projection of the polar plane. Coplanar epipolar constraint can be expressed as $l_{O_2q_{2i}}(l_{OO_2} \times l_{Oq_{1i}}) = 0$, the corresponding algebra is expressed as: $q'^{T}_{2i} Q q'_{1i} = 0$.

In the formula Q___ the nature of matrix, from the above analysis we can see, Q contains all the parameters, through the decomposition P_2 can be obtained. Then according to the upper equations, and projection coordinates obtained by the two digital cameras, using least-squares method we can calculate the three-dimensional coordinates of space point[4-6].

Solving the relationship between the relative positionof two digital cameras

First, examine the nature of the nature of the matrix.

Theorem 1: If the projection matrix of the left and right digital cameras are respectively $A_1 = (M_1 | -M_1 T_1)$ and $A_2 = (M_2 | -M_2 T_2)$, and then corresponding to the nature of matrix these two digital cameras are $Q \approx (M_2^T)^{-1} M_1 [M_1 (T_2 - T_1)]_x$. From this theorem we know, the essential matrix Q can be decomposed into a full rank matrix R with a symmetric matrix product S, that is $Q \approx RS$. The following theorem states that such a decomposition is almost unique (up to a difference of one scale factor).

Theorem 2: If a 3 × 3 square Q can have two kinds of decomposition method $Q = R_1 S_1 \approx R_2 S_2$, in which S_i is non-zero antisymmetric matrix, R_i is full rank matrix, then there must be $S_1 = S_2$, and if $S_i = [t]_\times$, then $R_2 \approx R_1 + at^T$, a is a three-dimensional vector.

Theorem 3: If the relative position of the left and right digital cameras is fixed, $\{A_1, A_2\}$ and $\{A'_1, A'_2\}$ is two different projection matrices, then, $\{A_1, A_2\}$ and $\{A'_1, A'_2\}$ correspond to the same essential matrix Q if and only if there exists a full rank matrix of 4 × 4's H to make and $P_1 H \approx P'_1$ and $P_2 H \approx P'_2$.

Using the nature of matrix calibration to define a digital camera binocular

The above theorem reflects the fact: the relationship of the relative position of matrix between two digital cameras and the nature is almost equivalent. That is, to set the relative position of the two digital cameras $\hat{R}$ and $\hat{T}$,that is to give the system of matrix Q; On the contrary, given the

nature of matrix Q, two digital cameras, the relationship between the rotation $\hat{R}$ and $\hat{T}$ the translation is almost set up, which is at most have a difference of a constant of proportionality factor. "Almost" means that the identification of $\hat{R}$ and $\hat{T}$ calibration of a scale factor needed k. Reference book (5)gives a convenient method Q by solving, $\hat{R}$ and $\hat{T}$, that is let Q eigenvalue decomposition, solution (SVD), $Q = UDV^T$.

According to theorem 1 in the previous section, we get $R = UGV^T$, $[t]_x = VZV^T$, Among them $G = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $Z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, it is to shift the matrix as a vector of three elements.

So, we have identified the basic form of the relationship between rotation and translation R and T. But, there is a difference of one scale factor between $\hat{R}$ and $\hat{T}$ and R and T, namely, the calibration can make use of the known world coordinate distance in two and then obtain the relationship between the relative position of two cameras through the least-squares method.

Acknowledgment

This work is supported by the Natural Scientific Fund Project in Hebei Province (No.A2009000735), Scientific Technology Research and Development Plan Project of Tangshan(No.09130202C).The authors are grateful for the anonymous reviewers who made constructive comments.

Referneces

[1] HARTLEY R．Estimation of Relative Camera Positions forUncalibrated Cameras l A J．Proceedings of IEEE Conferenceon Computer Vision and Pattern Recognition[C]．[s．1．]：[s．n．]，1992．761．764．

[2] LONGUET—HIGGINS H C．A Computer Algorithm for Re‘constructing a Scene from Two Projections[J]，Nature，293（42）：133—135，1981。

[3] SONKA M，HLAVAC V，et a1．Image Processing，Analy—sis and Machine Vision(2nd Edition)[M]，New York：Thommn Press，2002．460—469．

[4] Zhou Fuqiang，Hu Kun， Zhang Guangjun, Correction based on linear feature points of Video Camera detortion *Mechanical Engineering Journal*，Volume 42, Journal 9：174-177，2006。

[5] Song Keou，Wang Fengti, *Search for the quick decrease of the target Centroid of two values*,Volume 7,Journal 2：56-62，1994.

[6] Zhou Fuqiang，Zhang Guangjun，Wei Zhenzhong，Jiang Jie, The targetlocation of two eyes visual sensor of one-mansion based on unknown motion,*Mechanical Engineering Journal*，Volume 42, Journal 9: 3-42006.

©2010 Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.478

Research and Application on the Fixation Task Model of the Doubled-sight Digital Camera

Yang Aimin[1], Wenjun Zhang[2], Chen Huanhuan[1],Wu Haiming[1]

[1] College of Sciences Hebei United University Hebei Tangshan 063009,China
aimin_heut@163.com
[2] Institute of Applied Computer Technology, China Women's University, Beijing, 100101, China
voicefromzhwj@yahoo.com.cn

Key words: Geometric Model; Method of Abstracting Center of a Circle; Non-line able Equations; Newton Repeating Substitution Method or Newton Raphson Method.

Abstract: According to the principle of the digital camera's forming image, we measure the distance between the points on the target and digital camera respectively, then we construct the limited equation, and calculate the position relation between target and digital camera through this limited equation, then by using the coordinate's transformation, we transform the coordinate which takes digital camera as original points into the coordinate which takes target as original points. After that step, we could fix the coordinate which compare to the target. Finally, we finish the fixation task of the doubled-sight digital camera.

Introduction

Currently, with the spreading and application of the digital camera, the domain of its application expands gradually (taking electronic police for example). Therefore, the problem of relative position between photos taken by digital camera and the original entity turns to be very important.

This paper mainly focuses on the fixation of digital camera and constructs the line able geometric model by using reflecting relation between the coordinate of target's surface and the coordinate of target's image. We take the optical center of digital camera as the original point of coordinate, meanwhile, we also take two straight lines which parallelized among the median line of the target's surface as X axis and Y axis. Then, we construct the space coordinate system by taking optical axis as Z axis. We make use of the method of abstracting the center of a circle to find out image's coordinate of the target's center of a circle which exits on the surface of image. (Unit: mm)

According to the imagery coordinate's mapping relation between the three points'(A,C,E) geometric limited condition on target surface (length, perpendicular) and the image's coordinate on the imagery plane surface to construct non- lineable equation models, we make use of the method of abstracting the center of a circle to get the coordinates of the three points A', C', E'.

In order to improve the precision and decrease the quantity of operation, we take the accelerated Newton repeating substitution method (it known as Newton Raphson method) to evaluate the original numerical value $\vec{x}$ by adding its repeating times to increase the precision of model, then to solve out the numerical value of coordinates of the three points A,C,E on the target's surface, on the basis of this numerical value, we can fix the solution of the target's surface's equation. We also can make use of the geometry's limited qualification to obtain the coordinates of B, D on the target's surface,

moreover, we take advantage of the constructed mathematical model to calculate the coordinates of B' , D' which situate on the target's imagery surface. Then we get the remainder value between the calculation value of coordinates B' , D' and its gauge value , after that, we choose differently combinations and repeat that procedure, then we get a series of that kind of erroneous value. Try to assure the precision of the constructed model through these erroneous values, meanwhile, to make out the concrete analysis about the model's stability.

Firstly, we construct the mapping relation between the coordinate of object's surface and the coordinate of image's surface, testify and deduce the mathematical equation of a circle on target's surface which mapped on the image's surface. Then measure the coordinate of its center of circle according to the geometric characteristics of the mathematical equation[1]. In terms of the fixation of the target's image's coordinate on the image's plane surface, after we construct the space systematic coordinates, then this problem converts into the solution problem of the target's coordinate of its center of circle on image's plane surface. Meanwhile, we construct geometric model according to the geometric relationship between target's plane surface and its image's plane surface internally.

Secondly, it turns the fixation problem of approximately ellipse's center of a circle which situates on the target's imagery plane surface. On the basis of the pixel points we got from the target's image surface, we get the sum figure by using all of the pixel points' coordinate value which located in the approximately ellipse, then we take the average figure of it, the we can obtain the center of a circle of the approximately ellipse, following that way, the technician compile procedure by taking MATLAB method, after $N(N \geq 5)$ times of measurement, let's give up its maximum figure and minimum figure. And let's calculate its average figure, following that method, we can get the coordinate value of the approximately ellipse's center of a circle. Furthermore, the technician could solve out the value of image's coordinate on the target's image's plane surface[2].

However, the doubled-sight fixation of digital camera means that under the situation that we known the relative position of two digital cameras to take photos of the same target by using two digital cameras simultaneously try to record the position at sometime, and to find the intersected-point in an opposite direction which is constructed by the two digital camera's images through the geometric knowledge, then we restore the position of real images.

This paper mainly focuses on the fixation of digital camera and its analysis, to construct the properly space systematic coordinate, and construct the mathematic model by using the mathematic physical optical knowledge, therefore to analyze and get solution of digital camera's fixation problem.

Measuring Distance Model of Doubled-sight Digital Camera

As picture 7 showed, , L , R represent two video cameras, point O represent characteristic point on target, the video camera's focal distance is f' , the distance between two video cameras is b , L_0 is the photo's center taken by left video camera L (we called it as left photo). L_x Is the formed image of point O on the left photo, R_0 is the photo's center taken by right video camera R , (we called it as right photo), R_x is the formed image of point O on right photo. According to triangle relation, it can fix the value of c , then it can calculate the vertical distance between point O and the connecting line of L and R , it is

$$d = \frac{bf}{|x_1 - x_R|}$$

In it, x_l is the photo of the point L_x below,{ J 's lateral coordinate, x_R is the lateral coordinate of point R_x on the right photo }.

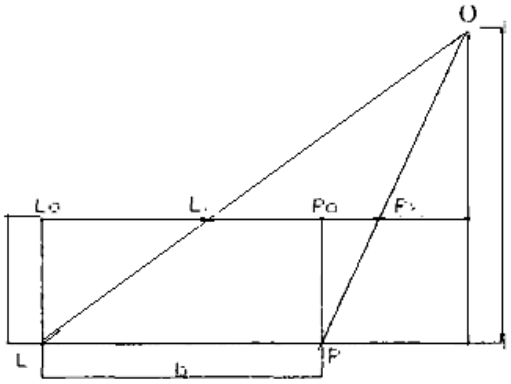


Fig.1 Photo of Doubled-sight Measuring Position

The key point of measuring distance by using doubled-sight stereoscopic vision technique is the choosing and matching of characteristic point (that is point O). The common method of it is choosing a target's shape in one photo as its characteristic point, and to make the match in another photo on the basis of its degree of grey. While, the counting method of matching is very complex, and the amount of calculation is very large, and it requires highly about the quality of taken photos. Otherwise, the probability of making erroneous matching is greatly[3]. To take problems we got above, and consider the concrete target in this research, this paper proposes such method, that is, to adopt the matching of the shape's center to replace the choosing and matching of ordinary characteristic point.

As picture 8 showed, circle 4 is the circle on target. L_0 is the center of left photo, L_x is the shape's center of left photo on target. R_0 is the center of right photo, R_x is the shape's center of right photo on target. Here, we regard L_x and R_x are the formed photo, they are the same point but takes on two differently photos on target respectively. Obviously, if it really exists such point, it must be the center of circle O_1.

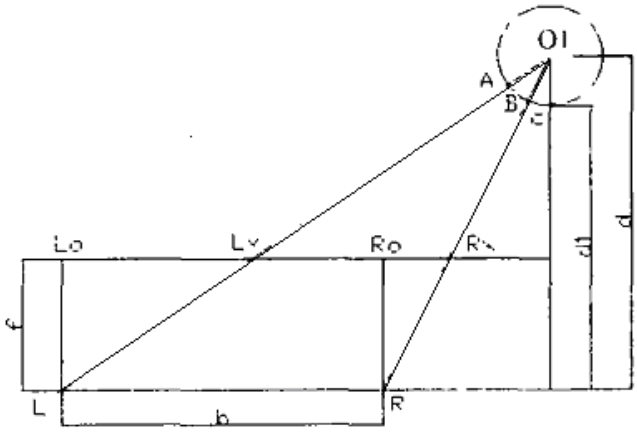


Fig.2 Photo of Measuring Center of a Circle

According to the result of recognizing photo, it is easy to obtain the coordinates of shape's center which belong to the circle in the two photos[4]. Then it can get the distance d between O_1 and the connecting line LR by using equation $(5.4.1)$, but what we really need is the distance d_1 between point C on target and the connecting line LR, then the measuring distance formula should convert into the equation:

$$d_1 = \frac{bf}{|x_1 - x_R|} - c$$

The conversion method of circle's space coordinate

Let's suppose A, B, C are three points in stereoscopic space, it is known that $\overrightarrow{AB} \perp \overrightarrow{BC}$, it can fix the coordinates of A, B, C through lineable equations.

Let's suppose A as original coordinate point, then change its coordinate with its original coordinate, its changing matrix is R, and then it has:

$$\begin{bmatrix} x_A \\ y_A \\ z_A \end{bmatrix} = R \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Therefore, it can obtain the position of digital camera's coordinate which takes A as its original point.

Conclusion

This model has strictly theory evidences, it can measure the distance of digital camera accurately by using target. The characteristic of this model is its strongly practicability. This model can apply into the fixation of digital camera commonly. When it applies into practice, if we know he distance between the object within the two digital cameras, then we can get the concrete coordinate of a third part. It can also apply into all kind of the measuring distances which require lower precision, but because it neglects the influence of the lens to the light, therefore, it may exceed the erroneous domain under the condition that needs highly precision. Besides that, because the amount of model's calculation is very large, so it is better to use the mathematical software.

It is possible to reduce the number of equations by using geometric knowledge to simplify the equations. It is also possible to reduce the quantity of equations and conclusions by using the relation between geometry to combine the equations, and then it is much simpler to calculate the equations. In addition to that, this paper mainly focuses on the fixation of object's position in space by using the doubled-sight measuring method, we can make use of the same theory to measure the space structure

of an object by using many digital cameras, the purpose of this is to restore the real shape of an object. Because it is very complicated to settle such problem, so we don't discuss the problem here.

Acknowledgment

This work is supported by the Natural Scientific Fund Project in Hebei Province (No.A2009000735), Scientific Technology Research and Development Plan Project of Tangshan(No.09130202C).The authors are grateful for the anonymous reviewers who made constructive comments.

References

[1] Song Keou,Wang Fengti, Search for the quick decrease of the target Centroid of two values,Volume 7,Journal 2:56-62,1994.

[2] Zhou Fuqiang,Zhang Guangjun,Wei Zhenzhong,Jiang Jie, The targetlocation of two eyes visual sensor of one-mansion based on unknown motion,Mechanical Engineering Journal,Volume 42, Journal 9: 3-4,2006.

[3] LONGUET—HIGGINS HC. A Computer Algorithm for Reconstructing a Scene from Two Projections[J],Nature,293:133-135,1981

[4] Zhou Fuqiang,Hu Kun, Zhang Guangjun, Correction based on linear feature points of Video Camera detortion Mechanical Engineering Journal,2006,Volume 42, Journal 9:174-177

doi:10.4028/www.scientific.net/AMM.50-51.483

Canny Edge Detection Method and It's Application

Liang Yanbing, Meng Xiaoli, Shujiang An

Hebei United University College of Science Hebei Tangshan 063009,China

snoopy6408665@163.com

Keywords: Multi-Vision Inspecting Technique scale factor, Canny edge detection method, distortion error

Abstract: Digital Cameras positioning has a wide range of application in the aspect of traffic monitoring (e-police).In this paper, the author builds and solves the mathematical model of positioning of monocular by edge detection methods and physical principles of optical imaging of Gauss, and offers a distortion error algorithm to test models, and finally sets up to solve the problem of relative position of multi-cameras. The introduction of distortion error algorithm, could be used to quantitatively examine the models in the first two steps. In accordance with the image situation of multi-image planes, the relative position between the cameras could be determined. This model of camera generates Multi-Vision Inspecting Technique of general distribution of the relative position. Relative position can be figured out if only the parameters of the pictures to be determined are available to determine the inner and outer parameters of the camera.

Introduction

Based on the perspective projection transformation, the track of space ellipse (circle as a special case) in the digital camera plane is elliptic curve. In order to obtain accurate center of space ellipse, the current visual inspection method is to get camera he two-dimensional center of the space ellipse on the plane of digital camera and then calculate the three-dimensional coordinate which is the center of space ellipse based on detection model of visual. And the elliptical center through being fit out the elliptical curve on digital camera's plane is the accurate two-dimensional image point[1].

However, this approach is only an approximate method. And for the inherent characteristics of perspective projection transformation, this method is accurate only in space with an oval cross-section plane parallel to digital camera's plane. In most cases, there are some deviations which are called distortion error. If in practical applications, we substitute the center of being fit out on the image plane for the center of the space ellipse, such introduction of errors in measurement will be inevitable as.

When straight line AD is determined, their corresponding point coordinates can be identified. In accordance with the principle of convex lens imaging, the point of intersection of line GF and the line going through E and the middle point on the line of AD is the point J. The middle point of GF if I, and the line going through I and E intersects with the line AD at C[2]. And $|BC|$ is regarded as the distortion error of point which could also expand the space. The error's level of points can be expressed by the distortion error.

About algorithm error distortion

Define error distortion theoretically and from data:

For this problem, based on the perspective projection transformation and spatial analytic geometry theory, we can deduce three-dimensional visual inspection of the mathematical model of imaging

error-deformation of the center in the oval [3]. Elliptical space imaging camera schematic is as shown in Fig.1.

For the convenience of calculation, general consideration is advised. The position relation of $O_1A_1B_1C_1D_1$ are as follows: one axis: A_1C_1 of $O'A_1B_1C_1D_1$ is orthogonal to Vector $\overrightarrow{O'O_1}$, and an inclined elliptic cone which is through the centre of perspectives O is regarded as $O'A_1B_1C_1D_1$ formed on $O'x'y'z'$, and $OA_1B_1C_1D_1$ on $Oxyz$. Then an isosceles triangular $\Delta A_1O'C_1$ is formed by the visible outline of $O'A_1B_1C_1D_1$. If $\theta = \angle A_1O'C_1 = \angle C_1O'O_1$, $\angle O'O_1D_1 > \frac{\pi}{2}$, then $\phi = \angle D_1O'O_1$; if the angle between Vector $\overrightarrow{O'B_2}$ and the positive half of the x'-axis anticlockwise= φ, then the image plane curve Γ of $O_1A_1B_1C_1D_1$ can be established.

In $Oxyz$, Γ is the line of section between π and $OA_1B_1C_1D_1$. If the equation of π is: $F(x,y,z)=0$, and the equation of $OA_1B_1C_1D_1$ is $G(x,y,z)=0$, then the equation of Γ is as follow.

In the plane $z'=z_0'$, the equation can be expressed as

$$\frac{[(x'\cos\varphi + y'\sin\varphi) - (x_0'\cos\varphi + y_0'\sin\varphi)]^2}{\left|\overrightarrow{B_1O_1}\right|^2} + \frac{[(-x'\sin\varphi + y'\cos\varphi) - (x_0'\sin\varphi + y_0'\cos\varphi)]^2}{\left|\overrightarrow{A_1O_1}\right|^2} = 1$$

For Vector $\overrightarrow{O'O_2}$, such relations are true: $\begin{cases} x_0' = \left|\overrightarrow{O'O_1}\right|\cos\alpha \\ y_0' = \left|\overrightarrow{O'O_1}\right|\cos\beta \\ z_0'' = \left|\overrightarrow{O'O_1}\right|\cos\gamma \end{cases}$, Then $\overrightarrow{O'O_1} = \frac{z_0'}{\cos\gamma}$.

Then $\begin{cases} \overrightarrow{A_1O_1} = \overrightarrow{O'O_1}\tan\theta = \frac{z_0'}{\cos\gamma}\tan\theta \\ \overrightarrow{B_1O_1} = z_0'\tan(\gamma+\phi) - \overrightarrow{O'O_1}\sin\gamma = z_0'\tan(\gamma+\phi) + \frac{z_0'}{\cos\gamma}\sin\gamma \end{cases}$

After simplifying, and set z_0' as the variable, that is to substitute z' for z_0'.then the equation of $O'A_1B_1C_1D_1$ on $O'x'y'z'$ is

$$\frac{[\cos\varphi x'\cos\gamma - z'\cos\alpha + \sin\varphi(y'\cos\gamma - z'\cos\beta)]^2}{[\tan(\gamma+\phi)\cos\gamma - \sin\gamma]^2} + \frac{[-\sin\varphi(x'\cos\gamma - z'\cos\alpha) + \cos\varphi(y'\cos\gamma - z'\cos\beta)]^2}{(\tan\theta)^2} = z'^2$$

From Figure 8, there is revolving relation between $Oxyz$ and $O'x'y'z'$. Established by the following formula:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

If $u = tan(\gamma + \phi)cos\gamma - sin\gamma$, from (6) and (5), the equation of $OA_1B_1C_1D_1$ in $Oxyz$ can be changed into as follows:

$$\begin{aligned}
&\tan^2\theta\{[\cos\varphi(r_1\cos\gamma - r_7\cos\alpha) + \sin\varphi(r_4\cos\gamma - r_7\cos\beta)]x + \\
&[\cos\varphi(r_2\cos\gamma - r_8\cos\alpha) + \sin\varphi(r_5\cos\gamma - r_8\cos\beta)]y + \\
&[\cos\varphi(r_3\cos\gamma - r_9\cos\alpha) + \sin\varphi(r_6\cos\gamma - r_9\cos\beta)]z\}^2 + \\
&u^2\{[-\sin\varphi(r_1\cos\gamma - r_7\cos\alpha) + \cos\varphi(r_4\cos\gamma - r_7\cos\beta)]x + \\
&[-\sin\varphi(r_2\cos\gamma - r_8\cos\alpha) + \cos\varphi(r_5\cos\gamma - r_8\cos\beta)]y + \\
&[-\sin\varphi(r_3\cos\gamma - r_9\cos\alpha) + \sin\varphi(r_6\cos\gamma - r_9\cos\beta)]z\}^2 \\
&= [(ur_7\tan\theta)x + (ur_8\tan\theta)y + (ur_9\tan\theta)z]
\end{aligned}$$

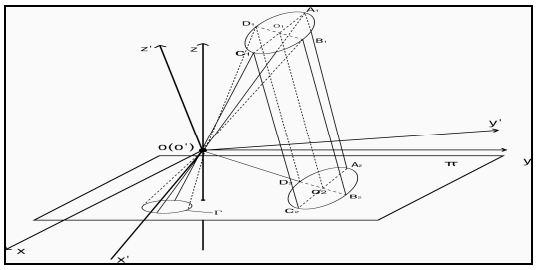


Fig.1 Space ellipses' imaging schematic of camera

Then the equation of perspective projection Γ formed on π by $O_1A_1B_1C_1D_1$ in $Oxyz$ is

$$\begin{aligned}
&(k^2 + n^2 - r^2)x^2 + 2(kl + np - rs)xy + (l^2 + p^2 - s^2)y^2 + 2(mk + nq - rt)x + \\
&2(ml + pq - st)y + m^2 + q^2 - t^2 = 0
\end{aligned}$$

From the second question, we can know Γ is in the form of an ellipse or a circle. Then the coordinates of the center (a,b) of Γ is

$$\begin{cases}
a = \dfrac{[(l^2 + p^2 - s^2)(mk + nq - rt) - (kl + np - rs)(ml + pq - st)]}{[(kl + np - rs)^2 - (k^2 + n^2 - r^2)(l^2 + p^2 - s^2)]} \\
a = \dfrac{[(k^2 + n^2 - r^2)(ml + pq - st) - (kl + np - rs)(mk + nq - rt)]}{[(kl + np - rs)^2 - (k^2 + n^2 - r^2)(l^2 + p^2 - s^2)]}
\end{cases}$$

The model of actual pinpointed coordinates of the center

The actual pinpointed coordinates of the center O_1 of $O_1A_1B_1C_1D_1$ can be identified through the intersection of line OO_1 and π. First establish the equation of OO_1. The intersection between OO_1 and $\pi : z = f$ is the actual image point of O_1[3]. Take $z = f$ into (11), then

$$(r_1\cos\gamma - r_7\cos\alpha)x + (r_2\cos\gamma - r_8\cos\beta)y = (r_9\cos\gamma - r_3\cos\alpha)f$$

$$(r_4\cos\gamma - r_7\cos\alpha)x + (r_5\cos\gamma - r_8\cos\beta)y = (r_9\cos\gamma - r_6\cos\alpha)f$$

Based on (12), the coordinates of actual image point $(a^{'}, b^{'})$ can be expressed as

$$\begin{cases} a' = \dfrac{q_1 l_1 - m_1 p_1}{k_1 p_1 - n_1 l_1} \\ b' = \dfrac{m_1 n_1 - k_1 q_1}{k_1 p_1 - n_1 l_1} \end{cases}$$

In order to verify the error of model in the second part, take one ellipse as the object in the study, the shape of any circle of the five can be identified according to the length of its major and minor axes. Reversibility of optical path can be used to identify the center of the circle and to search the trajectory of the round in the target plane[4].

Get the image coordinate of the center in the image plane and calculate the distance between the image point and elliptical center and then identify the distortion error.

Determination of the camera by the corresponding quadratic curve relation between the target and the image

The information proves that the trajectory of the target is in the shape of round, but according to the second issue, the corresponding trajectory is expected to be an elliptical track. What the paper is going to do next is to determine the camera by the corresponding imaging relation between the target and the image.

First of all, consider the constraint of the intrinsic parameters of digital camera when there is only one plane. Such signals are used to signify certain meanings. $\boldsymbol{m} = [\boldsymbol{u}, \boldsymbol{v}]^{\boldsymbol{T}}$:two-dimensional point in the image;

If the circle of plane curve is located in the target plane, its projection in the plane is ellipse Q, which has been proved, then they can be expressed as : $\begin{cases} \tilde{M}\, O\, \tilde{M} = 0 \\ \tilde{m}\, Q\, \tilde{m} = 0 \end{cases}$, take (2) into the formula, $M^T H^T Q H \tilde{M} = 0$ can be figured out, then $s' C = H^T Q H$ and $C = H^T Q H$.

If $G = U\begin{bmatrix} \sqrt{\hat{\sigma}_i} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} g_{11} & 0 & 0 \\ g_{12} & 0 & 0 \\ g_{13} & 0 & 0 \end{bmatrix}$ then H_{x2}, H_{x3} can also be figured out. After the solution of H, $\begin{cases} h_1^T K^{-T} K^{-1} h_2 = 0 \\ h_1^T K^{-T} K^{-1} h_1 = h_2^T K^{-T} K^{-1} h_2 \end{cases}$ can be known（for $r_1 \perp r_2$）.

The two equations above are two basic constraints of the intrinsic parameters of camera from one homographic matrix. 8 degrees of freedom exist in one homographic matrix, and 6 external parameters (3 rotations and 3 translations). Therefore only 2 constraints of the intrinsic parameters of camera can be known from one homographic matrix. $K^{-T} K^{-1}$ express the projection of the absolute conic on the image plane.

Solution of the parameters

If $B = K^{-1}K^{-T} = \begin{bmatrix} -B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{32} \\ B_{13} & B_{23} & B_{33} \end{bmatrix}$

It can express the Six-dimensional vector $\vec{b} = [B_{11} \quad B_{12} \quad B_{32} \quad B_{13} \quad B_{23} \quad B_{33}]^T$ if the vector of the rank i H is $h_i = [h_{i1} \quad h_{i2} \quad h_{i3}]^T$, then $h_i^T B h_i = v_{ij}^T b$。

Two homogeneous equations with b as the unknown quantity can be identified:

$$\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0$$

Solution and Conclusion

From the simultaneous system of equations about n pictures of the target plane, $vb = 0$ can be known. v is a matrix of $2n \times 6$ .if $n \geq 3$, generally, b can be identified under the condition with the difference of one scale factor. $n = 2$ is what the paper is going to solve. One additional constraint $\gamma = 0$ which is [0 1 0 0 0 0]b=0, is introduced as one additional equation of (a), then the solution of (a) is the corresponding feature vector of the smallest characteristic value of Matrix $v^T v$ .When b is solved, K^{-1} can also be identified by Cholesky decomposition algorithm of matrix, and K can be identified through inversion. The external parameters of every pictures are as follows: $\begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ t \end{bmatrix} = \begin{bmatrix} \mu K^{-1} h_1 \\ \mu K^{-1} h_2 \\ r_1 \times r_2 \\ \mu K^{-1} h_3 \end{bmatrix}$ and $\mu = \frac{1}{\|K^{-1}h_1\|} = \frac{1}{\|K^{-1}h_2\|}$.

Till now, the intrinsic and extrinsic parameters are both identified. The whole process is as follows:

（1）put the target into the space rectangular coordinate system

（2）take pictures from two different positions by two digital cameras(same way as one camera taking two pictures in different positions)

（3）extracting two images

（4）Identify the matrix of the intrinsic and extrinsic parameters of the camera according to the method (identifying the relative position of the two pictures)

The result proves that once the parameters of the characteristic locus can be identified in the two photographic planes, the relative position of the two cameras can be known, for the image is formed inside the camera.

References

[1] Minghua XU, “The Reoffered GMRES(m) Algorithm with Proper Parameters”, *Academy Newspaper of Jiangsu Petrochemical College*, 1999, 11(3): pp. 52-55

[2] Xin Yin, “Three-dimensional Elastic BEM Parallel Calculation and Its Engineering Application”, *Engineering Doctor Degree Papers from Tsinghua University*. 2000: 1-30

[3] Chen T, Chen W. “Frictional Contact Analysis of Multiple Cracks by Incremental Displacement and Resultant Traction Boundary Integral Equations”, *Engn. Anal.Bound.Elem*, 1998, 21:385-392

[4] Xiao H, etc. “Development of Simulator for Rolling In: M.Tanaka and Z.Yao eds. Boundary Element Method”, *Proceedings of the 7th Japan-China Symposium in Boundary Element Methods*, Fukuoka Japan, 1998. Ameatardam: elsevier, 1996:287-296

doi:10.4028/www.scientific.net/AMM.50-51.488

On the dimension of bivariate C^1 cubic spline space with homogeneous boundary conditions over a CT triangulation

Dian-Xuan Gong[1] and Feng-Gong Lang[2,a]

[1]College of Sciences, Hebei United University, Tangshan, Hebei 063009, China
[2]School of Mathematical Sciences, Ocean University of China, Qingdao, Shandong 266100, China

[a]email address: langfg@yahoo.cn, (Corresponding author).

Keywords: bivariate C^1 cubic spline; homogeneous boundary conditions; CT triangulation

Abstract. A bivariate spline is a piecewise polynomial with some smoothness defined on a partition. In this paper, we mainly study the dimensions of bivariate C^1 cubic spline spaces $S_3^{1,0}(\Delta_{CT})$ and $S_3^{1,1}(\Delta_{CT})$ with homogeneous boundary conditions over Δ_{CT} by using interpolating technique, where Δ_{CT} stands for a CT triangulation. The dimensions are related with the numbers of the inter vertices and the singular boundary vertices. The results of this paper can be applied in many fields such as the finite element method for partial differential equation, computer aided design, numerical approximation, and so on.

Introduction

For a simply connected domain $D \subset R^2$, let Δ be a partition of the domain D given by finite irreducible curves, and $D_i (i = 1, 2, \ldots, N)$ the cells. For integers k and μ with $k > \mu \geq 0$, define a bivariate spline space with degree k and smoothness μ over the domain D with respect to the partition Δ as follows([1, 2]):

$$S_k^{\mu}(\Delta) = \{s \in C^{\mu}(D) \mid s|_{D_i} \in P_k, \forall D_i, i = 1, 2, \ldots, N\},$$

where P_k denotes the set of all real bivariate polynomials with total degree k.
Recently, many scholars have studied bivariate spline spaces with homogeneous boundary conditions, defined as:

$$S_k^{\mu,r}(\Delta) = \{s \in S_k^{\mu}(\Delta) \mid \frac{\partial^d s}{\partial n^d}|_{\partial D} \equiv 0, d = 0, 1, \ldots, r\},$$

where ∂D denotes the boundary of the domain D and $r \leq \mu$. Obviously, $S_k^{\mu,r}(\Delta)$ is a linear subspace of $S_k^{\mu}(\Delta)$. The spline in $S_k^{\mu,r}(\Delta)$ is also important and useful in CAGD, plate bending problem, scientific computation, and so on. There are currently many research papers on this topic. In [3] and [4], the bivariate spline spaces $S_3^{1,r}(\Delta_{mn}^{(1)})$ and $S_2^{1,r}(\Delta_{mn}^{(2)})$ were well investigated, where $\Delta_{mn}^{(1)}$ and $\Delta_{mn}^{(2)}$ are uniform type-1 and type-2 triangulations respectively. The bivariate spline space $S_2^{1,r}(\Delta_{mn}^{(2)})$ over a non-uniform type-2 triangulation $\widetilde{\Delta_{mn}^{(2)}}$ was well studied in [5]. While in [6], the dimension of the bivariate spline spaces $S_k^{1,1}(\Delta_{mn}^{(1)})$ with $k \geq 4$, and $S_k^{1,1}(\Delta_{mn}^{(2)})$ with $k \geq 3$ are given. In [7], the dimension and the bases of the so-called Powell-Sabin spline space subjected to homogeneous Dirichlet and Neumann boundary conditions were comprehensively investigated. In [8], Liu et al. determined the dimensions and constructed the locally supported bases for the C^1 cubic spline space and its subspaces with boundary conditions over a non-uniform type-2 triangulation.

In this paper, we are aimed to study bivariate C^1 cubic spline spaces $S_3^{1,0}(\Delta_{CT})$ and $S_3^{1,1}(\Delta_{CT})$ with homogeneous boundary conditions over a CT triangulation Δ_{CT}, where Δ_{CT} (Figure 2) is introduced by Clough and Tocher originally in [9], is a kind of remarkable refinement in the relative literature. For arbitrary triangle in a given triangulation Δ (Figure 1), we divide it into three sub-triangles at any interior point of the triangle. And it has been proved that there exists a unique bivariate C^1 cubic spline on Δ_{CT} (Figure 2) interpolating the function values and the gradient values at the vertices of all triangles, and the normal derivatives at the midpoints of all edges of the triangulation Δ.

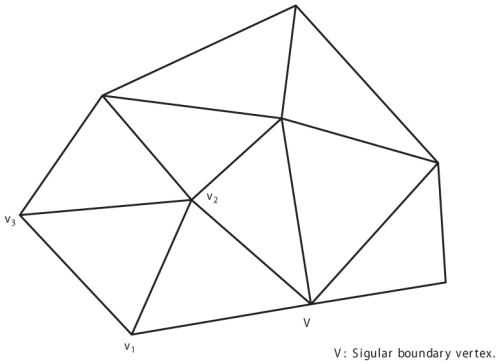


Figure 1. The original triangulation Δ.

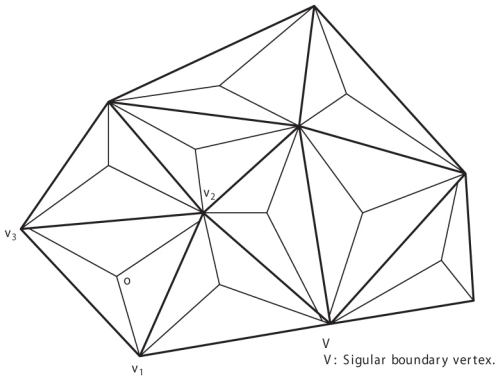


Figure 2. The CT refinement Δ_{CT}.

The remainder of paper is organized as follows: In Section 2, $S_3^1(\Delta_{CT})$ is briefly introduced for the sake of integrity. In Section 3, we mainly study the dimensions of $S_3^{1,0}(\Delta_{CT})$ and $S_3^{1,1}(\Delta_{CT})$ respectively. Finally, this paper concludes in Section 4.

Preliminaries

We begin by introducing the notations which we shall use throughout the paper. For a simply connected polygonal domain D with a regular triangulation Δ, let

E_I : number of the interior edges of Δ, E_B : number of the boundary edges of Δ,
$E = E_I + E_B$: number of the edges of Δ, V_I : number of the interior vertices of Δ,
V_B : number of the boundary vertices of Δ, $\widetilde{V_B}$: number of singular boundary vertices of Δ,
$V = V_I + V_B$: number of the vertices of Δ.

For example, in Figure 2, $V_I = 2$, $\widetilde{V_B} = 1$, $E = 17$ and $E_I = 10$. Let $\mathbf{E_s}$ and $\mathbf{V_s}$ denote the set of the edges and the vertices of Δ, $\mathbf{E_{Is}}$ and $\mathbf{V_{Is}}$ denote the set of the interior edges and the interior vertices of Δ, $\widetilde{\mathbf{V_{Bs}}}$ denotes the set of the singular boundary vertices of Δ, respectively. (See Section 3 for the definition of a singular boundary vertex.)
For a triangle $\triangle(v_1v_2v_3)$, split it into three sub-triangles. For any given $f(x,y) \in C^1(\triangle(v_1v_2v_3))$, there uniquely exists a bivariate C^1 cubic spline $s(x,y)$ satisfies the following conditions

$$s(v_i) = f(v_i),\ \nabla s(v_i) = \nabla f(v_i),\ \tfrac{\partial}{\partial n_i} s(m_i) = \tfrac{\partial}{\partial n_i} f(m_i),$$

where ∇f denotes the gradient of f, $\frac{\partial}{\partial n_i} f(m_i)$ denotes the normal directional derivative of f at the midpoint m_i on the edge $e_i := v_iv_{i+1}$ $(i = 1, 2, 3$ and $v_4 = v_1)$. Similarly, for a CT triangulation Δ_{CT} (Figure 2), we can get a unique bivariate globally C^1 cubic spline interpolates the following conditions ([2, 10]):

$$s(v) = f(v),\ \nabla s(v) = \nabla f(v),\ \tfrac{\partial}{\partial n_e} s(m_e) = \tfrac{\partial}{\partial n_e} f(m_e),$$

where $v \in \mathbf{V_s}$, m_e is the midpoint of the edge $e \in \mathbf{E_s}$, n_e is the normal direction at m_e.

By the unique existence of the interpolation spline, the dimension of $S_3^1(\Delta_{CT})$ is $3V+E$. Its locally supported bases $V_i^j(x,y)$ and $E_k(x,y)$ can be constructed by the following interpolation conditions:

$$\begin{cases} V_i^j(v_l) = \delta_{il}\delta_{j0} \\ \nabla V_i^j(v_l) = \delta_{il}(\delta_{j1}, \delta_{j2}) \\ \frac{\partial}{\partial n_{e_h}} V_i^j(m_{e_h}) = 0 \end{cases} \text{ and } \begin{cases} E_k(v_l) = 0 \\ \nabla E_k(v_l) = (0,0) \\ \frac{\partial}{\partial n_{e_h}} E_k(m_{e_h}) = \delta_{kh} \end{cases}$$

where $i = 1, 2, \cdots, V$; $l = 1, 2, \cdots, V$; $h = 1, 2, \cdots, E$; $k = 1, 2, \cdots, E$; $j = 0, 1, 2$, and $\delta_{\alpha\beta}$ are the Kronecker symbols.

Main Results

Bivariate spline space $S_3^{1,0}(\Delta_{CT})$ From the definition of bivariate spline space with boundary conditions, we have:

$$S_3^{1,0}(\Delta_{CT}) = \{s | s \in S_3^1(\Delta_{CT}), s|_{\partial D} \equiv 0\}.$$

For any spline $s \in S_3^{1,0}(\Delta_{CT}) \subset S_3^1(\Delta_{CT})$, it can be expressed as follows:

$$s(x,y) = \sum_{i=1}^{V} (s(v_i)V_i^0(x,y) + \frac{\partial s(v_i)}{\partial x} V_i^1(x,y) + \frac{\partial s(v_i)}{\partial y} V_i^2(x,y)) + \sum_{h=1}^{E} \frac{\partial s(m_{e_h})}{\partial n_{e_h}} E_h(x,y), \quad (1)$$

where $v_i \in \mathbf{V_s}$ $(i = 1, 2, \cdots, V)$, m_{e_h} is the midpoint of the edge $e_h \in \mathbf{E_s}$ $(h = 1, 2, \cdots, E)$. The restriction of $s(x,y)$ on each edge of Δ_{CT} is a univariate cubic polynomial. In particular, for a boundary edge e_b, let $v_1 := (x_1, y_1)$ and $v_2 := (x_2, y_2)$ be its two endpoints, $\mathbf{e_b} := (x_2 - x_1, y_2 - y_1)$ the directional vector. Suppose that the restriction of $s(x,y)$ on e_b is $s(t)$. So, the necessary and sufficient conditions for $s(t) \equiv 0$ are that the function values and the directional derivatives along $\mathbf{e_b}$ at the endpoints of e_b vanish simultaneously, this implies:

$$\begin{cases} s(v_1) = s(v_2) = 0 \\ \frac{\partial s(v_1)}{\partial x}(x_2 - x_1) + \frac{\partial s(v_1)}{\partial y}(y_2 - y_1) = 0 \\ \frac{\partial s(v_2)}{\partial x}(x_2 - x_1) + \frac{\partial s(v_2)}{\partial y}(y_2 - y_1) = 0 \end{cases}.$$

For a representative boundary vertex $v_0 := (x_0, y_0)$, let $e_L := v_L v_0$ and $e_R := v_0 v_R$ be the two boundary edges emanating from it, where $v_L := (x_L, y_L)$ and $v_R := (x_R, y_R)$ are the left and right neighboring points, respectively. From the above discussion, at each boundary vertex v_0, we have:

$$\begin{cases} s(v_0) = 0 \\ \frac{\partial s(v_0)}{\partial x}(x_0 - x_L) + \frac{\partial s(v_0)}{\partial y}(y_0 - y_L) = 0 \\ \frac{\partial s(v_0)}{\partial x}(x_R - x_0) + \frac{\partial s(v_0)}{\partial y}(y_R - y_0) = 0 \end{cases}. \quad (2)$$

If v_L, v_0 and v_R are not collinear, then we get $s(v_0) = \frac{\partial s(v_0)}{\partial x} = \frac{\partial s(v_0)}{\partial y} = 0$. v_0 is called a non-singular boundary vertex. Conversely, if v_L, v_0 and v_R are collinear, then v_0 is called a singular boundary vertex. The second equation and the third equation in (2) are linearly dependent. Thus, the gradient of s at v_0 does not necessarily vanish. We study this special case particularly. Let $(\cos\theta_0, \sin\theta_0)$ and $(\sin\theta_0, -\cos\theta_0)$ be the unit directional vector and the unit norm vector of the line $\overline{v_L v_0 v_R}$ respectively, we construct two aided splines:

$$F_{v_0}^u(x,y) = \cos\theta_0 V_{v_0}^1(x,y) + \sin\theta_0 V_{v_0}^2(x,y),$$

$$F^n_{v_0}(x,y) = \sin\theta_0 V^1_{v_0}(x,y) - \cos\theta_0 V^2_{v_0}(x,y).$$

It is easy to observe that the directional derivatives of $F^u_{v_0}(x,y)$ and $F^n_{v_0}(x,y)$ along $(\cos\theta_0, \sin\theta_0)$ and $(\sin\theta_0, -\cos\theta_0)$ are $1,0$ and $0,1$ respectively. We have:

$$V^1_{v_0}(x,y) = \cos\theta_0 F^u_{v_0}(x,y) + \sin\theta_0 F^n_{v_0}(x,y),$$

$$V^2_{v_0}(x,y) = \sin\theta_0 F^u_{v_0}(x,y) - \cos\theta_0 F^n_{v_0}(x,y).$$

Submitting into (1), noting that $\frac{\partial s(v_0)}{\partial x}\cos\theta_0 + \frac{\partial s(v_0)}{\partial y}\sin\theta_0 = 0$, we get:

$$\begin{aligned} s(x,y) \;=\; & \sum_{\widetilde{i}=1}^{\widetilde{V_B}} \left(\frac{\partial s(v_{\widetilde{i}})}{\partial x}\sin\theta_{v_{\widetilde{i}}} - \frac{\partial s(v_{\widetilde{i}})}{\partial y}\cos\theta_{v_{\widetilde{i}}}\right) F^n_{v_{\widetilde{i}}}(x,y) + \sum_{h=1}^{E} \frac{\partial s(m_{e_h})}{\partial n_{e_h}} E_h(x,y) \\ & + \sum_{i=1}^{V_I} \left(s(v_i)V^0_i(x,y) + \frac{\partial s(v_i)}{\partial x}V^1_i(x,y) + \frac{\partial s(v_i)}{\partial y}V^2_i(x,y)\right), \end{aligned}$$

where $v_i \in \mathbf{V_{Is}}$, $i = 1,2,\cdots,V_I$; $v_{\widetilde{i}} \in \widetilde{\mathbf{V_{Bs}}}$, $\widetilde{i} = 1,2,\cdots,\widetilde{V_B}$; and m_{e_h} is the midpoint of the edge $e_h \in \mathbf{E_s}$, $h = 1,2,\cdots,E$. Notice that the right-hand splines in the above linear combination are still linearly independent, so we have:

Theorem 1. *The dimension of $S^{1,0}_3(\Delta_{CT})$ is $3V_I + E + \widetilde{V_B}$. $V^j_i(x,y)$, $F^n_{v_{\widetilde{i}}}(x,y)$ and $E_h(x,y)$ are the basis splines, where $i = 1,2,\cdots,V_I$, $j = 0,1,2$, $\widetilde{i} = 1,2,\cdots,\widetilde{V_B}$ and $h = 1,2,\cdots,E$.*

Bivariate spline space $S^{1,1}_3(\Delta_{CT})$ By the definition of bivariate spline space with boundary conditions, we have:

$$S^{1,1}_3(\Delta_{CT}) = \{s | s \in S^{1,0}_3(\Delta_{CT}), \frac{\partial s}{\partial n}|_{\partial D} \equiv 0\}.$$

For an arbitrary spline $s(x,y) \in S^{1,0}_3(\Delta_{CT})$, $\frac{\partial s(x,y)}{\partial n}|_e$ is a univariate quadratic polynomial, where $\frac{\partial s(x,y)}{\partial n}|_e$ denotes the restriction of $\frac{\partial s(x,y)}{\partial n}$ to the boundary edge $e := v_1v_2$. $\frac{\partial s(x,y)}{\partial n}|_e \equiv 0$ is equivalent to

$$\frac{\partial s(v_1)}{\partial n} = \frac{\partial s(v_2)}{\partial n} = \frac{\partial s(m_e)}{\partial n} = 0.$$

Therefore, if $s(x,y)$ also belongs to $\in S^{1,1}_3(\Delta_{CT})$, we conclude that for any boundary vertex v_0 and any boundary edge e, we have:

$$\begin{cases} s(v_0) = \frac{\partial s}{\partial n}(m_e) = 0 \\ \frac{\partial s(v_0)}{\partial x}(x_0 - x_L) + \frac{\partial s(v_0)}{\partial y}(y_0 - y_L) = 0 \\ \frac{\partial s(v_0)}{\partial x}(x_R - x_0) + \frac{\partial s(v_0)}{\partial y}(y_R - y_0) = 0 \\ \frac{\partial s(v_0)}{\partial x}(y_0 - y_L) - \frac{\partial s(v_0)}{\partial y}(x_0 - x_L) = 0 \\ \frac{\partial s(v_0)}{\partial x}(y_R - y_0) - \frac{\partial s(v_0)}{\partial y}(x_R - x_0) = 0 \end{cases}.$$

This implies:

$$s(v_0) = \frac{\partial s}{\partial x}(v_0) = \frac{\partial s}{\partial y}(v_0) = \frac{\partial s}{\partial n}(m_e) = 0.$$

Hence, for an arbitrary spline $s(x,y) \in S_3^{1,1}(\Delta_{CT})$, it can be expressed as follows:

$$s(x,y) = \sum_{i=1}^{V_I}(s(v_i)V_i^0(x,y) + \frac{\partial s(v_i)}{\partial x}V_i^1(x,y) + \frac{\partial s(v_i)}{\partial y}V_i^2(x,y)) + \sum_{h=1}^{E_I}\frac{\partial s(m_{e_h})}{\partial n_{e_h}}E_h(x,y),$$

where $v_i \in \mathbf{V_{Is}}$, $i = 1, 2, \cdots, V_I$; and m_{e_h} is the midpoint of the edge $e_h \in \mathbf{E_{Is}}$, $h = 1, 2, \cdots, E_I$. By the linearly independence of the splines in the above linear combination, we obtain the following theorem:

Theorem 2. *The dimension of* $S_3^{1,1}(\Delta_{CT})$ *is* $3V_I + E_I$. $V_i^j(x,y)$ *and* $E_h(x,y)$ *are the locally supported basis splines, where* $i = 1, 2, \cdots, V_I$, $j = 0, 1, 2$; *and* $h = 1, 2, \cdots, E_I$.

Conclusions

In this paper, we obtain the the dimension of bivariate C^1 cubic spline space with homogeneous boundary conditions over a CT triangulation. The results of this paper can be applied in many fields such as the finite element method for partial differential equation, computer aided geometric design, numerical approximation, and so on.

References

[1] R.H. Wang, The structural characterization and interpolation for multivariate splines, Acta Math. Sinica, 18(1975)91-106; English transl., 18(1975)10-39.

[2] R.H. Wang, *Multivariate Spline Functions and Their Applications*. Science Press/Kluwer, (1994/2001).

[3] C.K. Chui, L.L. Schumaker and R.H. Wang, On spaces of piecewise polynomials with boundary conditions II, Type-1 triangulations, In Second Edmonton Conference on Approximation Theory, CMS Conference Proceedings, Volume 3, 1982, pp. 51-66.

[4] C.K. Chui, L.L. Schumaker and R.H. Wang, On spaces of piecewise polynomials with boundary conditions III, Type-2 triangulations, In Second Edmonton Conference on Approximation Theory, CMS Conference Proceedings, Volume 3, 1982, pp. 67-80.

[5] R.H. Wang and T.X. He, On spline spaces with boundary conditions over nonuniform Type-2 triangulations (in Chinese), Kexue Tongbao, 9(1985)249-251.

[6] A.B. Le, On spaces of bivariate splines with boundary conditions (in Chinese), Math. Numer. Sin., 12(1990)42-46.

[7] R.H. Wang, W.B. Wang, S.M. Wang and X.Q. Shi, The C^1 quadratic spline spaces on triangulations, Math. Appl., (1-2)(1988)123-132.

[8] H.W. Liu, D. Hong and D.Q. Cao, Bivariate C^1 cubic spline space over a nonuniform type-2 triangulation and its subspaces with boundary conditions, Comp. Math. Appl. 49(2005)1853-1865.

[9] R.W.Clough and J.L. Tocher, Finite element stiffness matrics for analysis of plates in bending, *Proceedings of the Conference on Matrix Methods in Structural Mechanics*, Ohio, 1965.

[10] O.C. Zienkiewicz, *The Finite Element Method*. McGraw-Hill, (1977).

Project supported by Educational Commission of Hebei Province of China (No.2009448, Z2010260),
Natural Science Foundation of Hebei Province of China (No.A2009000735,A2010000908).

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.493

A Parallel Evolutionary Algorithm for Job Shop Scheduling*

Ling Wang[1,a], Dianxuan Gong[1,b], Chuan'an Wei[2,c,*] and Shiqiu Zheng[1]

[1] College of Sciences, Hebei United Universtiy, Tangshan 063009, China

[2] Department of Information Technology, Hainan Medical College, Haikou 571101, China

[a]wangl1216@heut.edu.cn, [b]dxgong@heut.edu.cn, [c]weichuanan@yahoo.com.cn

Keywords: job shop scheduling; parallel evolutionary algorithm; neighborhood-based evolutionary

Abstract. The proposed hybrid parallel evolutionary method for Job Shop scheduling with parallel machines was verified by application in a steam turbine factory. The population diversity was improved, and premature elimination of reasonable evolutionary patterns could be avoided. The results indicated that the energy consumption was saved by using this method while the limits of time and costs were satisfied.

Introduction

Current market trends such as variety of consume demands; shorter product life cycles, lower production costs and energy consumption have resulted in the need of rapid responding in production. Effective and accurate production schedule is viewed as a competitive tool by the manufacturing firms to respond to the market nowadays [1]. Therefore many experts focus on the methods for scheduling.

An integrated optimization model of production planning and scheduling for batch production was constructed in [2]. The function objective of the proposed model was to minimize the sum of total setup cost, stock cost and production cost. These ants coordinated with each other by simulating the ant foraging behavior of spreading pheromone on the trails, by which they could make information available globally, and further more guide ants made optimal decisions. An evolutionary tabu search (GTS) was presented in [3]. An effective neighborhood structure was proposed which was based on the famous TSAB algorithm, to achieve the right balance between intensification and diversification searches. A proposition concerning the neighborhood swapping strategy was presented and proved in [4], which could be used to verify whether a neighborhood swapping was acceptable or not. A nonlinear mix integer programming model was presented and solved by a hybrid genetic algorithm in [5]. In the hybrid algorithm, the heuristic rules are used to improve the initial solutions. The proposed algorithm in [6] applied the parallel hybrid architecture of collaborative evolutionary algorithm and genetic algorithm, in which a kind of migration operator was designed to dynamically associate the coevolved subpopulations and the independently evolved common population.

In this paper, parallel hybrid architecture was applied in hybrid collaborative evolutionary algorithm, in which the collaborative evolutionary algorithm and the genetic algorithm were included. A kind of migration operator was designed to dynamically associate the coevolved subpopulations and the independently evolved common population. This gives full play to the role of both algorithms. The best complete solution resulting from the coevolved subpopulations would be reserved and migrated into common publication for further evolution. And this algorithm could avoid premature elimination of proper evolutionary patterns, which was decided by the dynamic property of collaborative evolutionary algorithm. A neighborhood-based evolutionary strategy was also employed to improve the population diversity. Both mutation operator and crossover operator were designed according to the characters of practical problem

* Project supported by Educational Commission of Hebei Province of China (Z2010260), Natural Science Foundation of Hebei Province of China (A2010000908).

* corresponding email：weichuanan@yahoo.com.cn (Chuan'an Wei)

Scheduling Optimization Model in Manufacturing Cell

The problem of Job Shop scheduling with parallel machines considering energy consumption is to determine which machine will be used in machining process, what is the starting time of process and processing sequence, and then make the best scheduling performance under the conditions. Here, scheduling indexes include completion time C_{max}, number of tardiness work pieces n_T and energy consumption ec.

$$C_{\max} = \max\{C_i \mid i = 1,2,...,n\} \tag{1}$$

$$n_T = \sum_{i=1}^{n} U_i \tag{2}$$

$$ec = \sum_{i=1}^{m} ec_l p_l \tag{3}$$

Where p_l is total working time of the machine M_l (subscript l is machine ID).

The scheduling target is to minimize completion time, reduce the number of tardiness work pieces, and decrease energy consumption. The objective function is as follow:

$$\min f = \min\left(C_{\max} + n_T + ec\right) \tag{4}$$

Solving this kind of scheduling problem comprises two decision processes. One is to assign operation $O_{ij}(i=1,2,...,n; j=1,2,...,n_i)$ of work piece $J_i(i=1,2,...,n)$ to a machine belonged to machinery group $\mu_{ij} \in \{M_l^c \mid l=1,2,...W\}$. When $|\mu_{ij}|=1$, the operator O_{ij} will be assigned to the only machine in that group. The other is to decide the processing sequence of the machine $M_l\ (l=1,2,...,m)$, where $m = \sum_{i=1}^{W} |M_l^c|$ is the number of machines

Flow of Algorithm

In general, the population diversity has a great affect on the searching speed of genetic algorithm. The population with higher diversities can avoid the local optimum by producing new evolution patterns successively. Here a neighborhood-based evolutionary strategy was adopted to improve the population diversity. Operation allocation subpopulation P (1) , operation sequencing subpopulation P(2) and common population P(3) were all designed based on a 2D square grid architecture, where $NP[k]_{ij}(k=1,2,3)$ is a 3x3 neighborhood including individual (i, j) and 8 neighbors which belonged to population P (k).

The proposed evolutionary schema includes these steps: First a random evolutionary center (i, j) is selected and the 3x3 neighborhood $NP[k]_{ij}(k=1,2,3)$ constructed; collaborative evolutionary algorithm is used in the neighborhoods of $NP[1]_{ij}$ and $NP[2]_{ij}$. Independently evolution schema is used in the neighborhood of $NP[3]_{ij}$. At last, repeat this process through choosing another evolution center (i, j) until the termination criterion is satisfied. The algorithm flow is as follow:

Step 1: Initialize and generate 3 kinds of population P (k)(1,2,3) , where P(1) is operation allocation subpopulation, P(2) is operation sequencing subpopulation and P(3) is common population.

Step 2: Evaluate the fitness values of the initial individuals in every population. In common population, the evaluation of initial individual can be executed directly to determine the best fitness value. However a symbiotic companion is needed in the evaluation of sub-populations P (1) and P (2). The symbiotic companion is the individual with the same order in the other subpopulation.

Step3: Choose a evolution center (i, j) and construct neighborhood $NP[k]_{ij}(k=1,2,3)$, set q =1.

Step 4: Execute the hybrid collaborative evolutionary based on neighborhood $NP[q]_{ij}$.

Choose parent individuals from neighborhood $NP[q]_{ij}$ by fitness based roulette mechanism. And generate two children individuals according crossover operator. Then replace two individuals of lowest fitness in the neighborhood $NP[q]_{ij}$.

Execute mutation operation on the individuals in the neighborhood $NP[q]_{ij}$ according the mutation probability. Evaluate the fitness values of evolved individual. Meanwhile, the evolved individuals in common population can be evaluated directly. The best fitness value f_{best} will be renewed if the evolved individual has a higher fitness value.

Set q←q+1, where q is the number of populations. Back to step 4 if q ≤3, otherwise step to 5.

Step 5: Replace the individuals of lowest fitness in neighborhood $NP[3]_{ij}$ with the co-evolved individuals of the neighborhood $NP[q]_{ij}(q=1,2)$. That means migrate operation allocation subpopulation and processing sequence subpopulation according the migration operator.

Step 6: Stop the evolutionary procedure if the terminal criteria are satisfied; otherwise back to step 3 and start a new round of evolution.

Algorithm Design

Encoding

Use combinations of machine operation set $\tau_1, S_u=(\tau_1,\tau_2,...,\tau_l,..,\tau_m)$, to indicate the operation allocation solution to the scheduling problem. Here $\tau_l=(l=1,2,...,m)$ is operation set allocated to machine M_l . Use combinations of processing sequences of machine $M_l(l=1,2,\ldots,m)$, $S_v=(\tau_1,\tau_2,...,\tau_l,..,\tau_m)$, to indicate the operation sequence resolution of scheduling problem as well as the final resolution to whole scheduling problem . Where $\pi_l(l=1,2,...,m)$ is the sequence of operations to be executed on the machine M_l in set τ_1 .

Initialization of population

Get the initial solution of operation allocation problem by using load balancing strategy. First, the operations in $\{O_{ij}\,|\,\mu_{ij}=M_l^c;i=1,2,...,n;j=1,2,...,n_i\}$, which need .to be executed on the parallel machines $M_l^c\in\{M_i^c\,|\,|M_i^c|>1;i=1,2,...W\}$, were sorted according to the ascending order of operation time. Then according to the sorted sequence, every operation in the set will be allocated to the least loaded machine in parallel machines respectively. Here, machine load is the sum of operation time of operations to be executed on it. At last, the operation in $\{O_{ij}\,|\,|\mu_{ij}|=1;i=1,2,...,n,j=1,2,...,n_i\}$ would be assigned to the only machine on which the operation can be executed.

Selection

A roulette mechanism based operator was adopted to select the parent individuals for crossover operation. The higher fitness value of the parent individual is, the higher possibility it will be chosen.

Crossover Operation

In crossover operation, arbitrary two random operations in set τ_j and τ_k were chosen respectively, then their operation orders would be exchanged in parent individuals. Where τ_j and τ_k are operation sets of machines $M_j,M_k\in M_l^c$, both belonged to same machinery group $M_l^c\in\{M_i^c\,|\,|M_i^c|>1;i=1,2,...W\}$. In subpopulations P (1) and P (2), crossover operations were performed respectively according to this strategy. In common population P(3) , crossover operations in both parts were performed simultaneously according operation allocation strategy and sequencing schema.

Mutation

In operation allocation subpopulation, arbitrary two operations in set τ_j and τ_k were chosen respectively and the operation in set $\tau_k(\tau_j)$ would mutate to the other set $\tau_k(\tau_j)$.Where, τ_j and τ_k are operation sets of machines $M_j, M_k \in M_l^c$, both belonged to same machinery group $M_l^c \in \left\{ M_i^c \,|\, |M_i^c > 1|; i = 1,2,...W \right\}$. A position-exchanging mutation operator was introduced in sequencing subpopulation. In the process of mutation, first, two random positions in an arbitrary processing sequence $\pi_l (l = 1,2,...,m)$ will be chosen; then their orders will be mutated to each other. Just like that in crossover operation, mutation of individual in common population P (3) was executed in both parts simultaneously.

Migration

A migration operator was proposed here in order to keep the best complete solution individual as the substitute of the worst one in the neighborhood of common publication. One priority of this schema is that the reasonable evolutionary patterns deduced from coevolved subpopulations will be reserved for further evolution. The other is that these complete solution individuals coming from outside of common populations will enhance the population diversity, avoid local optimum, and improve the whole performance of the proposed algorithm.

Fitness Value

Symbiotic companion was introduced because that the individuals of subpopulations were not complete solution. The method for symbiotic companion selection is various and that depends on the practical problem. The easy schema used here is to choose a random individual in the neighborhood. After that, the complete solution, combination of individuals in subpopulations and their symbiotic companion will be substituted in (1) - (4), then objective equation values will get. The objective equation value was used as fitness values in this paper.

Case Analysis

In this case there are 5 manufacturing tasks, 15 work pieces and 6 available machines. Partial data of tasks in a manufacturing unit are showed in Table I . The unit of costs is hundred yuan, the unit of energy consumption is kilogram standard coal, and the unit of time is working day. "—" means the operation cannot be supported by this machine. Scheduling scheme in the manufacturing unit is showed as Fig.1, and data of the manufacturing tasks are showed in Table II. The unit of costs is yuan, and the unit of time is working day.

Table 1 Partial Data of Tasks in A manufacturing Unit

Manufac-turing Task	Workpiece	Due Date	Operation	$M_1/M_2/M_3$			M_4			M_5			M_6		
				C	*EC*	*T*	*C*	*EC*	*T*	*C*	*EC*	*T*	*C*	*EC*	*T*
1	J_1,J_2,J_3	20	1	7	28	7	—	—	—	8	18	6	—	—	—
			2	8	32	8	8	24	8	9	21	7			
2	J_4,J_5	10	1	6	24	6	—	—	—	6	15	5	8	12	4
			2	4	16	4	—	—	—	—	—	—	6	9	3
3	J_6,J_7	27	1	7	28	7	—	—	—	8	18	6	—	—	—
			2	6	24	6	6	18	6	7	15	5	8	12	4
			3	4	16	4	4	12	4	4	9	3	6	9	3
4	J_8,J_9	20	1	7	28	7	7	21	7	8	18	6			
			2	5	20	5	5	15	5	6	15	5	8	12	4
			3	—	—	—	—	—	—	—	—	—	6	9	3
5	$J_{10},J_{11},J_{12},$ J_{13},J_{14},J_{15}	21	1	3	12	3	3	9	3	3	6	2	4	6	2

Tables 2 Data of the Manufacturing Tasks in the Unit

Manufacturing Task	Total Costs	Average Costs	Total Energy Consumption	Average Energy Consumption	Average Energy Consumption in History	Completion Time	Work Time
1	3600	1200	152	50.7	61.3	20	20
2	1860	930	79	39.5	44.5	10	10
3	2980	1490	113	56.5	67.9	27	21
4	2600	1300	93	46.5	53.7	20	20
5	5256	876	224	37.3	43.7	21	11

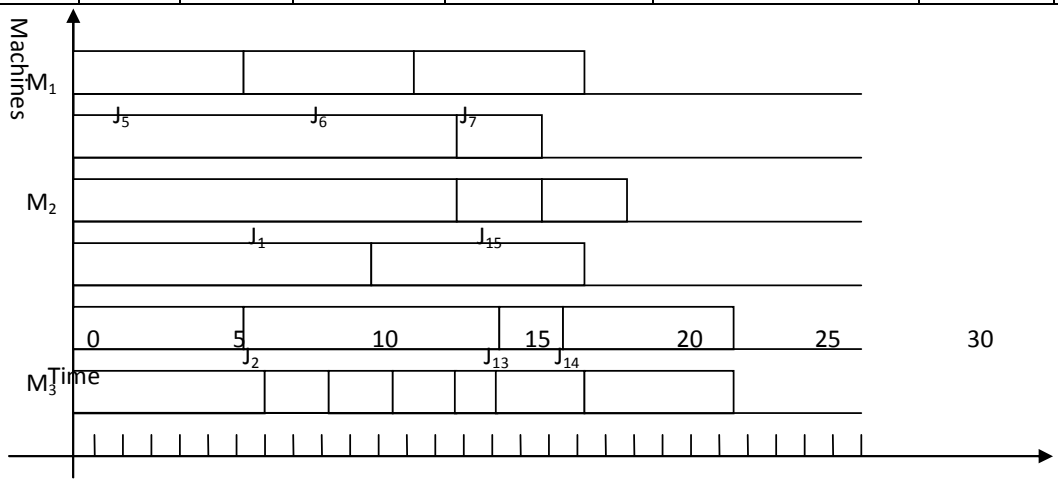


Figure 1 Scheduling scheme in the manufacturing unit

Table II shows that about 16% energy consumption is saved by using this method, at the same time the limits of time and costs are satisfied.

Conclusions

In order to solve the problem of Job Shop scheduling with parallel machines considering energy consumption, an optimization model of Job Shop scheduling with parallel machines was constructed, in which the energy consumption factor was included. The objective function of the model was to minimize the sum of completion time, number of tardiness work pieces and energy consumption. A hybrid collaborative evolutionary method was adopted to design the algorithm of Job Shop scheduling with parallel machines. In the algorithm a neighborhood-based evolutionary strategy was adopted to improve the population diversity. Finally, feasibility and effectiveness of the method are verified by an application case.

References

[1] ZHU Qiong, WU Lihui, ZHANG Jie. “A dynamic Job Shop scheduling method based on ant colony coordination system”, Journal of Donghua University (English Edition), Vol.26, No.1, 2009, pp. 1-4.

[2] SHANG Wenli, FAN Yushun. “Integrated optimization modeling of production planning & scheduling for batch production”, Computer Integrated Manufacturing Systems, Vol.11, No.12, 2005, pp. 1664-1667

[3] Zhang Chaoyong, Gao Liang, Li Xinyu, Shao Xinyu. “Solving Job-Shop scheduling problem using genetic tabu algorithm”, Journal of Huazhong University of Science and Technology (Natural Science Edition), Vol.37, No.8, 2009, pp. 80-84,95

[4] CUI Jianshuang, LI Tieke. “Global neighborhood algorithm for Job Shop scheduling problem”, Computer Integrated Manufacturing Systems, Vol.15, No.7, 2009, pp. 1383-1388.

[5] ZHANG Xiaodong, YAN Hongsen. “Integrated optimization of production planning and scheduling for a kind of Job-Shop”. Control and Decision, Vol.18, No.5, 2003, pp. 581-584

[6] ZHOU Hong, WANG Jian, TAN Xiaowei. “Hybrid collaborative evolutionary algorithm for integrated production planning”, Computer Integrated Manufacturing Systems, Vol.13, No.7, 2007, pp. 1413-1418.

doi:10.4028/www.scientific.net/AMM.50-51.498

Fuzzy Ideal of Hilbert Algebras in BCK-algebras

Qiuna ZHANG[1], Dianchuan JIN[2], Junna JIANG[3]

[1] College of Light Industry, Hebei United University, Hebei Tangshan China 063000

[2,3] College of Science, Hebei United University Hebei Tangshan China 063009

[1]zhangqiuna@yahoo.com.cn, [2]jindianchuan@yahoo.com.cn,[3]jiangjunna131313@126.com

Key words: Hilbert Algebras; BCK-algebras; ideal; fuzzy ideal

Abstract: In [1] it gives an ideal of Hilbert Algebras in BCK-algebras, as well as some propositions. So in this paper, we will give a fuzzy Ideal of Hilbert Algebras in BCK-algebras and it's propositions.

Introduction

The notion of BCK-algebras was formulated first in 1966 by K. Iséki, Japanese Mathematician. This notion is originated from two different ways. One of the motivation is based on set theory, another motivation is from classical and non-classical propositional calculi. The notion of ideals in BCK-algebras was introduced by K. Iséki in 1975.The ideal theory plays a fundamental role in the general development of BCK-algebras, Y. L. Liu and J. Meng discussed fuzzy ideal, fuzzy positive implicative and fuzzy implicative ideal in BCI-algebras. So in this article we will give a fuzzy ideal of Hilbert Algebras in BCK-algebras, and some propositions.

Definition 1.1 Let X be a subset with a binary operation $*$ and a constant 0. Then $(X;*,0)$ is called a BCK-algebra if it satisfies the following conditions:

BCI-1 $((x*y)*(x*z))*(z*y)=0$,

BCI-2 $(x*(x*y))*y=0$,

BCI-3 $x*x=0$,

BCI-4 $x*y=0$ and $y*x=0$ imply $x=y$,

BCK-5 $0*x=0$.

Definition1.2 Suppose H is a nonempty set, $\to$ is a binary operation on H, $1\in H$. Then $(H,\to,1)$ is a Hilbert algebras if it satisfies the following conditions for any x,y,z in H:

H_1 $x\to(y\to x)=1$;

H_2 $(x\to(y\to z))\to((x\to y)\to(x\to z))=1$;

H_3 If $x\to y=1$ and $y\to x=1$, then $x=y$.

Lemma1.1 Suppose $(H,\to,1)$ is a Hilbert algebras, the following conditions are satisfies for any

x, y, z in H :

H_4 $x \to (y \to z) = (x \to y) \to (x \to z)$;

H_5 $x \to (y \to z) = y \to (x \to z)$;

H_6 $(x \to y) \to ((y \to z) \to (x \to z)) = 1$;

H_7 $x \to ((x \to y) \to y) = 1$;

H_8 $x \to x = 1$;

H_9 $1 \to x = 1$.

Let $(X; *, 0)$ be a BCK-algebra, a fuzzy set A in X is a map $A : X \to [0,1]$. If ξ is the family of all fuzzy sets in X, $x_\lambda \in \xi$ is a fuzzy point if and only if $x_\lambda(y) = \lambda$ when $x = y$; and $x_\lambda(y) = 0$ when $x \neq y$. We denote by $\widehat{X} = \{x_\lambda | x \in X, \lambda \in (0,1)\}$ the set of all fuzzy points on X and define a binary operations on $\widehat{X}$ as following: $x_\lambda * y_\mu = (x * y)_{\min(\lambda,\mu)}$

It is easy to verify that $(\widehat{X}, *)$ satisfies the following conditions:

For any $x_\lambda, y_\mu, z_\alpha \in \widehat{X}$

BCK-(1') $((x_\lambda * y_\mu) * (x_\lambda * z_\alpha) * (z_\alpha * y_\mu) = 0_{\min(\lambda,\mu,\alpha)}$

BCK-(2') $(x_\lambda * (x_\lambda * y_\mu)) * y_\mu = 0_{\min(\lambda,\mu,)}$

BCK-(3') $x_\lambda * x_\lambda = 0_\lambda$

BCK-(4') $0_\lambda * y_\mu = 0_{\min(\lambda,\mu,)}$

The condition BCK-5 is not true in $(\widehat{X}, *)$. So the partial order $\leq$ in X can not be extend in $(\widehat{X}, *)$. We can also eatablish the following conditions: for any $x_\lambda, y_\mu, z_\alpha \in \widehat{X}$

(1) $x_\lambda * 0_\mu = x_{\min(\lambda,\mu)}$

(2) $x_\lambda * y_\mu = 0_{\min(\lambda,\mu)}$ and $x_\lambda * z_\alpha = 0_{\min(\lambda,\alpha)}$ imply $y_\mu * z_\alpha = 0_{\min(\mu,\alpha)}$

(3) $x_\lambda * y_\mu = 0_{\min(\lambda,\mu)}$ imply $(x_\lambda * z_\alpha) * (y_\mu * z_\alpha) = 0_{\min(\lambda,\mu,\alpha)}$

(4) $(x_\lambda * y_\mu) * z_\alpha = (x_\lambda * z_\alpha) * y_\mu$

$(5)(x_\lambda * y_\mu) * x_\lambda = 0_{\min(\lambda,\mu)}$

$(6)x_\lambda * (x_\lambda * (x_\lambda * y_\mu)) = x_\lambda * y_\mu$

$(7)(x_\lambda * y_\mu) * z_\alpha = 0_{\min(\lambda,\mu,\alpha)} \text{imply} (x_\lambda * z_\alpha) * y_\mu = 0_{\min(\lambda,\mu,\alpha)}$

$(8)[(x_\lambda * z_\alpha) * (y_\mu * z_\alpha)] * (x_\lambda * y_\mu) = 0_{\min(\lambda,\mu,\alpha)}$

$(9)[((x_\lambda * z_\alpha) * x_\lambda) * (y_\mu * z_\alpha)] * [(x_\lambda * y_\mu) * z_\alpha] = 0_{\min(\lambda,\mu,\alpha)}$

$(10)(x_\lambda * z_\alpha) * [x_\lambda * (x_\lambda * z_\alpha)] = (x_\lambda * z_\alpha) * z_\alpha$

If A is a fuzzy subset of a BCK-algebra X, $\widehat{A}$ is a subset of $\widehat{X}$ then we have the following: $\widehat{A} = \{x_\lambda \mid A(x) \geq \lambda, x_\lambda \in \widehat{X}, \lambda \in (0,1)\}$

For any $\lambda \in (0,1)$ $\widehat{X}_\lambda = \{x_\lambda \mid x_\lambda \in \widehat{X}\}$, $\widehat{A}_\lambda = \{x_\lambda \mid x_\lambda \in \widehat{A}\}$

Fuzzy Ideal and Propositions

Definition2.1 Given a Hilbert algebras $(H, \to, 1)$, a nonempty subset I of H is said to be a Hilbert ideal if it satisfies for all x, y, z in H.

(I) $1 \in I$,

(II) $z \to (y \to x) \in I$ and $z \to y \in I$ imply $z \to x \in I$

Definition2.2 A fuzzy subset A of Hilbert Algebras in BCK-algebras X is a fuzzy subalgebra if and only if for any $x, y \in X$, $A(y \to x) \geq \min(A(x), A(y))$

Definition2.3 A fuzzy subset A of Hilbert Algebras in BCK-algebras X is a fuzzy ideal if and only if:

(1) For any $x \in X, A(1) \geq A(x)$

(2) For any $x, y \in X, A(x) \geq \min(A(y \to x), A(y))$

Definition2.4 $\widehat{A}$ is a weak ideal of $\widehat{X}$ if and only if:

(1) For any $u \in \text{Im}(A), 1_u \in \widehat{A}$;

(2) For any $x_\lambda, y_\mu \in \widehat{X}$ such that $y_\mu \to x_\lambda \in \widehat{A}$ and $y_\mu \in \widehat{A}$, we have $x_{\min\{\lambda,\mu\}} \in \widehat{A}$

Thorem2.1 Let A is a fuzzy subset of a Hilbert Algebras in BCK-algebras X, then the following conditions are equivalent:

(1) A is a fuzzy subalgebra of X;

(2) For any $\lambda \in (0,1]$, $\widehat{A}_\lambda$ is a subalgebra of $\widehat{X}$;

(3) For any $t \in (0,1]$, the t-level subset $A^t = \{x \in X \mid A(x) \geq t\}$ is a subalgebra of X, when $A^t \neq \phi$;

(4) $\widehat{A}$ is a subalgebra of $\widehat{X}$.

Proof (1) $\Rightarrow$ (2) Let $x_\lambda, y_\lambda \in \widehat{A}_\lambda$,because A is a fuzzy subalgebra of X, we have $A(y \to x) \geq \min(A(x), A(y)) \geq \lambda$, hence $(y \to x)_\lambda = y_\lambda \to x_\lambda \in \widehat{A}_\lambda$

(2) $\Rightarrow$ (3) Let $x, y \in A^t$, because $\widehat{A}_\lambda$ is a subalgebra, we have $(y \to x)_t = y_t \to x_t \in \widehat{A}_t$, so $y \to x \in A^t$

(3) $\Rightarrow$ (4) Let $x_\lambda, y_\mu \in \widehat{A}$,and $t = \min(\lambda, \mu)$,since A^t is a subalgebra, we have $y \to x \in \widehat{A}^t$,so $(y \to x)_t = y_t \to x_t \in \widehat{A}$

(4) $\Rightarrow$ (1) Let $x, y \in X$ and $t = \min(A(x), A(y))$,since $\widehat{A}$ is a subalgebra, we have $(y \to x)_t = y_t \to x_t \in \widehat{A}$, hence $A(y \to x) \geq t = \min(A(x), A(y))$

Thorem2.2 Suppose that A is a fuzzy subset of a Hilbert Algebras in BCK-algebras X ,then the following conditions are equivalent:

(1) A is a fuzzy ideal;

(2) For any $x_\lambda, y_\mu \in \widehat{A}, x_\lambda \to (y_\mu \to z_\alpha) = 1_{\min(\lambda,\mu,\alpha)}$ imply $1_{\min(\lambda,\mu,\alpha)} \in \widehat{A}$;

(3) For any $t \in (0,1]$, the t-level subset $A^t = \{x \in X \mid A(x) \geq t\}$ is an ideal when $A^t \neq \phi$

(4) $\widehat{A}$ is a weak ideal.

Proof (1) $\Rightarrow$ (2) Let $x_\lambda, y_\mu \in \widehat{A}$, $x_\lambda \to (y_\mu \to z_\alpha) = 1_{\min(\lambda,\mu,\alpha)}$,Since A is a fuzzy ideal, we have $A(1) \geq A(x) \geq \lambda \geq \min(\lambda, \mu, \alpha)$,so $1_{\min(\lambda,\mu,\alpha)} \in \widehat{A}$

(2) $\Rightarrow$ (3) Since $A^t \neq \phi$,let $x \in A^t$ and $\lambda = A(x)$, $x_\lambda \to (x_\lambda \to 1_\lambda) = 1_\lambda$, then we obtain $1_\lambda \in \widehat{A}$, hence $1 \in A^t$

Let $y \to x \in A^t$,and $y \in A^t$, $y_t \to [(y_t \to x_t) \to x_t] = 1_t$,then we obtain $x_t \in \widehat{A}$ .Hence $x \in A^t$

(3) $\Rightarrow$ (4) we can follow from theorem2.1.

(4) $\Rightarrow$ (1) it is clear that for all $x \in X$, $A(1) \geq A(x)$

$t = \min(A(y \to x), A(y))$, $(y \to x)_t = y_t \to x_t \in \widehat{A}$,and $y_t \in \widehat{A}$,since $\widehat{A}$ is a weak ideal., $x_t \in \widehat{A}$,so $A(x) \geq t = \min(A(y \to x), A(y))$

References

[1] Zhang Qiuna. An Ideal of Hilbert Algebras in BCK-algebras.[J] PROCEEDINGS OF 2009 CONFERENCE ON COMMUNICATION FACULTY:310-311 2009

[2] Jie Meng, Young Bae Jun. BCK-algebras. [M].Seoul: K YUNG MOON SA CO. 1994.：64~71

[3] D. Busneag. Hilbert algebras of fractions and maximal Hilbert algebras of quotients. Kobe [J]. Math. 1988,5:161~172

[4] LiuFang, LiJizu. Hilbert algebras is an anti-positive implicative BCK-algebras (Chinese). [J].Shanxi college of mining and technology. 1997,15(2):214~217

[5] Jie Meng, Jun Y, Kim H. Fuzzy implication ideal of BCK-algebras[J]. Fuzzy Sets and Systems, 1997(89):243~248

[6] Jie Meng, On Ideals in BCK-algebras[J]. Math Japan, 1994(1):143~154

doi:10.4028/www.scientific.net/AMM.50-51.503

Relationships between some Ideals of Hilbert Algebras in BCK-algebras

Qiuna ZHANG[1], Cuilan MI[2], Aimin YANG[3]

[1] College of Light Industry, Hebei United University, China Hebei Tangshan 063009

[2,3] College of Science, Hebei United University China Hebei Tangshan 063009

[1] zhangqiuna@yahoo.com.cn, [2] micuilan@163.com, [3] aimin_heut@163.com

Key words: BCK-algebras; ideal; fuzzy ideal; commutative fuzzy ideal; weak ideal

Abstract: The concept of ideals play an important role in Hilbert Algebras in BCK-algebras, so here we will give some relationships between ideal, fuzzy ideal, commutative fuzzy ideal and weak ideal.

Introduction

A BCK-algebra is an important class of logical algebras and was extensively investigated by several researchers. The concept of fuzzy set was introduced by Zadeh in 1965. The notion of BCK-algebras was formulated first in 1966 by K. Iséki，Japanese Mathematician. Y. B. Jun and E. H. Roh introduced the concept of fuzzy commutative ideals and discussed the fuzzy characteristic commutative ideals in BCK
-algebras. Y. B. Jun and Jie Meng discussed the fuzzy commutative ideals in BCI-algebras. In this paper will give some relationships between ideal, fuzzy ideal, commutative fuzzy ideal and weak ideal.

In this paper, given a Hilbert Algebras in BCK-algebras $(X,\to,1)$ and a fuzzy subset A on X, we construct the set $(\hat{X},\to)$ of all fuzzy points on X and the subset $\hat{A}$ of $\hat{X}$. $\hat{A}=\{x_\lambda \mid A(x)\geq\lambda, x_\lambda\in\hat{X}, \lambda\in(0,1]\}$

Definition1.1 Suppose H is a nonempty set, $\to$ is a binary operation on H, $1\in H$. Then $(H,\to,1)$ is a Hilbert algebras if it satisfies the following conditions for any x,y,z in H:

H_1 $x\to(y\to x)=1$;

H_2 $(x\to(y\to z))\to((x\to y)\to(x\to z))=1$;

H_3 If $x\to y=1$ and $y\to x=1$, then $x=y$

Definition 1.2 Let X be a subset with a binary operation $*$ and a constant 0. Then $(X;*,0)$ is called a BCK-algebra if it satisfies the following conditions:

BCI-1 $((x*y)*(x*z))*(z*y)=0$,

BCI-2 $(x*(x*y))*y=0$,

BCI-3 $x*x=0$,

BCI-4 $x*y=0$ and $y*x=0$ imply $x=y$,

BCK-5 $0*x=0$.

Definition1.3 Given a Hilbert algebras $(H,\rightarrow,1)$, a nonempty subset I of H is said to be a Hilbert ideal if it satisfies for all x,y,z in H.

(I) $1\in I$,

(II) $z\rightarrow(y\rightarrow x)\in I$ and $z\rightarrow y\in I$ imply $z\rightarrow x\in I$.

Definition1.4 A fuzzy subset A of Hilbert Algebras in BCK-algebras X is a fuzzy ideal if and only if:

(1) For any $x,\in X, A(1)\geq A(x)$

(2) For any $x,y\in X, A(x)\geq\min(A(y\rightarrow x),A(y))$

Definition1.5 A non empty subset I of X is called a commutative ideal if it satisfies:

(1) $1\in I$

(2) $z\rightarrow(y\rightarrow x)\in I$ and $z\in I$ imply $[(x\rightarrow y)\rightarrow y]\rightarrow x\in I$

Definition1.6 A fuzzy subset A of a Hilbert Algebras in BCK-algebras X is a fuzzy commutative ideal if and only if:

(1) For any $x\in X, A(1)\geq A(x)$

(2) For any $x,y,z\in X$, $A(((x\rightarrow y)\rightarrow y)\rightarrow x)\geq\min(A(z\rightarrow(y\rightarrow x)),A(z))$

Definition1.7 $\widehat{A}$ is a weak ideal of $\widehat{X}$ if and only if:

(1) For any $u\in \mathrm{Im}(A), 1_u\in\widehat{A}$;

(2) For any $x_\lambda, y_\mu\in\widehat{X}$ such that $y_\mu\rightarrow x_\lambda\in\widehat{A}$ and $y_\mu\in\widehat{A}$, we have $x_{\min\{\lambda,\mu\}}\in\widehat{A}$

Definition1.8 $\widehat{A}$ is a commutative weak ideal of $\widehat{X}$ if and only if:

(1) For any $u\in \mathrm{Im}(A), 1_u\in\widehat{A}$;

(2) For any $x_\lambda, y_\mu, z_\alpha\in\widehat{X}$ such that $z_\alpha\rightarrow(y_\mu\rightarrow x_\lambda)\in\widehat{A}$ and $z_\alpha\in\widehat{A}$, we have

$((x_{\min(\lambda,\mu)}\rightarrow y_\mu)\rightarrow y_\mu)\rightarrow x_{\min(\lambda,\mu)}\in\widehat{A}$

Relationships

Theorem2.1 Let A be a fuzzy subset of a Hilbert Algebras in BCK-algebra X, then the following conditions are equivalent:

(1) A is a fuzzy commutative ideal;

(2) $\widehat{A}$ is a commutative weak ideal.

Proof (1) $\Rightarrow$ (2) Let $\lambda \in \mathrm{Im}(A)$,suppose that $\lambda = A(x)$, since A is a fuzzy commutative ideal, we have $A(1) \geq A(x) = \lambda$,so $1_\lambda \in \widehat{A}$

Let $z_\alpha \to (y_\mu \to x_\lambda) \in \widehat{A}$, and $z_\alpha \in \widehat{A}, A(z \to (y \to x)) \geq \min(\lambda, \mu, \alpha)$ and $A(z) \geq \alpha$, since A is a fuzzy commutative ideal, we have

$$A(((x \to y) \to y) \to x) \geq \min(A(z \to (y \to x)), A(z))$$

$$\geq \min(\min(\lambda, \mu, \alpha), \alpha) = \min(\lambda, \mu, \alpha)$$

So $(((x \to y) \to y) \to x)_{\min(\lambda,\mu,\alpha)} = ((x_{\min(\lambda,\alpha)} \to y_\mu) \to y_\mu) \to x_{\min(\lambda,\alpha)} \in \widehat{A}$

(2) $\Rightarrow$ (1) Let $x \in X$ and $\lambda = A(x)$, $\lambda \in \mathrm{Im}(A)$. Since $\widehat{A}$ is a commutative weak ideal, we have $1_\lambda \in \widehat{A}$, so $A(1) \geq \lambda = A(x)$

If $x, y, z \in X$, let $A(z) \geq \alpha$ and $A(z) = \alpha$,

$(z \to (y \to x))_{\min(\beta,\alpha)} = z_\alpha \to (y_\alpha \to x_\beta) \in \widehat{A}$, and $z_\alpha \in \widehat{A}$.

Since $\widehat{A}$ is a commutative weak ideal, we have

$((x_{\min(\beta,\alpha)} \to y_\alpha) \to y_\alpha) \to x_{\min(\beta,\alpha)} = (((x \to y) \to y) \to x)_{\min(\alpha,\beta)} \in \widehat{A}$

So $A(((x \to y) \to y) \to x) \geq \min(\alpha, \beta) = \min(A((z \to (y \to x)), A(z))$

Theorem2.2 A commutative weak ideal must be a ideal, but the converse does not hold in general.

Theorem2.3 A fuzzy commutative ideal must be a fuzzy ideal, but the converse does not hold in general.

Theorem2.4 Suppose that $\widehat{A}$ is a weak ideal, then the following conditions are equivalent:

(1) A is commutative;

(2) For all $x_\lambda, y_\mu \in \widehat{X}$ such that $y_\mu \to x_\lambda \in \widehat{A}$, we have

$((x_{\min(\lambda,\mu)} \to y_\mu) \to y_\mu) \to x_{\min(\lambda,\mu)} \in \widehat{A}$

(3) For any $t \in (0,1]$, the t-level subset $A^t = \{x \in X \mid A(x) \geq t\}$ is a commutative ideal when $A^t \neq \phi$;

(4) $\widehat{A}$ is commutative.

Proof (1) $\Rightarrow$ (2) Let $y_\mu \to x_\lambda \in \widehat{A}$, since A is a fuzzy commutative ideal, we have

$A(((x \to y) \to y) \to x) \geq \min(A((y \to x) \to (y \to x)), A(y \to x))$

$\geq \min(A(1), A(y \to x)) = A(y \to x)) \geq \min(\lambda, \mu)$

So $(((x \to y) \to y) \to x)_{\min(\lambda,\mu)} = ((x_{\min(\lambda,\mu)} \to y_\mu) \to y_\mu) \to x_{\min(\lambda,\mu)} \in \widehat{A}$

(2) $\Rightarrow$ (3) It is clear that $1 \in A^t$

Let $z \to (y \to x) \in A^t$ and $z \in A^t$, we have

$(z \to (y \to x))_t = z_t \to (y_t \to x_t) \in A^t$ and $z_t \in A^t$. Since $\widehat{A}$ is a weak ideal, we have

$(y_t \to x_t) = (y \to x)_t \in \widehat{A}$, then

$(((x_t \to y_t) \to y_t) \to x_t) = (((x \to y) \to y) \to x)_t \in \widehat{A}$

So $(((x_t \to y_t) \to y_t) \to x_t) \in \widehat{A}$

(3) $\Rightarrow$ (4) Let $\lambda \in \mathrm{Im}(A)$,it is clear that $1_\lambda \in \widehat{A}$

If $z_\alpha \to (y_\mu \to x_\lambda) \in \widehat{A}$ and $z_\alpha \in \widehat{A}$. Let $t = \min(\lambda, \alpha, \mu)$, then

$z \to (y \to x) \in A^t$ and $z \in A^t$. Since A^t is a commutative weak ideal, we obtain

$((x \to y) \to y) \to x \in \widehat{A}$.

So $(((x \to y) \to y) \to x)_t = (((x \to y) \to y) \to x)_{\min(\lambda,\mu,\alpha)}$

$$((x_{\min(\lambda,\alpha)} \to y_\mu) \to y_\mu) \to x_{\min(\lambda,\alpha)} \in \widehat{A}$$

(4) $\Rightarrow$ (1) follows from theorem2.1

References

[1] Zhang Qiu na. An Ideal of Hilbert Algebras in BCK-algebras.[J] PROCEEDINGS OF 2009 CONFERENCE ON COMMUNICATION FACULTY:2009:310-311

[2] Jun Y B. Fuzzy commutative ideal of BCK-algebras[J] Fuzzy Sets and Systems, 1994(63):1~26

[3] Jie Meng, On Ideals in BCK-algebras[J]. Math Japan, 1994(1):143~154

[4] Jie Meng, Jun Y, Kim H. Fuzzy implication ideal of BCK-algebras[J]. Fuzzy Sets and Systems, 1997(89):243~248

[5] Zadeh L. Fuzzy Sets[J]. Information and Control 1965(36):936~942

[6] Xi O G. Fuzzy BCK-algebras[J]. Math Japan, 1991(36):338~353

Part 2

© (2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.511

Analysis on the Dynamic Responses of the Broken Conductors in Transmission Lines

Fengli Yang [1, a], Jingbo Yang [2, b]

[1]China Electric Power Research Institute, Beijing, China

[2]China Electric Power Research Institute, Beijing, China

[a]yangfl1@epri.sgcc.com.cn, [b]yjb@epri.sgcc.com.cn

Keywords: transmission line, broken conductor, dynamic response, damping, phase spacer

Abstract: Impact loads from the broken conductors are common for transmission lines, which can bring threaten to the safe operation of the transmission lines. Dynamic analysis of the conductors in transmission lines under broken load was carried out. A finite element model of seven span conductors in transmission line was established in general software ANSYS. The insulator and the phase spacer were considered in the FEA model. The broken load case can be realized by the birth-death element method in ANSYS. Stiffness of the broken conductor or insulator element was changed to be a near zero value in a very short time. Effect of the damping property of the conductors was considered by the Rayleigh damping method. Dynamic responses of displacements at the broken points and the reaction forces of the insulators were obtained. Dynamic responses for the broken conductors with different damping ratios and bundle numbers were compared.

Introduction

Conductors rupture is common in transmission lines. Conductors will be broken due to the heavy accreted ice, fatigue or intensive galloping. Transient tension of conductors will vary intensively and the conductors will jump severely induced by the broken conductors, which can cause severe electrical or mechanical accidents of transmission lines.

M.B.Thomas et al. [1] developed an effective program for calculating dynamic response of ruptured conductors in transmission lines. Peak broken conductor loads on supporting towers can be predicted. Alian H.Peyrot et al. [2] studied on the longitudinal loads on a suspension tower and the dynamic response of a transmission line following a broken conductor. Predicted values of the peak tension force were compared to the recorded test values and published data. Damping properties of the conductors were not considered in the above broken analysis.

In this paper, a three dimensional FEA model of conductor-insulator system in transmission line was established by ANSYS. The phase spacer and the damping of the conductors were considered in the FEA model. The rupture load case of insulator or conductor can be realized by the birth-death element method in ANSYS. Dynamic responses of displacements at the broken points and the reaction forces at the end of insulators were obtained.

FEA Model of the Conductor-Insulator System

FEA model of a continuous seven-span conductor-insulator system was shown in Fig.1. The longitudinal flexibility of the towers is not modeled and assumed to be rigidly fixed at the end of insulators. The end points of the model were fixed to simulate the strengthen towers or the dead-end towers.

In the analysis model shown in Fig.1, the first four span conductors (①-④) and the last two span conductors (⑥-⑦) have the same span length. Span ④ and span ⑤ has a span length difference. The first four span conductors have an elevation difference with the last three span conductors.

Diameter of the insulator is 30mm and its length is 5.4m, and the final elastic modulus is 51.83GPa. The type of the conductors is LGJ400/50，with a total cross-sectional area of 451.55mm^2 and a

self-weight of 1.51kg/m for single bundle conductor. The final elastic modulus of the conductor is 69.00GPa and its maximum rated tensile strength is 117.23kN. The safety coefficient for the design requirement of the conductors is 2.50. The maximum working tension force is 46.89kN.

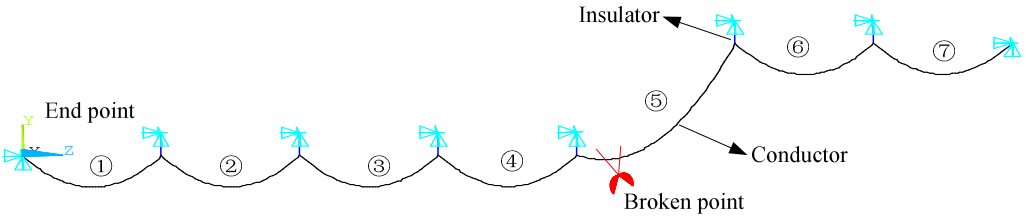


Fig.1. FEA model of conductor-insulator system

Conductors Modeling

The conductors were simulated by element Link10. Link10 is a three dimensional spar element with the unique feature of a bilinear stiffness matrix resulting in a uniaxial tension-only (or compression-only) element. With the tension-only option, the stiffness is removed if the element goes into compression. This feature is useful for static guy-wire applications where the entire guy wire is modeled with one element.

In the conductor-insulator system FEA model, four bundles conductor can be simplified to a single bundle with the equivalent area of cross section and stiffness. The element length of the conductor is 10 meters. The initial tension of the conductor can be realized by applying initial strain or decreasing the temperature of the elements, and the latter method was used for the analysis. A very small initial tensile strain (such as 10^{-8}) was specified for each element of conductors in order to avoid singularities in the initial stiffness matrix of the cables.

Before the broken conductor analysis, the form-finding analysis of conductors under self-weight load was carried out. Firstly, the initial tensions of the conductors can be determined according to the representative span length [3, 4] of the typical transmission line section. Then the parabolic geometry models of the conductors were obtained on the basis of tension-sag formula. The geometry forms of the conductors were updated repetitively through nonlinear static analysis, and the form-finding process [5-7] will be terminated when the deformations of the conductors were closed to zero.

Insulator Modeling

The insulators were also simulated by element Link10. The insulator is assumed to be made of elastic material. Self-weight of the insulator was ignored in the analysis. The loads and displacements at the end of the insulator were mainly focused in the rupture analysis, so one single element was divided for each insulator.

Damping

Determination of the conductor damping is a very complex problem. In the broken analysis, for the quick convergence of the transient solution, damping of the conductors was considered by Rayleigh damping ratio.

In reference [8], based on the field measurements, the natural vibration frequencies and damping ratios of the first two lateral vibration modes were identified by Hilbert-Huang transform. The first order damping ratio of the tower-line system is 9.93%, and the second order damping ratio is 21.92%. In this paper, damping ratios of the conductors are set from 5% and 20%.

Dynamic Responses of the Broken Conductors

For the broken conductors load case, FEA model of a seven-span conductor-insulator system was established. The lengths of different spans are all equal to 500m, except the length of the fifth span is 575m which has a length difference of 15 percent with other spans. The first four span conductors have a 15 percent elevation difference with the last three spans. As presented in Fig.1, The fifth span conductor near the fourth insulator will be ruptured in the broken analysis. As the peak dynamic responses were focused in this analysis, contact between the broken conductor and the ground was

ignored. The detailed solution process is as follows.

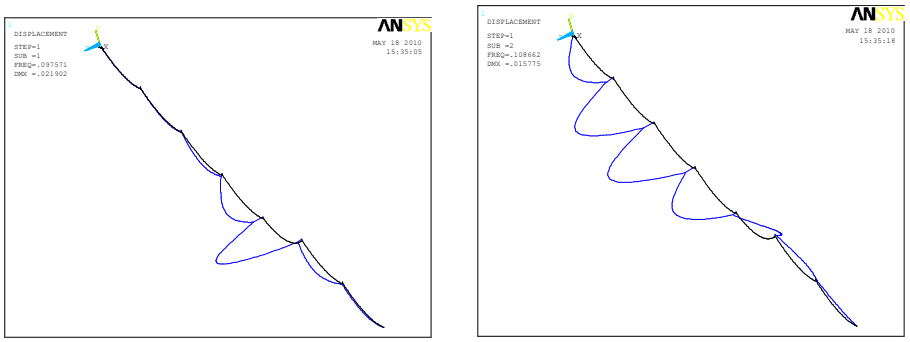
Firstly, modal analysis of the conductor-insulator system was completed by Block Lanczos method. The first order and the second order natural frequency were obtained. So the Rayleigh damping ratios can be calculated for the broken analysis. Secondly, rupture of the insulator or the conductor can be realized by birth-death element method in ANSYS with *ekill* command. Stiffness of the broken insulator or conductor element was changed to be a near zero value in a very short time. Lastly, dynamic responses of displacements at the broken points, tension forces of the conductors and the reaction forces at the end of the insulators were analyzed. Effect of the bundle numbers of the conductors and the broken location were considered in the analysis.

Effect of Damping Ratios

Rayleigh damping can be calculated by the following equations.

$$C = \alpha M + \beta K \tag{1}$$

where $\alpha = \dfrac{2\omega_1\omega_2(\xi_1\omega_2 - \xi_2\omega_1)}{\omega_2^2 - \omega_1^2}$, $\beta = \dfrac{2(\xi_2\omega_2 - \xi_1\omega_1)}{\omega_2^2 - \omega_1^2}$. ξ_1 and ξ_2 are the first order damping ratio and the second order damping ratio. M and K are the mass matrix and the stiffness matrix. The symbols of α and β represent the mass damping coefficient and the stiffness damping coefficient, which can be applied by *Alphad* and *Betad* command in ANSYS. The first mode and the second mode are presented in Fig.2. The first natural frequency and the second natural frequency are 0.098Hz and 0.109Hz.



(a) The first order (b) The second order

Fig.2. Vibration modes

For different values of the first order damping ratio and the second order damping ratio, the values of α and β were calculated by Eq.1. Then the transient analyses for the broken conductors with different damping ratios were carried out. The solving process continues 20 seconds. For the geometrical nonlinear transient analysis, the stiffness of the conductor will be strongly varied in the broken process. As shown in Table 1, the dynamic responses vary greatly with the variation of the stiffness damping ratio. ΔT_d and V_d represent the unbalanced tension and the vertical load at the end of insulator. Numbers with bracket in Table 1 indicate the location of the maximum loads at the end of the insulators.

Table 1 Dynamic responses from the broken insulator

ξ_1	ξ_2	α	β	Dropping height[m]	ΔT_d [kN]	V_d [kN]
2%	2%	0.0021	0.194	495.1	135.5(4)	-90.1(5)
5%	5%	0.0051	0.485	495.6	215.7(4)	-82.0(5)
5%	10%	-0.040	5.235	467.5	58.8(4)	-59.6(5)
10%	10%	0.0103	0.970	496.0	152.9(4)	-75.1(5)
10%	20%	-0.080	10.470	296.0	58.5(4)	-53.6(5)

Effect of Bundle Numbers

Dynamic responses from broken conductors by one equivalent model and four bundles conductors were compared. FEA model of the four bundles conductors with a partial view was shown in Fig.3. The vertical and longitudinal distances of the conductors are 450mm. There are 8, 8, 8, 8, 9, 8, 8 spacers for the seven conductors in turn. The spacer was simulated by Beam4 element in ANSYS. The initial tension force of single bundle conductor is 22.03kN.

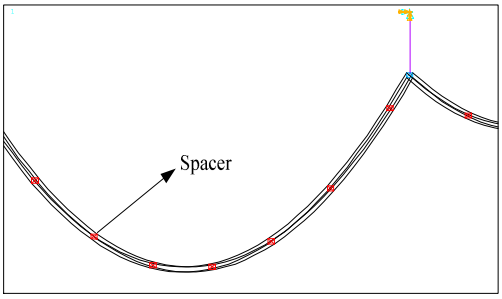


Fig.3. FEA model of the four bundle conductors with a partial view

Firstly, it is assumed that the broken location at the fourth insulator and four bundle conductors were broken at the same time. The first order damping ratio and the second damping ratio are valued 10% and 20% separately. The solve time is 120 seconds. Dynamic responses by three different models were shown in Table 2. The dropping height refers to the maximum dropping displacement of the broken point. The maximum value of ΔT_d and V_d occur at the third and the fifth insulator respectively. From the results in Table 2, for the rupture of the insulator, dynamic responses by one equivalent bundle model are very close to those by four bundles model. Phase spacer has little effect on the dynamic response with rupture of the insulator.

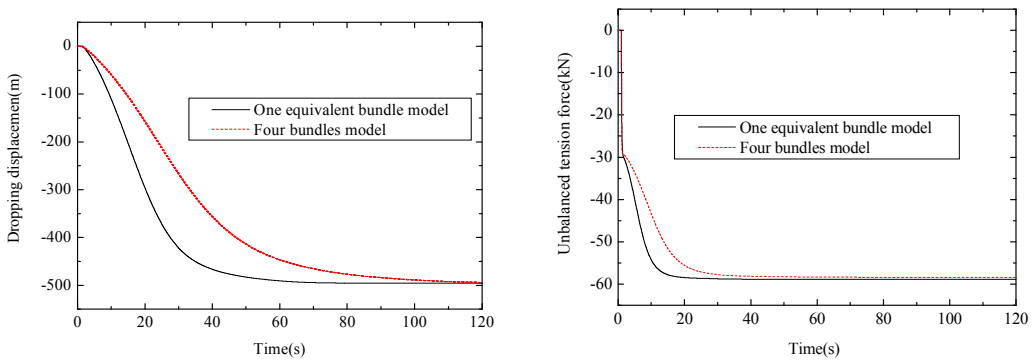
Table 2 Dynamic responses from the broken insulator

Model type	Dropping height[m]	ΔT_d [kN]	V_d [kN]
One equivalent bundle model	34.1	16.14(5)	-57.12(5)
Four bundles model(with phase spacer)	34.2	15.85(5)	-56.48(5)
Four bundles model(without phase spacer)	21.6	15.28(5)	-57.49(5)

Secondly, it is assumed that the broken location at fifth span conductor near the fourth insulator and four bundle conductors were broken at the same time. As shown in Table 3 and Fig.4, for the rupture of the conductor, dynamic responses by one equivalent bundle model are very close to those by four bundles model.

Table 3 Dynamic responses from the broken conductor

Model type	Dropping height[m]	ΔT_d [kN]	V_d [kN]
One equivalent bundle model	495.6	64.8(5)	-53.6(5)
Four bundles model	494.0	63.4(5)	-52.6(5)



(a) Dropping displacement (b) Unbalanced tension force

Fig.4. Comparison of the dynamic responses

Conclusions

In this paper, a three dimensional finite element model of conductor-insulator system was established in ANSYS, and the dynamic responses of from the broken insulator and conductor cases were analyzed for a seven continuous conductor-insulator system. For different analysis model, effects of the damping, bundles of the conductors as well as the phase spacer on the dynamic responses were studied. Some conclusions were drawn from the analysis results.

(1) Damping of the conductors has great effect on the dynamic responses from the broken conductors. For the geometrical nonlinear transient analysis, the stiffness of the conductor will be strongly varied in the broken process. The dynamic responses vary greatly with the variation of the stiffness damping ratio.

(2) For the rupture of the insulator, dynamic responses by one equivalent bundle model are very close to those by four bundles model. Phase spacer has little effect on the dynamic response with rupture of the insulator. For the rupture of the conductor, dynamic responses by one equivalent bundle model are very close to those by four bundles model.

References

[1]M.B.Thomas, A.H.Peyrot: Dynamic response of ruptured conductors in transmission lines, IEEE Transactions on Power Apparatus and Systems, vol.101 (1982), p.3022-3027.

[2]Alain H.Peyrot, Robert O.Kluge and Jun W.Lee: Longitudinal loads from broken conductors and broken insulators and their effect on transmission lines, IEEE Transactions on Power Apparatus and Systems, vol.99 (1980), p.222-231.

[3]North-east Electric Power Design Institute of State Power Corporation of China: Design manual for high-voltage transmission lines of electric power engineering, China Electric Power Press, Beijing (2007).

[4]F.L. Yang, J.B. Yang and D.J. Fu et al: Analysis on dynamic responses of transmission tower-line system under ice shedding, Journal of Vibration Engineering, vol.23 (2010), p.86-93.

[5]F.L. Yang, J.B. Yang and J.K. Han et al: Numerical Simulation on the HV Transmission Tower-Line System under Ice Shedding, in Proceedings of IEEE T&D Asia 2009, p.1-5.

[6]F.L. Yang, J.B. Yang and Z. Li et al: Study on the ice shedding and vibration suppression of transmission line. Journal of Vibration and Shock, vol.29 (2010), p.20-25.

[7]F.L. Yang, J.B. Yang and J.K. Han et al: Analysis on dynamic responses of transmission tower-line system under ice shedding, Journal of Structural Stability and Dynamics, vol.10 (2010): p.461-481.

[8]J. LI, Q. YAN and Q. XIE et al: Wind field measurements and wind-induced vibration responses of transmission tower during typhoon Wipha, J. Architecture and Civil Engineering, vol.26 (2009), p.1-7.

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.516

Analysis of subsynchronous oscillation in a multi turbo-generator sets power system based on electromechanical wave model

Chengbing He[1,a], Zhiyong Wang[2,b] , Jian Shen[3,c]

[1]Key Lab of Condition Monitoring and Control for Power Plant Equipment Ministry of Education (North China Electric Power University), Beijing, P, R. China

[2]Ningxia Guyuan Power Supply Bureau, Ningxia, P, R. China

[3]Guodian Solar Co., Ltd, Ningxia, P, R. China

[a]hcbyy@126.com, [b]wangzhiyong8660@126.com, [c]shenj@guodiansolar.com

Keywords: electromechanical wave model; subsynchronous oscillation (SSO); torsion; turbo-generator sets

Abstract. Electromechanical wave model is a new method in the research of generator-grid dynamic response of large scale electric power systems. Based on wave equations of elastic bearing transmit mechanical power and transmission line transmit magnetic power respectively, this paper built integrate electromechanical wave model of turbine-generator-grid with the example of No. 3 unit in Yimin Power Plant. With the help of electromechanical wave model, subsynchronous oscillation (SSO) example was given which indicated that the higher disturbance voltage, the bigger SSO amplitude; the larger damp, the smaller SSO amplitude, which will make the fluctuation become stable earlier. This conclusion has significant engineering value on reducing SSO amplitude of Yimin power plant.

Introduction

With the fast development of electric power system, High Voltage Direct Current(HVDC) transmission system, Thyristor Controlled Series Compensation(TCSC) transmission lines and new electrical equipment have been widely used. However, it has been tested that the transmission systems and the application of the equipment mentioned above would probably cause the problem of SSO which lead to serious accidents with dangerous to the safety and stability of electric system[1]. In May 2008, several cracks appeared in the coupling between low pressure shaft and generator shaft as well as in the generator rotors of No.3 unit in Yimin Power Plant. The electricity was transferred by transmission lines with TCSC. The analyzed results showed that the cracks were caused by SSO. It is clear that the research of SSO is of great significance.

The traditional method to analyze the effect of SSO to the stability of power systems was to build systematic mathematical models, then gave solutions by numerical method. But it was difficult to research on the transmission of electromechanical disturbance in large scale power systems through the traditional method[2]. Different to the traditional method, the continuum electromechanical wave model described the dynamic responses of electromechanical disturbance transmission in a large space system, which included generator, load and transmission line. It has become a new idea of generator-grid dynamic response in large scale power systems[3-5].

Integrated electromechanical wave model of torsion of multi turbo-generator sets

Elastic wave model of shafting torsion

The shafting torsion model of turbo-generator usually used multi-body lumped-mass model. Shafting was discretized into many rigid discs that connected through massless spring elements.

When the number of disk in shafting increased, the interval disk distance approached zero, then the discrete parameter would become continuous distribution equation. As for the multi-body lumped-mass model, wave equation can be written as[3]

$$J(x)\frac{\partial^2\theta}{\partial t^2}+d(x)\frac{\partial\theta}{\partial t}=K(x)\frac{\partial^2\theta}{\partial x^2}-\varepsilon(x)(\frac{\partial\theta}{\partial x})^2+T(x) \tag{1}$$

Where J denoted moment of inertia，d denoted dampling coefficient, K denoted shaft torsional stiffness， ε denoted internal dampling coefficient.

Electromechanical wave model of power systems

A chain power system composed by N sets of generators in which power was transmitted from the left to the right. In the chain power system, wave equation can be written as[3]

$$m(x)\frac{\partial^2\theta}{\partial t^2}+d(x)\frac{\partial\theta}{\partial t}=b(x)\frac{\partial^2\theta}{\partial x^2}-g(x)(\frac{\partial\theta}{\partial x})^2+P_m(x) \tag{2}$$

Where m denoted angular momentum，d denoted dampling constant，b denoted adjacency busbar susceptance，g denoted conductance.

Integrated wave model used for torsional vibration analysis

Eq.(1) was the wave equation of elastic shafting transmission mechanical power, where

$$T_e(x)=\varepsilon(x)(\frac{\partial\theta}{\partial x})^2-K(x)\frac{\partial^2\theta}{\partial x^2} \tag{3}$$

Both sides of Eq.(3) multiply Ω, Eq.(3) can be re-written as:

$$\Omega J(x)\frac{\partial^2\theta}{\partial t^2}+\Omega d(x)\frac{\partial\theta}{\partial t}=p(x)-p_e(x) \tag{4}$$

Where $p(x)$,$p_e(x)$ denoted the input and output power respectively in shafting coordinate x.

Eq.(2)was the wave equation of electromagnetism power in transmission line, where

$$P_e(x)=g(x)(\frac{\partial\theta}{\partial x})^2-b(x)\frac{\partial^2\theta}{\partial x^2} \tag{5}$$

$p_e(x)$denoted the power in the shafting coordinate x, then Eq.(2)can be re-written as:

$$m(x)\frac{\partial^2\theta}{\partial t^2}+d(x)\frac{\partial\theta}{\partial t}=P_m(x)-P_e(x) \tag{6}$$

Eqs.(4)and(6) were wave equations of transferring power in shafting and transmission line respectively. Both of their mathematical expressions and Physical meanings were the same. Therefore, the research of shafting of turbo-generator shafting and the connected transmission lines can be treated as a whole, then integrated wave equation of generator-grid can be given as[4]:

$$\Omega J(x)\frac{\partial^2\theta}{\partial t^2}+\Omega d(x)\frac{\partial\theta}{\partial t}=P(x)-P_e(x) \tag{7}$$

Torsion wave propagation in a multi turbo-generator sets with electric disturbance

N sets parallel operation turbo-generators in a power plant connected with power network through busbar. If there was a disturbance source pouring into the incident ray to the system at a point A, and the electromechanical wave would reflect and refract at a point B. The total power transilluminated from point B to the turbo-generators was :

$$P(t)=\frac{1}{Z_C}[\omega_2^+(t)-\omega_2^-(t)] \tag{8}$$

Transillumination power on *No.i* generator could be written as:

$$P_i^{'}=\frac{P(t)\prod_{i=1,j=1}^{n}Z_{Ci}Z_{Cj}}{Z_{Ci}\sum_{i=1,j=1}^{n}Z_{Ci}Z_{Cj}}(i\neq j) \tag{9}$$

Swing equation of the *No.i* generator could be represented in the form

$$M_i\frac{d\omega_2(t)}{dt}+D_i\omega_2(t)=P_i^{'}-\omega_2(t)/Z_{Ci} \tag{10}$$

Numerical example and discussion

Example introduction

There were 4 units in Yimin power plant, including 2×500MW（No.1, No.2 unit）and 2×600MW（No.3, No.4 unit）. No.1 and No.2 units were the same structure. No.3 and No.4 units were the same structure. The power plant was connected with Fengtun transformer substation through double circuit 500KV line in which Thyristor Controlled Series Compensation(TCSC) were employed to compensate the transmission loss. Obviously, the turbo-generators and transmission system composed a typical multi turbo-generator sets power torsional coupling system.

Main related parameters were shown in Tab.1and Tab.2.

Tab.1 Parameter nominal value of 500KV line

line	Positive sequence impedance	Negative sequence impedance	Positive sequence conductance	Negative sequence conductance
Yi-Feng 1st	0.0044+j0.0444	0.0132+j0.2038	2.1006	0.9233
Yi-Feng 2nd	0.0036+j0.0446	0.0108+j0.2009	2.3891	0.9233

Tab.2 natural frequency of No.3 unit torsional vibration

	The first-order	The second-order	The third-order
Natural frequency [Hz]	14.09	22.46	26.79

Simulation calculation

SSO phenomenon occuring in No.3 unit was corresponded the second order torsional vibration of shaft, thus the basic conditions of simulation computation were as follows: take No.3 unit as an object of study; inject a subsynchronous frequency(27.54Hz) voltage disturbance at series compensation point; assume 4% electricity consumption, 100% capacity.

Example1: distrubance source was 50KV subsynchronous voltage, and damping coefficient

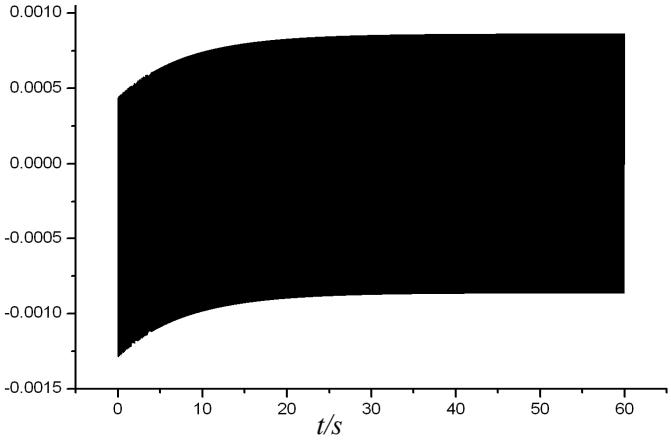


Fig.1. the response of relative speed of No.3 unit in 60 seconds

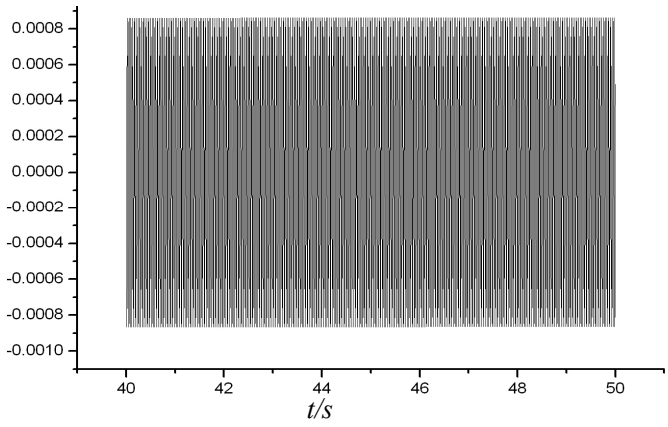


Fig.2. the response of relative speed of No.3 unit after rotative speed difference is steady

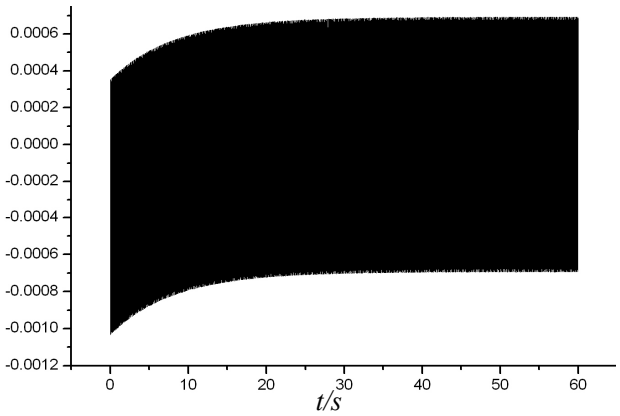


Fig.3. the response of relative speed of No.3 unit after reducing the excitation voltage

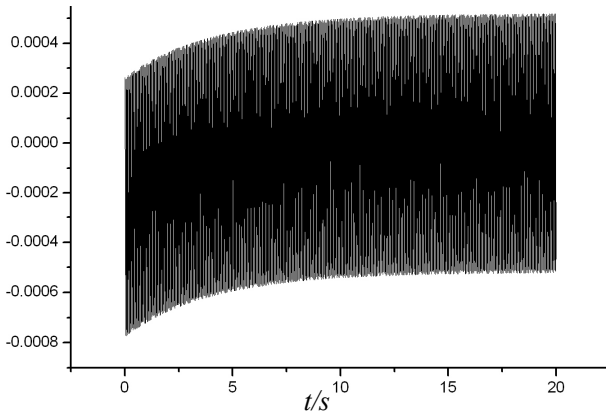


Fig.4. the response of relative speed of No.3 unit after increasing the damping

was 0.1/S.

Fig.1 showed the response of relative speed of No.3 unit. Fig.2 showed the response of relative speed of No.3 unit after rotative speed difference became steady. The vertical axis in all figures was the ratio of differential value with rated speed. The figures showed that the wave amplitude of relative speed of No.3 unit kept steady after injecting a distrubance, but the balance position of the wave had obviously changed. At the beginning, the balance position of the wave motion was less than zero. However, after adding a distrubance for 40s, it was taken rated speed as balance position of generator's ralative speed, and the wave amplitude kept steady.

Example 2: disturbance source was 40KV subsynchronous voltage, and damping coefficient was 0.1/S.

Fig.3 showed the response of relative speed of No.3 unit. Compared it with Fig.1, it could be seen that the wave amplitude of the relative speed of the generator decreased after reducing the excitation voltage. When the excitation voltage was reduced by 20%, the wave amplitude decreased from 1.65×10-3 to 1.4×10-3, down by 15%.

Example 3: disturbance source was 40KV subsynchronous voltage, and the damping coefficient was increased to 0.2/S.

Fig.4 showed the response of relative speed of No.3 unit after increasing the damping.Compared it with Fig.3, it could be seen that the wave amplitude decreased from 1.4×10-3 to1.0×10-3, down by 29%.

From the above examples, it was concluded that the amplitude of SSO would be increased when the disturbance voltage increases, and it decreases when the system damping increases, and

the wave amplitude becomes steady more quickly. Thus, reducing the excitation voltage (reducing the compensation degree of the series capacitors) and increasing the system damping (whether the mechanical damping or the electrical damping of the transmission lines) can reduce the amplitude of SSO effectively.

Conclusions

The continuum electromechanical wave model for power system is a new idea to explore the electromechanical dynamic response of large-scale electrical power network. The model also described the whole power system including generators, loads and transmission lines. Through establishing the wave equations for mechanical transferring power in elastic shaft and transferring electromagnetic in transmission line, the electromechanical wave model was developed. Based on the model, a numerical example for SSO was presented here, which took the No.3 turbo-generator unit in Yimin power plant for example. By analyzing the SSO response, it was concluded that the amplitude of SSO increases when the disturbance voltage increases; it reduces when the system damping increases. At the same time, the wave motion becomes steady. The conclusions of the paper have significant engineering value on reducing SSO amplitude of Yimin Power Plant.

Acknowledgment

The authors are grateful for the financial support from National Science Foundation of China (Grant No. 50905058), program for new century excellent talents in university(NCET-08-0769) and the Fundamental Research Funds for the Central Universities(Grant No. 10MG14).

References

[1] Oriane M. Neto, Donald C. Macdonald. Analysis of subsynchronous resonance in a multi-machine power system using series compensation[J]. *Electrical Power and Energy Systems*, 2006,28: 565-569

[2] Zhang Baohui, Yu Yinghui. Analysis for the effect of electrical operation on turbine generator's shaft torsional oscillation[J]. *Proceedings of the CSEE,*2001,21(9):24-28

[3] Wang Delin, Wang Xiaoru, James S.Thorp. Electromechanical Wave Continuum Model for Power System[J]. *Proceedings of the CSEE,* 2006,26(17):30-37

[4] Wang Delin,Wang Xiaoru. Study on Characteristics of Electromechanical Wave Propagation in the Continuum Model for Power Systems[J]. *Proceedings of the CSEE*,2007,27(16):43-48

[5] Parashar M, Thorp J S, Seyler C E. Continuum modeling of electromechanical dynamics in large electric power systems[J]. *IEEE Trans on CAS-I*, 2004,51(9):1848-1858

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.521

A Heuristic Resource Management Method in Grid Computing Environment

FANG Xian-Mei

Department of Computer and Information Science,Hechi University,Yizhou,Guangxi 546300,China
Email:791555595@qq.com

Keywords: grid computing, Web service, open grid services architecture, resource management

Abstract: Grid is an emerging infrastructure which enables effective coordinate access to various distributed computing resources in order to serve the needs of collaborative research and work across the world. Grid resource management is always a key subject in the grid computing. We first analyze the resource management in the grid computing environment, then according to the load imbalance question in the ant colony optimization algorithm, propose an improved algorithm that suits to be used in the grid environment.

Introduction

The goal of Grid is to realize share of homogeneous resources that locate all over the world. Grid resource management is always a key subject in the grid computing. Because the Grid is homogeneous, distributed and dynamic, traditional methods of resource management will not function well in Grid environment. Service with its original knowledge and how to enhance the scalability of grid resource management system. Information technology is changing rapidly, and now forms an invisible layer that increasingly touches every aspect of our lives. Power grids, traffic control, healthcare, water supplies, food and energy, along with most of the world's financial transactions, now depend on information technology. Grid computing is an element (Metacomputing), it is the physical distribution system, a variety of heterogeneous computing resources connected by high-speed networks together to solve large application problems.

Grid refers to the entire Internet into a single, large, integrated supercomputer to achieve computing sources, storage resources, data resources, information resources, knowledge resources, expert resource sharing, its size can be as large as a state, small to enterprises, local area networks, and even families and individuals. Grid areas are mainly distributed supercomputing has distributed instrumentation systems, data-intensive computing and remote immersion four class. Grid is integrated into the entire network a huge supercomputers, achieving computing resources, storage resources, data resources, information resources and knowledge resources, resource sharing, experts. In recent years, computing power with the computer the rapid growth of Internet popularity and the cost of a significant reduction in high-speed network as well as traditional methods and computer calculations of Shi Yong way of change, grid computing a super computing development an important trend. Grid computing is a new and important research area, which to a large granularity resource sharing, innovative application of high performance computing and the main features of the 21st century will become an important driving force of economic development.

Grid computing at least have the following three kinds of basic functions: Task Manager: the user submits the task, specify the resources for the task, delete the task and monitor tasks running .Task Scheduling: The user's task presented by the type of the function in accordance with

the mandate, the required resources, the resources available to run the program and policy arrangements. resource management: Identify and monitor network resources, gather the resources of the task run-time occupancy data. To achieve the goal of grid computing, we must address the following three key issues, which is now widespread Internet problems: heterogeneity: As the grid by the distribution of the same below the WAN management domain composed of a variety of computing resources, how to achieve the cooperation between heterogeneous machines and conversion is the primary problem. scalable nature: To network resources have been expanding, growing application of the situation, the do not degrade performance. dynamic adaptability: When a resource failure or the possibility of failure is high, the resource management must be able to dynamically monitor and manage network resources, from the available resources, select the best resources and services.

Web Service and Open Grid Services Architecture

OGSA (Open Grid Services Architecture), including two key techniques, namely the Grid and Web Service technology, which is based on the five hourglass structure, combined with Web Service technologies put forward to address two important issues - the definition of standard service interfaces and protocols recognition. To serve as the center is the basic idea of OGSA, everything is in service in OGSA. In the OGSA framework, will be all abstracted as services, including computer programs, data, equipment, and which is conducive to a unified standard interface through management and use of the grid. Web Service provides a framework based on service structure, but the Web Service the face of the generally permanent services, whereas grid environment, a large majority are temporary short-term services, such as a computing tasks execution. Taking into account the specific characteristics of grid environment, OGSA service concept in the original Web Service based on the proposed "Grid Service" concept, used to address service discovery, dynamic service creation, service management and life cycle temporary service related issues. Based on the concept of grid services, OGSA grid as a whole "Grid service" collection, but this collection is not static; it is extended, reflecting the dynamic nature of the grid. Grid service completed by defining interfaces to different functions, service data is the information on the Grid service instance, the grid services can be simply expressed as a "grid service = interface / behavior + service data." In the definition of OGSA, only Grid Service interface is required. Fig.1. is the Web service.

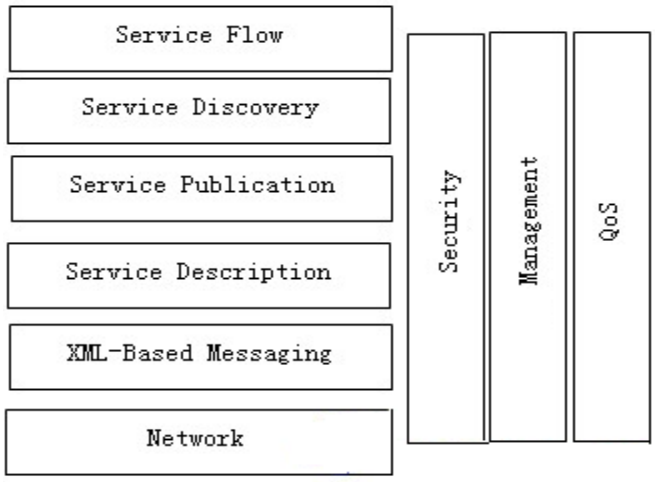


Fig.1. The Web service

OGSA basic idea is to "services" as the center. In the OGSA framework, will all be abstracted as services, including computing resources, storage resources, networks, procedures, databases, etc., in short, all services. This concept is conducive to adoption of a unified standard interface to

manage and use grid. OGSA defines a Grid Service concept of grid service is a Web Service, the service provides a set of interfaces to comply with the definition of clear and specific management to address service discovery, dynamic service creation, life cycle management, notification and other issues. In OGSA, the Grid services will all have seen, so the grid is scalable set of grid services. Grid services can be gathered in different ways to meet the needs of virtual organizations, virtual organization itself can be partially based on their shared services operations and defined. In short, the grid service = interface / behavior + service data. Figure 4 is a brief description of grid services.

Grid technology (such as the Globus package) and Web Service is OGSA two supporting technologies. Globus is the science and engineering has been widely accepted in the field of grid computing solutions. It is a community-based, open structure, open source set of services also support the grid and grid software library. The toolkit addresses security, information discovery, resource management, data management, communications, and error monitoring and portable and so on. Closely with the OGSA Globus components GRAM Grid Resource Allocation and Management Agreement and the guard (Gate Keeper) services, which provide a safe and reliable service creation and management capabilities, meta directory services through soft state registration, data model and the local registration to provide information discovery, GSI (Grid Security Infrastructure Grid Security Architecture) support a single login point in the map and trust. These features provide the structure necessary for the service element, but the versatility than OGSA smaller.

Web Service is a standard application of the framework of the access network. XML protocol related work is the basis of Web Service. Web Service in a few of the more important the protocol standard is SOAP (Simple Object Access Protocol, Simple Object Access Protocol), WSDL (Web Service Description Language, Web Services Description Language), and WS-Inspection, UDDI (Universal Description, Discovery & Integration, unified Description, Discovery and Integration). SOAP is XML-based RPC (Remote Process Call, remote process call) protocol, used to describe the common WSDL target. Support will be extended through the SOAP Web Service Framework for security. WSDL for describing services, including interfaces and access methods, complex services can be composed of several services; it is the Web Service interface definition language. WS-Inspection service description is given a definition of the practice, including a simple of the XML language and related management, used to locate the service provider's published.

Resource management

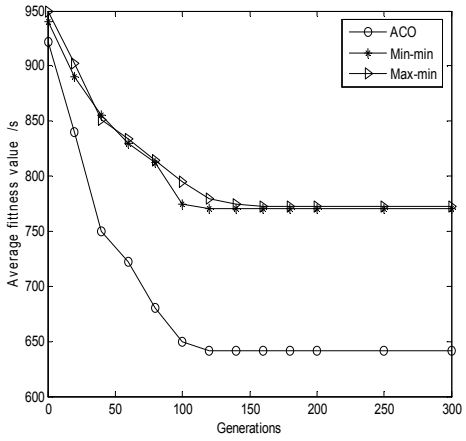
To better manage the grid resources, VO resource organization and management structure must be able to effectively reflect the grid computing environment, resources, organic makeup, so should generally reflect the structure of VO wide distribution, local concentration and local application management control features. In order to reflect the grid computing "service-oriented" design concept, VO forms of organization of resources and resource stimulate the process of treatment must be followed "for the job processing, task-oriented requests," the principle of instant excitation, and can work with treatment scale fluctuation and change their share of resources to adjust their own information processing methods to achieve the best possible methods to deal with consumption optimization. Since VO resources in the matching process, there is always such a situation exists: the existing grid computing environments in a region already have considerable scale segment operating group responsible for the operation scheduling and resource matching group VO groups have occupied a large section of the area of computing resources. If at this time by the regional section of an entity to submit a sizable operations to process the request T1, is produced by the prior

possession of the virtual organization VO1 appropriate computing resources to match the T1 treatment, however, paragraph in the region can offer idle resources, resource demand is only a small part, other 80% of the resources must be to obtain wide-area grid environment. If on the wide area under the jurisdiction of the management of resources is still limited to VO1, then not only will cause a lot of time delays and network load.

Globus is a global resource management structure, the local two-resource model. Globus resource model is mainly composed of 4 parts: resource broker Broker, cooperative distributor DUROC, Resource Information Service Component Information Service and Resource Allocation Manager GRAM. In addition, Globus also provides a resource description language RSL, is used to transfer resources between the various components of the request. Resource broker role is analysis application program resource requests submitted by the senior RSL description and converts it into more specific needs described to distributor on, collaborative task is divided into sections, GRAM resource processing. Information services function is to provide global broker and co-distributor of information resources. MDS is the grid computing environment information service center, using LDAP protocol to form a centralized directory tree hierarchy based on the information server. DUROC processing applications can be made on a number of resources requested for the application of co-allocation of different resources. GRAM is the Globus in shielding the heterogeneity of the various components of local resources. Functions including the provision of access to local resource management tool API interface; analysis and processing of RSL description; resources for remote application request processing, remote task scheduling processing, remote task management.

Application of ACO

Grid resources include computing resources, storage resources and communication resources. How to use these resources efficiently complete the task of grid computing research priorities. For grid users, submit it to the grid task, the grid scheduler according to some scheduling policy to the user submits the task assigned to the grid system resources available. Evolutionary Algorithms (EA), as a family of computational models inspired by the natural process of evolution, have been applied with a great degree of success to complex design optimization problems. In Genetic Algorithms (GA), Ant colony (ant colony optimization, ACO), also known as ant algorithms is used to find the optimal path in the graph the probability of technologies. It consists of Marco Dorigo in 1992 in his doctoral thesis in the introduction, its inspiration from the ants found in the course of the path in search of food behavior. Its main idea is: Each ant in the path traversed left pheromone, ant based on the current path of the pheromone according to the probability of a certain selection path. The shortest path is taking more than ants, and thus obtains the optimal solution. Figs.2-3. are the experimental result.



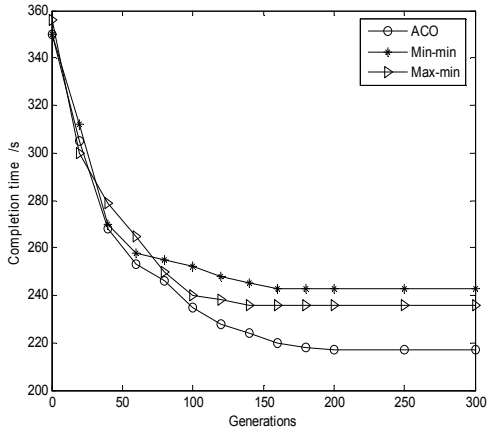


Fig.2. Average fitness values comparison of ACO with other two algorithms

Fig.3. Completion time comparison of ACO with other two algorithms

Conclusions

The technology of grid computing has been a new research area, which appeared in the 1990s. Grid system is composed of heterogeneous resources. In grid computing, a good algorithm for tasks scheduling should not only decrease the makespan of all tasks but also balance the load among the resources in the grid system. The framework offers a comprehensive solution for the efficient parallel evolutionary design of problems with computationally expensive fitness functions, by providing novel features that conceal the complexity of a Grid environment through an extended API and a meta scheduler for automatic resource discovery. An intelligent Grid environment will contain more intelligent functions for resource management, security, Grid service marketing, collaboration and so on . These functions will make the Grid environment more efficient, secure, economic in using resources, intelligent and easy to use. Intelligent resource management is a core component in the intelligent Grid environment.

References

[1] L. Smarr and C. Catlett, Metacomputing, Communication of the ACM, Vol.35. No.6(1992),pp. 44-52.

[2]Rajkumar Buyya, DavidAbramson, Jonathan Giddy.Grid Resource Management, Scheduling,And Computational Economy [A].WGCC 2000[C].Japan,March 15-17,2000.

[3] F.J. Villegas, T. Cwik, Y. Rahmat-Samii, M. Manteghi, A parallel electromagnetic Generic-Algorithm Optimization (EGO) application for patch antenna design, IEEE Transactions on Antennas and Propagation 52 (9) (2004) 2424–2435.

[4] Das Sajal K, Harvey Daniel J, Biswas Rupak. MinEX:A latency-tolerant dynamic partitioner for grid computing Applications[J].Future Generation Computer Systems, March 2002, 18(4):477-489.

[5] Foster, I.; Kesselman, C.; Nick, J.M.; Tuecke, S. “Grid services for distributed system integration”， Computer , Volume: 35 Issue: 6 , June 2002， Page(s): 37 -46 [5]Min R,Maheswaran M.SchedulingAdvance reservation with priorities in grid computing system[A].13AISTED International Conference on ParallelAnd Distributed Computing Systems(PDCS'01)[C].Aug 2001:172-176.

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.526

An Efficient Resources Management Model and Task Scheduling Algorithm in Grid Computing

GAO Xiao-Bo

Department of Computer and Information Science,Hechi University,Yizhou,Guangxi 546300,China
Email:791555595@qq.com

Keywords: grid computing, resources management, task scheduling, open grid service architecture, open grid services infrastructure

Abstract: From the perspective of resource sharing, grid computing is a system ranging from small kind of network system for home using to large-scale network computing systems even to the Internet. The management of resources in the grid environment becomes very complex as these resources are distributed geographically, heterogeneous in nature, and each having their own resource management policies and different access as well as cost models. In this paper, we bring forward an efficient resources management model and task scheduling algorithm in grid computing. The simulation results show that the proposed algorithm achieves resource load balancing, and can be applied to the optimization of task scheduling successfully.

Introduction

Grid technology emerged in the 1990s of an information technology, which represents Internet technology and Web technology, following after the third technology wave [1]. Widely used in IT industry, "Grid" and "grid computing" combines two terms and the related areas of information in the field of modern advanced technology. Grid is based on Internet-based new technology, to the Internet, online computers, large databases, remote devices that integrate the virtual "super computing system". To realize computing resources, storage resources, communication resources, software resources, knowledge resources sharing. According to the functionality provided by the grid, grid computing can be divided into two major categories of grid and storage grid, according to the different grid applications; grid computing can be divided into grids, data grids and grid services to three categories. Grid's goal is to share resources and work together to solve the multiple levels of resource sharing and collaboration of new technology, new standards, will interact with Internet information from the communications platform to a resource sharing platform for grid will become the essential IT infrastructure.

The key technology is the grid system, grid system software and operating systems, the core including resource management, information optimization, and job scheduling and grid safety technologies. Specifically showed: grid node, that is the provider of grid computing resources, which includes high-end server, cluster systems, MPP systems, and databases, large storage devices; broad band network system is to provide high performance grid computing environment in communication of necessary means to significantly affect the properties of the grid; resource management and task scheduling, resource management tool which is responsible for calculating the description of resources, organization and management, task scheduling system is currently under the load on the system within dynamic scheduling tasks; monitoring tool, its role is to monitor through performance analysis and monitoring system resources and operation; visualization

tools, the calculation results into intuitive graphical information, and provides a friendly user interface; grid security, safety issues related to confidentiality of authentication and authorization issues shared problem that requires human intervention, or to strengthen safety measures must be orderly grid control and management[2].

Grid computing is rapidly along with the Internet developed, specifically for the new scientific computing complex calculation mode. This calculation model is the use of the Internet to geographically dispersed computers organized into a "virtual supercomputer", each of which involved a computer calculation is a "node", but the calculation is made of thousands “node "consisting of" a grid ", so this calculation is called grid computing [3]. It organized a "virtual supercomputer" has two advantages, one is powerful data processing capabilities; the other is to make full use of the Internet's idle processing power. First, the grid computing model to calculate the data is divided into a number of “small pieces”, and calculates the "small piece" of software is usually a pre-programmed screen saver, and then the computer can be different nodes according to their capacity to download one or multiple data segments, and this screen saver. So "show began," As long as the node of the computer users do not use the computer, the screen saver will work, so this computer was idle computing power fully mobilized.

The rest of this paper is organized as follows: grid architecture is discussed in Section 2.Section 3 describes the resource management .Section 4 describes the task scheduling. Finally, we conclude this paper and show future work direction in section 5.

Grid Architecture

Grid architecture is the structure of the overall framework of the infrastructure. It is about how to build grid technology, which gives the grid basic components and functions, describing the grid part of relation and the method them or manner, depicts the support effective the mechanism of the grid [4]. The more important current grid architecture has two, one is proposed by Foster and other five in the earlier hourglass structure (Five-Level Sandglass Architecture), which is based on "TCP / IP protocol" as the central emphasis on the service and API (Application Programming Interface) and SDK (Software Development Kits) importance. Five components of the hourglass structure according to the distance and sharing of resources will be functions scattered in five different levels, from bottom to top are: structural layer, link layer, resource layer, collective layer and application layer, the more the lower The more close to the physical shared resources; and the number of different parts of the agreement is different, for the most part of the core, upper layer protocol to be able to achieve the mapping to the core protocol, while achieving other key agreement protocol to lower the mapping, so the formation of the core protocol of the agreement hierarchy bottlenecks in five-story structure, the resource layer and link layer together constitute the core. The other is represented in the IBM under the influence of the industrial sector, taking into account the development of Web technology and the implications, Foster and other Web Service proposed combining OGSA (Open Grid Service Architecture, OGSA). Open Grid Services Architecture OGSA is the second most important after five hourglass structure is the latest of a grid architecture. Fig .1. is grid computing model.

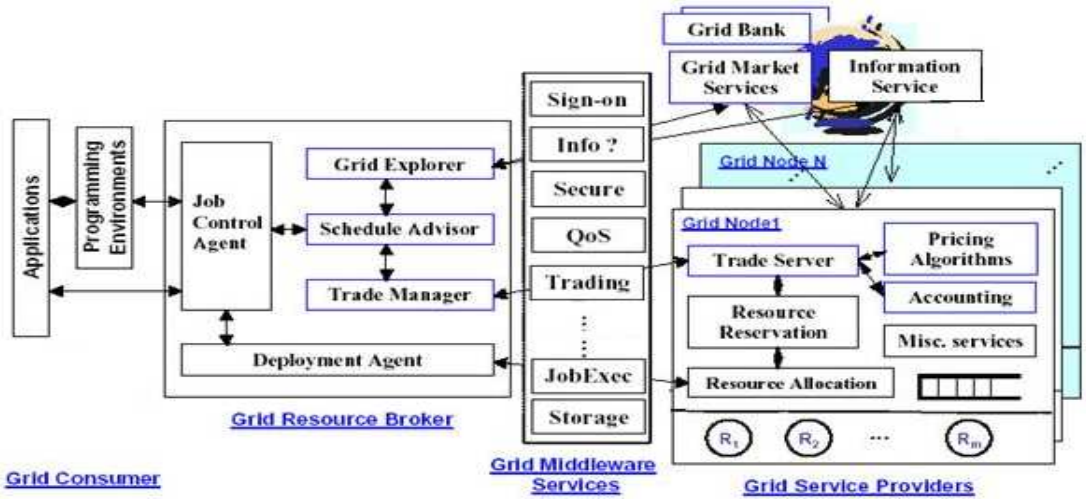


Fig 1 .Grid computing model

OGSA framework is four main layers. From bottom to top were: 1. Resource layer: The concept of resource level and in the normal sense OGSA center of grid computing, including physical and logical resources. Physical resources including servers, storage and networks. Logical resource management through the polymer layer of virtualization and resources to provide additional functionality; 2.Web services and the definition of grid services, OGSI (Open Grid Services Infrastructure, OGSI) expansion: OGSA all grid resources (logical and physical) are understood as services. OGSI specification defines grid services and the establishment of a standard Web service technology. OGSI use such as XML and Web Services Description Language (Web Services Description Language, WSDL) Web service mechanisms such as, for all grid resources to develop the standard interfaces, behaviors and interactions. OGSI extends the definition of Web services provides a dynamic, stateful, and manageable Web services capabilities, which is modeled on the grid resources are necessary; 3. OGSA architecture based services: Web services OGSI extension for the next layer and the layer provides the infrastructure: the Grid service-based architecture. GGF is currently working in areas such as program execution, data services and the definition of core services in areas such as service-based grid architecture. With these services, the new structure began to appear, OGSA will become a more useful service-oriented architecture (Service Oriented Architecture, SOA); 4. Grid procedure: With the passage of time, a rich web-based lattice structure of the services to be developed, using one or more service-based grid architecture will be also a new grid procedure. These applications constitute the fourth major OGSA architecture layer. Construction of two support technology OGSA Globus and Web Service. Globus is the science and engineering has been widely accepted in the field of grid computing to solve the program. Web Service is a standard framework for accessing network applications. At present, the grid services interface is still relatively limited, the next major consideration OGSA expansion management, security and other aspects.

Resource Management

In grid computing, resource and resource management are two very broad concepts, resources include hardware resources (computers, storage, network infrastructure, etc.), software resources (virtual services) and data. Resource management is a resource that consumers use the resources available throughout the life cycle of resources to identify needs, match resources, allocating

resources, scheduling and monitoring of resources a process. Grid resource management involves four roles: resource consumers, resource scheduler, and grid middleware and grid resource providers. Grid ultimate aim is to focus on the development of resources from the simple to the data sharing, and finally to collaborate and quality of service (Quality of Service). Resources so companies can focus on the company's entire IT infrastructure as a computer, according to their needs can not be used to find the resources. Data sharing to enable companies to access remote data [5]. This is particularly useful for certain life science projects because these projects, companies and other companies need to share human genetic data. Grid computing to enable collaboration through broad dispersed organizations can collaborate on some projects, integrate business processes, and share from engineering blueprints to the software application and all other information, co-processing project in question. QoS refers to different users or different data streams adopt correspondingly differentiated the priority, or is application requirements. The same network node to provide the guaranteed quality of service.

Grid scheduler is resource management to implement the module, grid middleware is similar Globus, Condor, Legion and other grid software infrastructure, and some grid middleware itself contains the scheduler, so figure shows the grid scheduler and grid middleware surrounded with a dotted line. Grid scheduler implementation of resource management process is called "scheduling." Resource scheduling in three phases: resource discovery, system selection, operation execution. Currently in the field of grid computing, there are three stages for this made a lot of research.

4 Task scheduling

If you stand on the user's point of view, grid scheduling, also known as "application schedule" or task scheduling. Task scheduling is to allocate parallel program interdependent tasks to resources Ikegami, to achieve the minimum overall execution time of the target. This view that scheduling is a more specific level that is fine-grained scheduling. In a traditional parallel distributed environment, task scheduling management system has been studied very good. But the grid compared to traditional networks, with higher heterogeneity, instability and the need for scheduling tasks across administrative domains, so the fact that task scheduling which is facing new challenges.

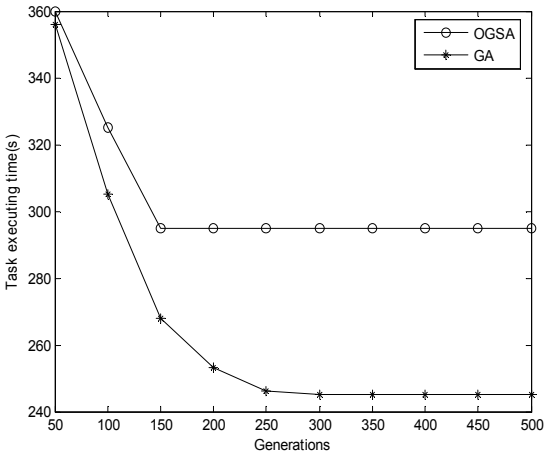


Fig.2. Task executing time comparison of GA with scheduling based on OGSA

When a user submits a job to job control agent, the agent will permit control agent scheduling, admission control agents and grid information system communication, checks whether the resource requirements are met, whether submitted to the scheduler, the scheduler uses the network Grid

service discovery resources and information, and domain control agent negotiation to determine the current status and availability of resources, and then make a decision and transmission to implementation of the mapping agent, agents and resources to implement the domain where the domain control agent negotiation using the resources, and sometimes may need to pre- resources. Monitor the process of surveillance, if the performance is lower than expected, then decide whether to re-schedule.Fig.2. is the experimental result.

Conclusions

Modern science, life becoming more and more urgent need for more computing power the earth, many computational problems can no longer use a conventional supercomputer to complete. Meanwhile, as the computer enhancements, the emergence of high-speed Internet, allows us to be a large class of geographically far apart the various resources such as supercomputers, storage systems, databases, and some special equipment together as a unified resource to use, we call grid. However, the use of multiple geographically distributed organization-wide resources, grid computing environment network computing environment in general has a more complex features, such as the existence of multiple administrative domains, site autonomy, the system of dynamic, heterogeneous, mission uncertainty and higher communication delays and other uncertainties. The existence of these features makes the grid environment, resource allocation, task scheduling more difficult. The traditional round robin scheduling method can not effectively adapt to grid environment, there must be a reasonable management and scheduling algorithms to fully play to the characteristics of the grid.

Acknowledgments

This research work was supported by the National Natural Science Foundation of P.R.China under Grant No. 60902053. The authors would like to thank the anonymous reviewers for their insightful comments and constructive suggestions that have improved the paper.

References

[1] Peng Liu, Yao Shi, San-li Li. Computing Pool--a Simplified and Practical Computational Grid Model[C].*The Second International Workshop on Grid and Cooperative Computing,* 2003;(12):7-10.

[2] Abraham A,BuyyaR,Nath B. Nature's heuristics for scheduling jobs on computation grids. *In proceedings of the 8th IEEE international conference on advanced computing and communicationas, Conhin, India,* 2000, pp.45-52.

[3] I. Foster, C. Kesselman, S. Tuecke. The Anatomy of the Grid: Enabling Scalable Virtual Organizations [J].*International J. Supercomputer Applications, 2001*,15(3):1024-1031.

[4] Veysel Aslantas, Saban Ozer,and Serkan Ozturk. A Novel Clonal Selection Algorithm Based Fragile Watermarking Method[C].*Lecture Notes on Computer Science 4628.Sprnger-Verlag Berlin Heidelberg 2007*, pp.358-369.

[5] Singh, Carl Kesselman and Ewa Deelman. A Provisioning Model and its Comparison with Best Effort for Performance-Cost Optimization in Grids. *The 6th IEEE International Symposium on High-Performance Distributed Computing*, 2007, pp.117-126.

doi:10.4028/www.scientific.net/AMM.50-51.531

Research on Network Security and Data Encryption

Zhongqi Duan [1, a], Xiaofeng Xue [2,b]

[1] School of Information Science & Engineering, Changzhou University ,Changzhou, China

[2] School of Computer Engineering,Jiangsu Teachers University of Technology, Changzhou, China

[a]duan@cczu.edu.cn, [b]571537@qq.com

Keywords: network security, access control, data encryption, network firewall

Abstract: In the old days of the mainframe, information was centralized so security was not as big an issue. Nowadays, with the advent of information being distributed across many servers located in different areas, security is a major concern. Protecting the information is a high priority. A network with a good accounting and auditing system will ensure that all activities are logged, thereby enabling malicious activity to be detected.

Introduction

21 computers around the world together through the Internet together, the information content of security also has fundamentally changed. It not only from the general's defense become very common precautions, but also from specialized fields becomes ubiquitous. When the man entered the 21st century information society, network society of the time, China will establish a complete network security system [1]. Network security is that the network hardware, software and system data is protected, not because of accidental, or malicious reasons suffered, disclosure, system normal running of the network service was not interrupted. Network security from its essence, is the network information security. Broadly speaking, all information relating to the network the confidentiality, integrity, availability, authenticity and control technologies and theories are related to network security research. Network security is one related to computer science, network technology, communication technology, cryptography, information security technology, applied mathematics, number theory, information theory and other disciplines in the comprehensive discipline. With the rapid development of computer technology, information networks have become an important guarantee for social development. Many of them are sensitive information, or even a state secret. So inevitably attract all kinds of people from around the world attacking (such as information leakage, information theft, data tampering, data deletion Tim, computer viruses, etc.). Meanwhile, the network entity also suffering from such as flood, fire, earthquake, electromagnetic radiation and other aspects of the test.

Information security is the country's development an important issue facing. For this problem, we do not go from the planning system to consider it, from the technical, industry, policy development it came. Government should not only see the development of information security is high in China part of the technology industry, and should see the development of the security industry, information security policy is an important component of the system, and even the future of our country should see its electronic, information technology development will play a very important role.

Network Security Type

Operating system security is information processing and transmission systems to ensure safety. It focuses on ensuring normal operation of the system, because the crisis and expected messy to avoid damage to the system storage, processing and transmission unfavorably with the devastation and loss to avoid the electromagnetic leakage, generate information, interference of others. The security of the network system information [2]. Including user password authentication, user access control, and data access, methods of control, security audit, security tracking, computer virus prevention, and data encryption.

Information dissemination network security, that is, the safety consequences of information dissemination. Including information filtering. It focuses on prevention and control of illegal and harmful information to spread the consequences. To avoid a large number of free public network transmission of information out of control. Web content security. It focuses on protecting the confidentiality of information, authenticity and integrity. To avoid the attacker using the tapping system security vulnerabilities, impersonating, fraud, and undermine the legitimate user behavior. Essentially to protect the interests of users and privacy.

Network security features

Network security should have the following four characteristics: Confidentiality: information is not disclosed to unauthorized users, entities or processes, or for their use features. Integrity: data can not be changed without the authorization features [3]. That information stored or transmitted in the process of being amended to maintain, not to be damaged and missing features. Availability: authorized entities can access press needs to use. That is, when necessary, can access the information needed. For example, denial of service under the network environment of network and the normal operation of the

system belong to the availability of the attacks; controllable: the dissemination of information and content with the control.。

From the network operation and management perspective, they want access to information on the local network, reading and writing operations are protected and controlled, to avoid a "trap door", viruses, unauthorized access, denial of service and network resources and illegal control of the illegal occupation and other threats, to stop and defense network hackers. They want to illegal, harmful or confidential information involving state filtering and the prevention of and avoid disclosure of confidential information to prevent harm to society, causing huge losses to the state. Education and ideology from the social point of view, unhealthy content on the network, will social stability and obstruct human development and must be controlled. With the rapid development of computer technology, computer processing business also by mathematical calculations based on single, document processing, simple to connect the internal network based on internal business processing, office automation, development of complex internal network-based (Intranet), Extranet (Extranet), the global Internet (Internet) of enterprise-class computer processing system and the worldwide information sharing and business processes. Increase in system capacity, the system's connectivity also constantly improved. However, connectivity information, ability to increase circulation, while the connection-based network security issues become increasingly prominent, the overall network security mainly in the following areas: network, physical security, network topology, security, network security, application system security and network management security.

Network Security System, on the one hand due to authentication, encryption, monitoring, analysis, recording, etc., thus affecting the efficiency of the network, and reduce customer

application flexibility; on the other hand also increased administrative costs. However, from network security threats are real, particularly in the run key business networks; network security is the first problem to be solved. Select appropriate technologies and products for flexible network security policy, in case of network security, providing flexible access network services.

Security techniques

Physical measures: for example, protection of key network equipment (such as switches, mainframe computers, etc.), strict network security regulations, to radiation protection, fire protection and installation of uninterruptible power supply (UPS) and other measures. Access control: user access to network resources on a rigorous certification authority and control. For example, for user authentication, encryption of the password, updates, and identify, set user access to directories and file permissions, the permissions control network device configuration, and so on.

Data Encryption: Data encryption is also an important part of the network policy. Encrypting information being sent across the network can reduce the probability of that information being intercepted. Encryption is an important means to protect data security. The role of encryption is to protect the information being intercepted can not read its meaning. Computer network to prevent viruses, install anti-virus system network. Other measures include information filtering, fault tolerance, data mirroring, backup, and auditing. In recent years, network security issues surrounding a number of solutions, such as data encryption and firewall technology. Data encryption is a network of data transmission is encrypted, and then decrypt the destination to restore the original data, to prevent unauthorized users intercepted theft information. Through the network firewall technology is the isolation and restricted access to other methods to control network access.

Firewall Technologies

Network firewall technology is a used to enhance network access control between, to prevent the external network users to illegally enter the internal network through an external network, access to internal network resources, protect the internal network operating environment of the particular networking device. It two or number of data packets between networks such as the connection mode according to some security policy to implement checks to networks of communication is to be permitted, and to monitor network running. Firewall products bastion, and circuit level gateway, shielding the host firewall, dual-host and other types. Internet firewall is such a system (or set of systems), it can enhance the in-house network security. The Firewalls defines a firewall as "a system or group of systems that enforces an access control policy between two networks." In the context of home networks, a firewall typically takes one of two forms: Software firewall - specialized software running on an individual computer, or Network firewall - a dedicated device designed to protect one or more computers.

Firewall system determines which access internal services can be outside; the outside world who can access what services internal and external service which can be visited inside. To make an effective firewall, all the information from the Internet and to go through the firewall, the firewall checks accepted. The firewall allows only authorized data through, and from the firewall it must be able to penetrate. Not just the router firewall, bastion host, or any equipment to provide network security portfolio of firewall security policy is a part. The establishment of comprehensive security strategy defense system, even: tell the user should have the responsibility, the company provides network access, service access, local and remote user authentication, inbound and outbound, disk and data encryption, virus protection measures and employee training. All the possible places of

attack must be protected at the same level of security.

Intrusion detection is a measure of an effective defense system is an important factor in a complete, powerful integrated firewall, intrusion detection system can make up a relatively static defensive deficiencies. From external networks and campus networks of the various acts of real-time detection, to detect all possible attack attempts, and take corresponding measures. Specifically, intrusion detection is to switch on the engine access center. Set of intrusion detection intrusion detection systems, network management and network monitoring functions in one, capturing and transferred between internet and

intranet all the data in the built database of signatures, using matching and intelligent analysis, testing happening on the network intrusion and anomaly, and the incidents recorded in the database as a network administrator based on a post hoc analysis; If the situation is serious, the system can send real-time alarm, a school administrator to take timely response.

Use of firewall technology, careful configuration, usually between the internal and external networks to provide secure network protection and reduce network security risks. However, only the use of firewalls, network security is far follows: (1) might be behind a firewall an intruder may find the back door open. (2) The intruder may be inside the firewall. (3) The performance of the restrictions, anti-flame usually can not provide real-time intrusion detection capability. Intrusion detection systems in recent years, the new network security technology is designed to provide real-time intrusion detection and take necessary protective measures, such as recording evidence for tracking and recovery, broken network connections. Real-time intrusion detection capability is important first of all it can

to deal with the internal network from attacks, followed by the invasion it can shorten the time hacker. Intrusion Detection System can be classified into two categories: host-based and network-based. Host-based intrusion detection system for the protection of critical application server, real-time monitoring of suspicious connections, system log inspection, intrusion and other unauthorized access and to provide surveillance of the typical applications such as Web server applications. Network-based intrusion detection system for real-time monitoring information on the critical path network.

Network equipment and network applications in the market booming, driven by the network security market in recent years, ushered in rapid development period, on the one hand as an extension of the network, the network has expanded rapidly and become increasingly complex security issues, construction can control, can be control, reliable network is to further promote the development of the premise network applications; the other hand, the network carrying business with the increasing complexity of application-layer security to ensure the development of network security in a new direction.

Encryption technology is a basic safety measures to take e-commerce, transactions both sides can exchange information as needed in phase. Encryption technology is divided into two categories, namely symmetric encryption and asymmetric encryption:(1) Symmetric encryption. Symmetric encryption is also called private key encryption, the message sender and receiver use the same key to encrypt and decrypt data. Its biggest advantage is the encryption / decryption speed, suitable for large data is encrypted, but the key management problems. If the two sides can communicate to ensure the private key in the key exchange stage has not been compromised, then the confidentiality and integrity of messages can be encrypted by this encryption of confidential information, along with the message send the message digest or hash message value to achieve.(2) non-symmetric encryption. Asymmetric encryption, also known as public key encryption, using a pair of keys for encryption and decryption operations are completed, including a public release (ie public key),

another secret kept by the users themselves. Information exchange process is: Party A will generate a pair of keys which a public key to the other parties to the transaction as an open, get the public key of B using the key information is encrypted and then sent to the Party, Party then their information stored private key to decrypt encrypted.

Conclusions

With the rapid development of network technology, the original single stack type network threat protection measures have been difficult to effectively resist the growing hybrid security threats. Construction of a local security, global security, intelligence security, overall security system to provide users with multi-level, comprehensive protection system for three-dimensional construction of a new concept of information security. In this concept, the network security products will undergo a series of changes.

Authentication technology is to use electronic means prove identity of the sender and recipient and file integrity techniques, namely the confirmation of both identity information in the course of transmission or storage is not tampered with :(1) digital signature. Digital signature, also known as electronic signature, handwritten signature as to produce the same electronic documents can play a certification, approval and effective role. This is done through the hash function and public key algorithm combines the sender from the message text to generate a hash value, and with their own private key to encrypt this hash value to form the sender's digital signature; then , this digital signature as the message with the attachment and the message sent to the message recipient; message from the first recipient receives the original packet to calculate the hash value, then using the sender's public key to attach to the message to decrypt the digital signature; If the two hash values are the same, then the receiver can verify that the digital signature is the sender. Digital signatures provide a mechanism to identify ways to address the forgery, denial, pretending, tampering and other issues. (2) Digital certificates. Digital certificate is an authorized center by the digital signature certificate containing the public key owner of the file information, and public key digital certificate constitutes the most important include a user public key, together with the key owner's user identifier, and trusted third-party signature trusted third party is usually the user certification authority (CA), such as government departments and financial institutions. Users in a secure manner to the public key certificate authority and get his public key certificate, and then user can open the certificate. Any person who requires the user public key can get this certificate, signed by the related trust to verify the effectiveness of the public key. Parties to the transaction sign the digital certificate status information through a series of data, provides a way to verify their identity, the user can use it to identify each other's identity.

References

[1] Simmonds, A; Sandilands, P; van Ekert, L. Expression error: Missing operand for “An Ontology for Network Security Attacks". *Lecture Notes in Computer Science* 3285, 2004, pp: 317–323.
[2] Self-Defending Networks: The Next Generation of Network Security, *Duane DeCapite,* Cisco Press, Sep. 8, 2006
[3] Security Monitoring with Cisco Security MARS, *Gary Halleen/Greg Kellogg, Cisco Press, Jul. 6, 2007.*

©(2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.536

Dynamic Response of Supercavitating Underwater Vehicle Impacted by Tail-Slap Force

Qiankun He[a], Jiazhong Zhang[b], Yingjie Wei[c], Cong Wang[d], Wei Cao[e]

School of Astronautics, Harbin Institute of Technology, Harbin 150001, China

[a]heqiankun@foxmail.com, [b]zhangjz@hope.hit.edu.cn, [c]weiyingjie@gmail.com, [d]alanwang@hit.edu.cn, [e]vicdorcao@sohu.com

Keywords: Supercavitation, Underwater Vehicle, Tail-Slap, Dynamic Response, Finite Element Method

Abstract: Based on the movement characteristics of the supercavitating underwater vehicle, the tail impact force has been defined by summarizing related papers, then the dynamic responses of structure has been simulated by using Finite Element Method. The displacement and acceleration responses and corresponding spectrums have been gotten. The calculation results show that: the displacement response along impact force's direction oscillates periodically; the displacement response along body axis lengthens and shortens periodically; the displacement response along the 3rd axis increases over time. The acceleration response along impact force's direction and body axis oscillates periodically. The corresponding spectrums show that: accessories of the underwater vehicle should avoid some scaled frequencies in order to let them run reliably.

Introduction

When the underwater vehicle moves at sufficient speed through water, a low-density gaseous 'cavity' can form. When the underwater vehicle is entirely enveloped by cavities, the flows are called 'supercavitating'. Since the moving body through most of its length does not contact with the fluid, skin drag is almost negligible. Therefore, the underwater vehicle can achieve extremely high submerged speeds[1].

The hydrodynamic forces and cavity shape and the trajectory control for supercavitating underwater vehicles have been modeled by many researchers since many years ago[2-3], and the vehicle is always simplified as rigid body. However, the research of supercavitating underwater vehicle's dynamic characters, where the vehicle is considered as elastic body, is not too many[4-5]. However, any tiny disturbance will make a interaction between the vehicle and the water/cavity interface, and the interaction will make a violent high-frequency vibration of the supercavitating underwater vehicle[6], and this vibration can't be neglected.

The paper has first has summarized the related references and defined the tail impact force. Then, using the tail impact force, the underwater vehicle's dynamic responses and deformation has been solved by Finite Element Method.

1 Analysis Model

Referenced by a projectile's shape, the model of underwater vehicle is established The geometry parameters are defined as follows: underwater vehicle's length $L = 0.18\text{m}$, underwater vehicle's diameter $d = 10\text{mm}$, cavitator's radius $R_n = 2\text{mm}$. And, the material propertys are defined as: underwater vehicle's density $\rho_s = 7800\text{kg/m}^3$, elastic modulus $E = 210\text{GPa}$, Poisson's ratio $\nu = 0.3$, fluid's density $\rho = 1000\text{kg/m}^3$.

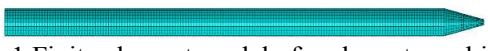


Fig.1 Finite element model of underwater vehicle

Structure dynamic equation is defined as:

$$M\ddot{x} + C\dot{x} + Kx = f \tag{1}$$

Hence, the dynamic responses of underwater vehicle can be solved, if the structure's geometry parameters and the load f is known.

2 Interactions with the Water and the Cavity: Tip Forces and Tail Impact Loads

Any tiny disturbance will make a interaction between the vehicle and the water/cavity interface, and the interaction will produce a huge impact force acting on the underwater vehicle enveloped by a supercavitation. The forces on the underwater vehile enveloped by supercavitation is assumed to be the force by water at the cavitator and the force by water/cavitation interface at the vehile's tail.

To begin with, any applied moment at the nose has been neglected here. Thus, the tip forces of the cavitator are the drag force F_D and the lift force F_L. So, the forces acting on the underwater vehicle(Fig.3) can be summarized as follows: drag force F_D and lift force F_L on cavitator, propulsion force F, R_L and R_D are components of the tail impact force.

The drag and lift force is defined as

$$F_D = \frac{1}{2}\rho A C_D V^2 \ , F_L = \frac{1}{2}\rho A C_L V^2 \tag{2}$$

where, ρ is the fluid density; A is the cross-sectional area of cavitator; V is the moving velocity of underwater vehicle; C_D and C_L is the drag coefficient and lift coefficient. The drag coefficient and lift coefficient for a disk with angle of attack angel is given by Wu's theory [7].

The tail impact force R_L and R_D can be defined as:

$$\begin{cases} R_D = 0 \ , R_L = 0 & \delta \le 0 \\ R_D = \rho A_1 V^2 (1-\cos\theta) \ , R_L = \rho A_1 V^2 \sin\theta & \delta \ge 0 \end{cases} \tag{3}$$

where, θ is the angel between cavitation axis and vehicle axis; $A_1 = \lambda\delta d$ is the cross-sectional area of fluid layer, λ is a coefficient; δ is the penetration distance of the tail into the surrounding fluid[8]:

$$\delta = L\sin\theta + (d/2)\cos\theta - h(y) \tag{4}$$

$h(y)$ is the diameter of cavitation that corresponds to y coordinate, which can be defined as[6]:

$$h(y) = \frac{1}{2}D(y) = \frac{1}{2}D_{\max}[1-(\frac{y-L_c/2}{L_c/2})^2]^{1/2.4} \tag{5}$$

where, $D_{\max}$ and L_c are the maximum radius and the length of the cavitation. They are defined as:

$$D_{\max} = D_n\sqrt{\frac{C_D(\sigma,0)}{\sigma}} \ , \ L_c = D_n\sqrt{\frac{C_D(\sigma,0)}{\sigma^2}\ln(\frac{1}{\sigma})} \tag{6}$$

where, D_n is the radius of cavitator. In Eq. 6, the influence of gravity and lift force to the shape deformation of cavitation is ignored. Hence, R_D, R_L can be written by the function of the attack angel θ of the vehicle.

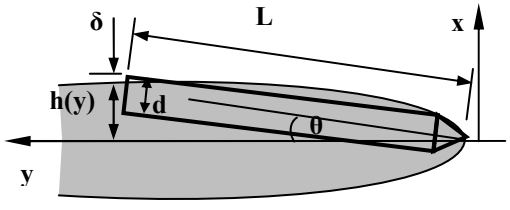


Fig.2 Geometry of impact: underwater vehicle enters the surrounding fluid

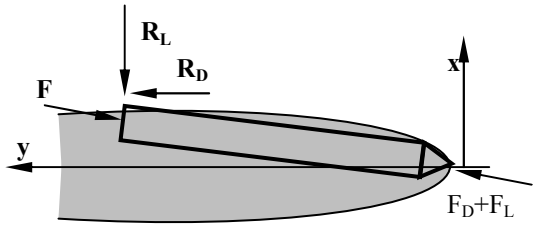


Fig.3 Fluid forces on the underwater vehicle

During the period of motion between successive impacts, the only force on the body is the tip force which is directed along the body axis. The propulsion force, also considered as applied along the

body axis, is assumed to be equaled to the tip force[9], while the supercavitating underwater vehicle is moving stably. By the equilibrium of moments $\sum M = I\ddot{\theta}$, we obtain:

$$I\ddot{\theta} = -R_D a\sin\theta - R_L a\cos\theta \tag{7}$$

where, I is the transverse moment of inertia of the vehicle about its center of mass($I = mL^2/12$ for cylinder), a is the distance between the tail and mass center of the vehicle. Simplification results by replacing θ by ϕ, defined as $\phi = \theta - [(2h-d)/2l]$ (ϕ is little). Ignoring high-level minim variable ϕ^2, the Eq. 7 can be defined as

$$\ddot{\phi} + \omega^2\phi = 0 \tag{8}$$

where, $\omega = (\rho\lambda da/I)(h-d/2)V^2$ is the angular frequency. Thus, the duration of the impact is half the period of the simple harmonic oscillator $T_0 = \pi/\omega$ (from the tail entering the cavity wall the cavity wall to the tail exiting cavity wall)[8].

When the tail have completed impact, it will rebounds to the cavity. Based on the conservation of angular momentum,

$$I\dot{\theta}_t = \int_{-T_0/2}^{T_0/2} \mathrm{M}dt \tag{9}$$

where, $M = -R_D a\sin\theta - R_L a\cos\theta$, $\dot{\theta}_t$ is the angular velocity of vehicle after impact. In the absence of impacts, the duration is

$$T_1 = 2\theta_0/\dot{\theta}_t \tag{10}$$

where, $\theta_0 = \arcsin[h(y)/L]$. If the initial impact depth is supposed to be δ_0, the periodic force can be gotten.

3 Analysis of Simulation Result

Assuming the underwater vehicle is moving stably with the velocity $V = 200\mathrm{m/s}$ and the initial maximum impact depth is $\delta_0 = d/4$, thus, the load f can be defined as Fig.4.

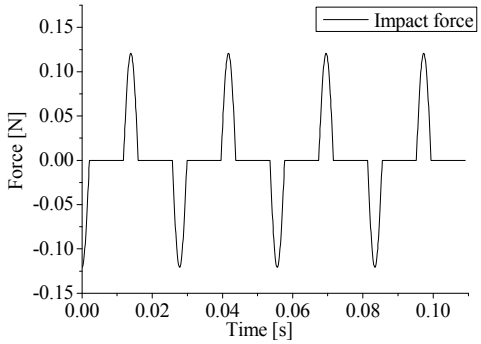


Fig.4 Tail impact load

Using the tail impact load of Fig.4, the finit element model has been simulated with the finit element method, and the structure responses have been gotten.

The displacement responses of tail are shown in Fig.5A-5C. It reveals that: displacement along x axis oscillates periodically, along y axis lengthens and shortens periodically and along z axis increases periodically. Hence, the tail-slap, simplified to move in xoy plan, is valid in earlier stage of vibration. Fig.6A-6C have shown the spectrum of displacement along x and y axises. It reveals that: the displacement along x axis is mainly distributed in the frequency of 36Hz, 107.5Hz, 180Hz; along y axis is mainly distributed in the frequency of 36Hz, 73Hz, 145.5Hz, 227Hz; and, along z axis is mainly distributed in the frequency of 36Hz, 895Hz. Where, the frequency of 36Hz contain the diplacement superposition of regid displacement and elastic displacement, and it mainly reveals the rigid displacement of underwater vehicle. Therefore, the natural frequence of accessories working on underwater vehicle should keep away from these values, in order to avoid the resonance.

The acceleration responses of tail are shown in Fig.7A-7B. It reveals that: acceleration along x and y axis are all oscillate periodically. Fig.8A-8B have shown the spectrum of acceleration along x and y axises. It reveals that: the acceleration along x axis is mainly distributed in the frequency of 36Hz, 107Hz, 178Hz, 249Hz, 886Hz; along y axis is mainly distributed within the frequency of 300Hz and nearby 900Hz. Also, the natural frequence of some accuracy accessories and easily destroyed accessories working on underwater vehicle should keep away from these values, in order to avoid the vibration damage effect.

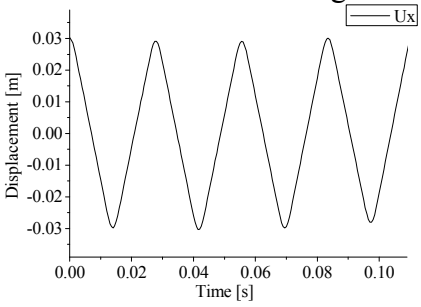


Fig.5A Displacement response along x axis

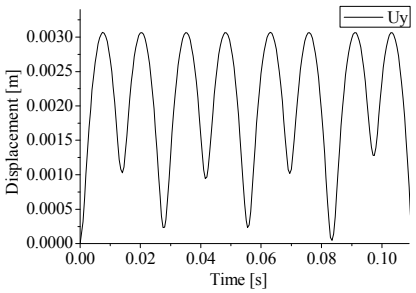


Fig.5B Displacement response along y axis

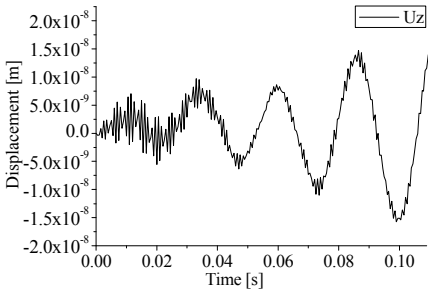


Fig.5C Displacement response along z axis

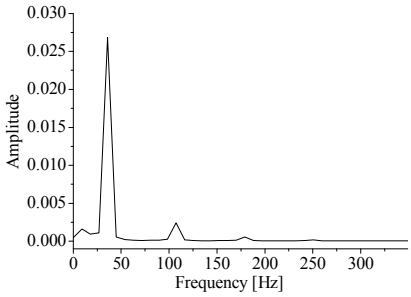


Fig.6A Spectrum of displacement response along x axis

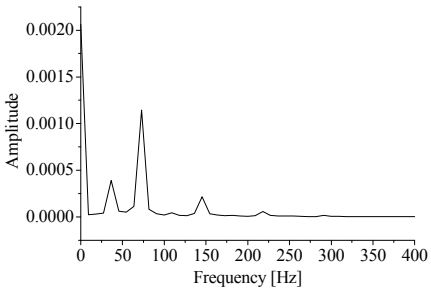


Fig.6B Spectrum of displacement response along y axis

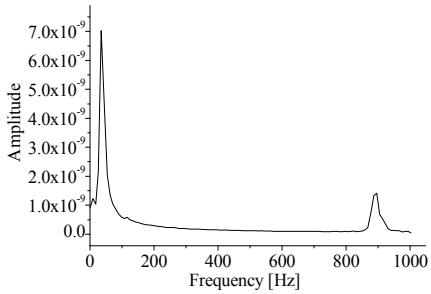


Fig.6C Spectrum of displacement response along z axis

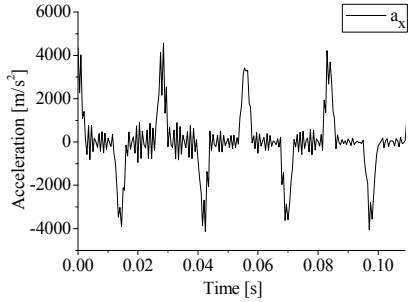


Fig.7A Acceleration response along x axis

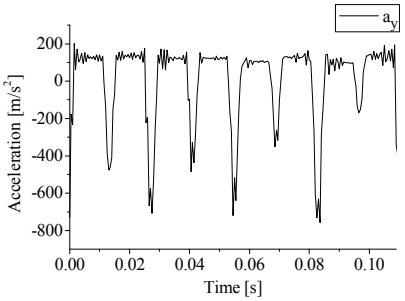


Fig.7B Acceleration response along y axis

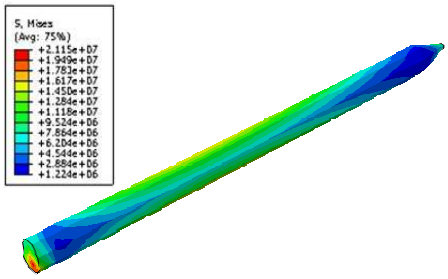


Fig.9 Mises stress during impact

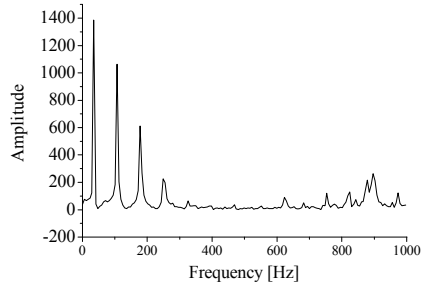


Fig.8A Spectrum of displacement response along x axis

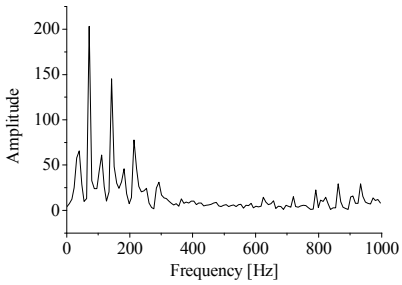


Fig.8B Spectrum of displacement response along y axis

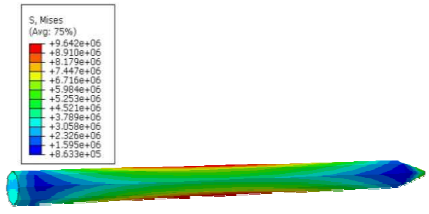


Fig.10 Mises stress between two successive impacts

Based on the conclusions, the frequency of accessories of underwater vehicle is required to be kept away from the value equal or less than 300Hz and the value nearby 900Hz in order to keep running reliably.

The mises stress of underwater vehicle has also been computed, Fig.9 and Fig.10 show the mises stress during impact and between two successive impacts respectively. It reveals the stress is mainly distributed concentratively over the body's tail during impact and the value increases instantly, but when the vehicle has completed impact successful, the stress is mainly distributed evenly over the middle part of vehicle.

Conclusions

The dynamic behavior of supercavitating bodies operating in periodic tail impact conditions is analyzed here. A numerical model is developed to simulate the body motion resulting from the impacts of the tail of the underwater vehicle with the water/cavity interface. Besides, the tail impact force of the model has been defined. The simulation results reveal that:

1. The displacement along x and y axis oscillates periodically under a periodic impact force, the displacement along z axis increases periodically.
2. The impact can't be simplified in one plan after long time impacting.
3. The frequency of accessories of underwater vehicle is required to be kept away from the value equal or less than 300Hz and the value nearby 900Hz in order to keep running reliably.
4. The mises stress is mainly distributed concentratively over the body's tail during impact, and distributed evenly over the middle part of vehicle during the interspace between two successive impacts.

Reference

[1] D. Miller: Supercavitation: Going to War in a Bubble. *International Defense Review*, (1995), p. 61-63

[2] A. Goel: Robust Control of Supercavitating Vehicles in The Presence of Dynamic and Uncertain Cavity Florida: *University Of Florida*. (2005).

[3] M. Ruzzene, R. Kamada, C. L. Bottasso and F. Scorcelletti: Trajectory Optimization Strategies for Supercavitating Underwater Vehicles. *Journal of Vibration and Control*, Vol. 14(5) (2008), p. 611-644

[4] J.Y. Choi, M.Ruzzene and O.A. Bauchau: Dynamic Analysis of Flexible Supercavitating Vehicles Using Modal-Based Elements, *SIMULATION*, Vol 80(11) (2004), p. 619-633

[5] I.N. Kirschner, B.J. Rosenthal and J.S. Uhlman: Simplified Dynamical Systems Analysis of Supercavitating High-Speed Bodies. *Fifth International Symposium on Cavitation* (CAV2003) Osaka, Japan, (2003).

[6] I.N. Kirschner, D.C. Kring, A.W. Stokes, N.E. Fine and J.S. Uhlman: Control strategies for supercavitating vehicles. *Journal of Vibration and Control.* Vol. 8 (2002), p. 219-242

[7] Wu, T. Y-T. A Free Streamline Theory for Two-Dimensional Fully Cavitated Hydrofoils. J.Math and Phys. Vol. 35 (1956), p. 236-265

[8] R. Rand, R. Pratap, D. Ramani, J. Cipolla and I. Kirschner: Impact Dynamics of A Supercavitating Underwater Projectile Proceedings of DETCp97. California: *ASME,* (1997), p. 1-11

[9] M. Ruzzene and F. Soranna: Impact Dynamics of Elastic Stiffened Supercavitating Underwater Vehicles *Journal of Vibration and Control.* Vol. 10 (2004), p. 243-267

©(2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.541

Normal Stress between Steel Wires in the Stay-cable

Gang Zheng[1,a] ,Hong Li[1,2,b*]

[1]Chongqing Communication Research & Design Institute, Chongqing 400067, China

[2]Department of Civil Engineering, Chongqing Jiaotong University, Chongqing 400074, China

[a]zhenggang@cmhk.com, [b]hongli_1129@163.com

* Corresponding author

Keywords: normal stress; steel wire; stay-cable

Abstract. The normal stress between wires is one of the main factors of the slipping and stress concentration in the stay-cables. In order to improve the research on the stay-cables, the mechanism of the normal stress has been studied in this paper. The normal stress can be resulted from two reasons. Firstly, the normal stress generate from the semi-parallel wires because of the existence of spiral angle. Secondly, the normal stress results from the pinch-effect generated by the hot-extruding PE protection. Both the formulation have been listed to analyse the normal stress in this paper. Result shows, the normal stress between the outermost layers is far smaller than the inner layers. Finally, the field survey surveys have been used to verify the analysis.

1. Background

In recent years, most studies on the flexural failure of stay-cable around the world are based on the theoretical assumption that the stay-cable has a constant flexural stiffness. According to this assumption, the cross section of cable keeps a constant during the loading procedure [1]. This hypothesis, however, could not fulfill the practical demand of the calculation of bending stress. In the literature [2], Professor Gimsing stated that the assumption of the uniform cross-section is only valid when the steel wires don't slide between each other [2]. Otherwise, the effect of frictional stress between steel wires must be taken into consideration during the analysis of stay-cable's bending property. The frictional stress determines whether the steel wires will slide easily or not, further influence the magnitudes of the local bending stresses.

As a general contact problem, the slide between steel wires can be assumed as Mohr-Coulomb failure. In this problem, the friction factor can be confirmed by test, so the key procedure is accurately calculating the normal stress between the steel wires when they are subjected to the longitudinal force. After that, the friction stress between wires can be attained, and the slide state of stands will be confirmed. In this paper, the normal stress will be attained by mathematical calculation, and an example is used to discuss the slide state in the stay-cable.

The general manufacturing process of stay-cable lists as follows. First, high-strength, cold-worked steel wires are put in rows after cut with specific size. After that the steel wires are twisted slightly and twined with spiral steel wires or fiber-reinforced polyester as steel wires. The steel wires then are coated with polyethylene jackets by plastic extruding machine and cut precisely according to the design. The cables will be finished after installing the matched chill anchor on both sides. The cross section of cable is closely arranged with normal hexagon or unfilled corner hexagon, made by slightly left twisted with a angle of 2° ~4° (as shown in Figure.1). [4]

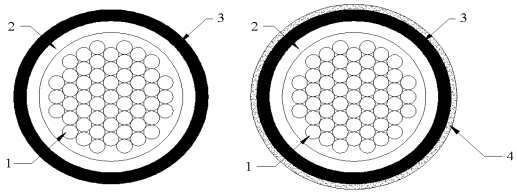


Fig. 1 Cross Section of Typical Stay Cable

1- steel wire,
2- circumvoluted small-gauge wire or fibre-reinforced polyester,
3- black polyethylene jacket,
4- multi-colour polyethylene jacket.

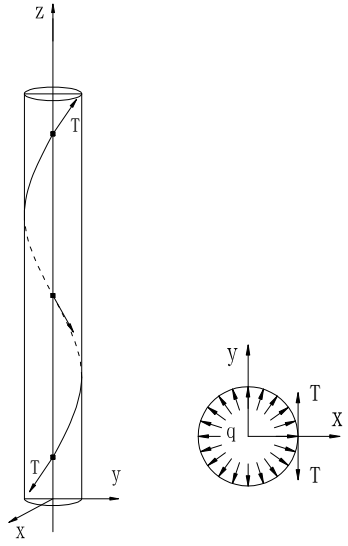
Because of the existence of spiral angle, when the stay-cable is subjected to the longitudinal force, the pinch-effect will be generated. Also, the normal stress will be caused by the change of temperature during the process of hot-extruding PE wrapping.

2. Pinch-effect generated by the semi-parallel steel wire's spiral angle

When shaping the cable, the steel wires are twisted slightly to make sure each spiral wire is coiled with the inner wire. Due to the spiral angle, the normal stress will be generated when a tangential tensile force T applied on the cable. Obviously, because the normal stress between the concentric steel wires is far less than the one between non-concentric steel wires, the calculation of normal stress at the outermost layer will be discussed particularly in this paper.

When the cable is subjected to a tangential tensile force, the normal stress among steel wires will transfers as linear uniformly distributed load, we assume it as q (KN/m). Based on the assumptions above, take the outermost wire as an example , shown in Figure 2 (a). When the tangential tensile force T is applied, the radial normal stress will occurred between the wire and the cylinder as Figure 2 (b).

We set the pitch of helix is *L*, the radius of the cylinder is r, the lay angle of helix is t, then the following equation can be attained:



(a) (b)

Fig.2 The force diagram of single spiral wire

$$\begin{cases} x = r\cos t \\ y = r\sin t \quad ,t \in [0,2\pi] \\ z = L \cdot t/2\pi \end{cases} \tag{1}$$

To differential the equation (1), we can gain:

$$\begin{cases} \dfrac{dz}{dx} = \dfrac{z'}{x'} = -\dfrac{L/2\pi}{r \cdot \sin t} \\ \dfrac{dz}{dy} = \dfrac{z'}{y'} = \dfrac{L/2\pi}{r \cdot \cos t} \end{cases} \tag{2}$$

2.1 Normal stress of single helix subjected to longitudinal force

Take a segment $\triangle s$ for mechanical analysis, shown in Figure 3[6]:

The tension forces acted on the ends of the segment can be listed in the vector way as following:

$$\vec{T} = \frac{2\pi \cdot T}{\sqrt{4\pi^2 r^2 + L^2}}\left(-r\sin t, r\cos t, \frac{L}{2\pi}\right)$$

$$\vec{T}' = \frac{2\pi \cdot r}{\sqrt{4\pi^2 r^2 + L^2}}\left(r\sin(t+\Delta t), -r\cos(t+\Delta t), \frac{L}{2\pi}\right)$$

Then the vector expression of the equation is:

$$-\vec{q} \cdot \Delta s = \vec{T} + \vec{T}'$$

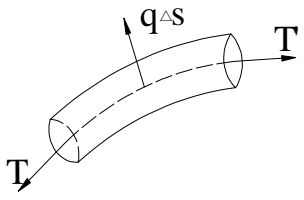


Fig. 3 The force diagram of the infinitesimal Δl

And the normal stress on Δs is:

$$|\vec{q}\cdot\Delta s| = \frac{2\pi\cdot T}{\sqrt{4\pi^2r^2+L^2}}\left|(r\sin(t+\Delta t)-r\sin t, r\cos t - r\cos(t+\Delta t),0)\right|$$

$$= \frac{2\sqrt{2}\pi\cdot Tr}{\sqrt{4\pi^2r^2+L^2}}\sqrt{1-\cos(\Delta t)}$$

The distributed load can be calculated as follows:

$$|\vec{q}| = \lim_{\Delta s\to 0}\frac{2\sqrt{2}\pi\cdot Tr}{\sqrt{4\pi^2r^2+L^2}}\frac{\sqrt{1-\cos\Delta t}}{\Delta s} = \lim_{\Delta s\to 0}\frac{2\sqrt{2}\pi\cdot Tr}{\sqrt{4\pi^2r^2+L^2}}\frac{\Delta t}{\sqrt{2}\Delta s} = \frac{2\pi\cdot Tr}{\sqrt{4\pi^2r^2+L^2}}\frac{dt}{ds}$$

Then substitute Equation (2) into the above equation:

$$|\vec{q}| = \frac{4\pi^2 Tr}{4\pi^2r^2+L^2} \tag{3}$$

Sameresults have been attained use integration method and differentiation method respectively. Also the following relationship can be attained between *a* and the pitch of helix L:

$$\tan\alpha = \frac{\mathrm{L}}{2\pi\cdot r}$$

Then, Equation(3)can be expressed as：

$$q = \frac{T}{r(1+\tan^2\alpha)} = \frac{\mathrm{T}}{r}\cdot\cos^2\alpha \tag{4}$$

2.2 Normal stress between each steel wires layers subjected to longitudinal force

The cross section of stay-cable is arranged by normal hexagon or unfilled corner hexagon multi-layer steel wires (as shown in Figure.1). If the normal stress of the steel wires can be fully transmitted from the outer layer to inner layer, we can calculate it layer-to-layer. But, the normal hexagon arrangement makes it impossible to calculate it by the above method due to the different spread pattern. So in the simplified calculation, we take the average value of radius as the helical radius.

Suppose the number of wire layers arranged in the cross section of the stay-cable is *n*, the tension applied at the both ends of each wire is *T*, set the outermost wire layer as the first layer, then the average radius of the N-layer steel wires would be:

$$r_i = \frac{n-2i}{2}\cdot d, (i=1,2,3,\ldots,\frac{n-1}{2}) \tag{5}$$

Substitute Equation (5) into Equation (4), the normal pressure under the tension T at both ends will be:

$$q_{i0} = \frac{T}{r_i}\cdot\cos^2\alpha \tag{6}$$

The pressure of the first layer is :

$$q_1 = q_{10} = \frac{2\mathrm{T}}{(n-2)\cdot d}\cdot\cos^2\alpha \ ,$$

The pressure of the Nth layer is

$$q_i = \sum_{k=1}^{i-1} q_{k0}, (i = 2,3,\ldots,\frac{n-1}{2}) \tag{7}$$

3. Pinch-effect generated by the hot-extruding PE protection

By using the extrusion technology, parallel wire cables are coated with black high density polyethylene jacket. During the process the polyethylene jacket will deform due to the heat-expansion and cold-contraction, which will result in the pinch-effect to the parallel steel wires.

We denote the linear thermal expansion coefficient of high density polyethylene under the room temperature is a .So, $\alpha = \frac{1}{L}\cdot\frac{\Delta L}{\Delta t}$ .Then the strain of unit volume polyethylene jacket generated by the temperature variation Δt is $\alpha\times\Delta t$, and the stress is $\alpha\cdot E\cdot\Delta t$, where E is modulus of elasticity of polyethylene. Take a unit length of cable, if the extrusion pressure of the unit-length wire generated by the polyethylene jacket was q_0, then:

$$q_0 = 2\pi\cdot r\cdot\alpha\cdot E\cdot\Delta t \tag{8}$$

4. Example

The cable which consists of 21 layers and 283 steel wires is arranged by unfilled corner hexagon, and slightly left twisted and the lay angle is 3° (as shown in Figure 4). The tensile strength of steel wire is $\sigma_b = 1670MPa$ .And the tensile force $T = 0.4A\sigma_b$ is applying at the both cable ends.

The temperature variation of polyethylene jacket caused by extrusion processing process is about 24 ℃. The elongation rate is 25%, tensile elasticity modules ≥150MPa.[2] Then the extrusion pressure transferred to the steel wires arranged in each layers can be calculated from Eq.(8) is about 35.578KN/m. The normal pressure between layers can be calculated by added the two extrusion effect together. The following table lists the maximum frictional force (static friction) the wire can afford:

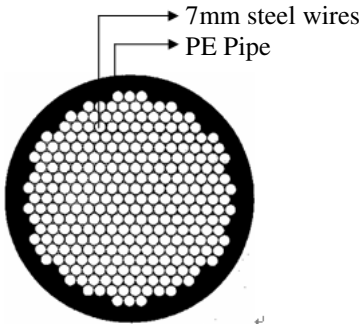


Figure 4 Cross-section of cable

Table 1. The normal strsee and the maximum static friction steel wires can resist

Layer No.(i)	Normal stress caused by the lay angle ((KN/m)	Normal stress caused by the PE protection (KN/m)	Frictional factor*	Maximum static friction (KN/m)
1	378.88	35.578	0.211	87.42
2	802.34	35.578	0.211	176.74
3	1282.26	35.578	0.211	277.97
4	1836.01	35.578	0.211	394.77
5	2490.44	35.578	0.211	532.81
6	3290.31	35.578	0.211	701.52
7	4318.70	35.578	0.211	918.44
8	5758.45	35.578	0.211	1222.12
9	8158.04	35.578	0.211	1728.26
10	15356.81	35.578	0.211	3246.68

* Measured by the friction test

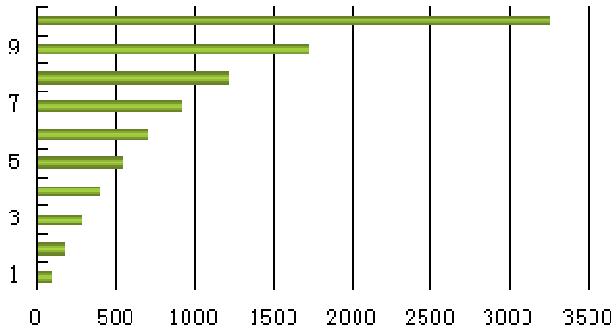


Fig.5 The tendency chart of the static friction force between layers

5. Conclusion

Normal stress will be generated due to the spiral angle of wires and pinch-effect of hot-extruding PE wrapping when the stay-cable is under longitudinal force, it will ultimately affect the local cable bending stresses. From the results of the above analysis, we can conclude that normal stress are mainly result from the pinch-effect because of the existence of the spiral angle, which is directly determined by the manufactory process and the tensile force of semi-parallel cable. That means the compressive stress is directly proportional to the longitudinal force and the square of the cosine of spiral angle, and inversely proportional to the diameter of cable.

It can be seen from Figure 5 that the maximum static friction between inner steel layers is much larger than that between the outer layers. Thus, the outer layer wire is much easier to be broken and slid than the inner layer.

The steel wires at the six corner points of the hexagonal array may be supported by the neighboring ones, so its compressive stress cannot be fully passed into the inner steel wires. Therefore, compressive stress suffered by these steel wires may be smaller than the others; the maximum static friction will be smaller as well, so they will be prone to slippage or broken.

The two conclusions above can be verified by the fatigue test of the parallel wire cables. The fractured part of the wire at the anchorage is almost always locates in the outermost row of the cable cross-section, as is shown in Figure 6. A signal slip has been found at the outermost row [6], which can be explained by the slip hypothesis listed in this phenomenon.

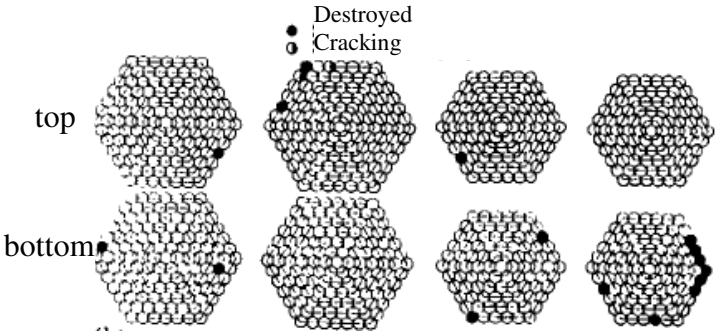


Fig.6 The fractured part of the wire at the anchor position

References

[1] Eggers J.C., Cable Stay Analysis for the Fred Hartman Bridge[M.S.Thesis],Austin,the University of Texas at Austin,2003

[2] Eggers J.C.,Translated by Yao Lingsen and Lin Changchuan, Cable-supported Bridges--Concept and Design(the second edition), Renmin Communication Press,2002

[3] X.Y.Zou, S.B.Li, The construction and application for Technical Conditions for Hot-extruding PE Protection Figh Strength Wire Cable-stayed Bridge, Highway, No.11, 2002

[4] National Standard, Technical conditions for hot-extruding PE protection high strength wire cable-stayed bridge GB/T 18365-2001, 2001

[5] Y.Z.Liu. Nonlinear mechanics of thin elastic rod. Tsinghua University Press. 2006

[6] Translated by Y.J.Xiang, The fatigue strength of large diameter stay-cable, Foreign bridges, No.1, 1994

©(2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.547

Experimental Evaluation of the Seismic Performance of Circular Reinforced Concrete Bridge Columns

Gang Zheng [1,a] Guiqian Li [1,2,b]

1. Chongqing Communications Research & Design Institute, Chongqing 400067, China；

2. Guangxi Communications Planning Surveying and Designing Institute, Nanning，530011, China

a.zhenggang@cmhk.com, b.lgqmancan@163.com

Key words: r/c bridge columns, orthogonal quasi-static test, ductility capacity, experimental evaluation

Abstract: Based on the basic requirements of current *Guidelines for Seismic Design of Highway Bridges*, the orthogonal quasi-static test of four factors (shear-span ratio, longitudinal reinforcement diameter, axial-load ratio and spiral reinforcement ratio) at three different levels for circular reinforced concrete bridge columns has been designed. With test data the damage state, displacement ductility, capacity of accumulative energy dissipation to ultimate displacement state of bridge columns subjected to low-cyclic loading have been analyzed systematically so as to investigate effects of factors such as shear-span ratio, axial-load ratio, longitudinal reinforcement ratio and spiral reinforcement ratio on ductility performance of bridge columns.

Introduction

With the rapid development of seismic techniques on bridges, the main bridge seismic design codes in the world have adopted ductility design. In comparison with the *Specifications of Earthquake Resistant Design for Highway Engineering* (JT004-89) which design based on quasi-strength and ductility seismic details are inadequately[1], the most important progress in current *Guidelines for Seismic Design of Highway Bridges* of China (JTG/T B02-01-2008) lies in introduction of the idea of ductility based seismic design[2]. In ductility based seismic design of reinforced concrete bridges, the hysteresis energy dissipated capacity and ductility performance of columns decide the integral seismic capability of bridges to some extent. However, as of now researches on reasonability of requirements for reinforcing bars and structural details of columns in the *Guidelines for Seismic Design of Highway Bridges* are particularly scarce. Therefore, it is necessary to carry out systematic and deep research on seismic performance and ductility damage mechanism of reinforced concrete bridge columns in compliance with the requirements in existing seismic codes for bridges.

Design of Orthogonal Quasi-static Test for Bridge Columns

Design of model specimens of columns

The quasi-static test for reinforced concrete bridge columns adopts orthogonal design of four factors at three different levels (3 series, 9 model specimens of columns), and mainly investigates the effects of factors such as shear-span ratio (L/D), longitudinal reinforcement diameter (d_b), axial-load ratio ($P/A_g f_c'$) and spiral reinforcement ratio (ρ_h), etc. on seismic performance of bridge columns. The specimens satisfy the relevant provisions and requirements in the *Guidelines for Seismic Design of Highway Bridges* (2008) in reinforcing bars and structural details. The summary of parameters for model specimens of columns is shown in Table 1,and the layout of strain gauges for longitudinal reinforcement and spiral reinforcement is show in Fig.1,where L_p is the plastic-hinge length in reference [2]. In addition, it should be pointed out that Column 7024 has many holes because of uncompacted vibration of concrete in bottom area of the column during construction, and these holes are only made up by means of cement slurry in later period.

Table 1 Summary of Parameters for Three-level Orthogonal Model Specimen Design

Row	1	2	3	4				
Factor / Column No.	L/D	d_b (mm)	$P/A_g f_c'$	ρ_h (%)	ρ_h' (%)	ρ_l (%)	Length (cm)	Spiral spacing (mm)
4508	4.5	10	0.1	0.8	0.4	0.8	180	63
4515	4.5	14	0.2	1.0	0.6	1.5	180	50
4524	4.5	18	0.3	1.2	0.9	2.4	180	42
5708	5.7	10	0.2	1.2	0.6	0.8	230	42
5715	5.7	14	0.3	0.8	0.8	1.5	230	63
5724	5.7	18	0.1	1.0	0.4	2.4	230	50
7008	7.0	10	0.3	1.0	0.7	0.8	280	50
7015	7.0	14	0.1	1.2	0.4	1.5	280	42
7024	7.0	18	0.2	0.8	0.7	2.4	280	63

Note: the diameter of all columns is 40cm, the concrete is C30, the number of longitudinal reinforcements is 12, the diameter of spiral reinforcement is 8mm. and ρ_h' is minimum spiral reinforcement ratio in reference [2].

Material properties

The moment-curvature analysis of section must take the strain strengthening effect of longitudinal reinforcements and confinement effect on core concrete by spiral reinforcement into account, the mechanical parameters for concrete and reinforcement should adopt measured values. The strengths of concrete for base of the column and column specimens are 33.9MPa and 31.9MPa respectively, reinforcements include 4 types with diameters of 8mm, 10mm, 14mm and 18mm, of which, the 8mm reinforcement is plain bar for spiral reinforcement, and the others are all rebar for longitudinal reinforcements. The material characteristics of reinforcements are shown in Table 2.

Table 2 Properties of Reinforcements

Diameter (mm)	8	10	14	18
Yield strength(MPa)	278	394	332	374
Ultimate strength (MPa)	451	532	545	606
*Yield strain	0.0013	0.0020	0.0017	0.0019
Ultimate strain	0.281	0.244	0.225	0.197

*** Obtained via calculation by taking 2.1×10^5MPa as elastic modulus for 8mm reinforcement and 2.0×10^5MPa as elastic modulus for the others.**

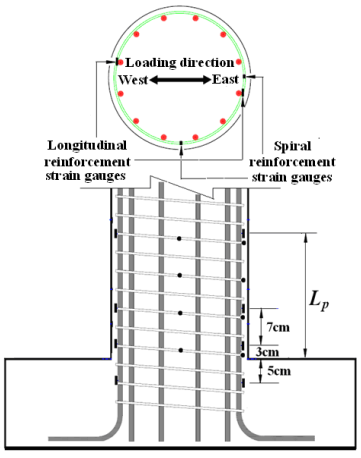


Fig.1 Layout of Strain Gauges for Longitudinal Reinforcement and Spiral Reinforcement

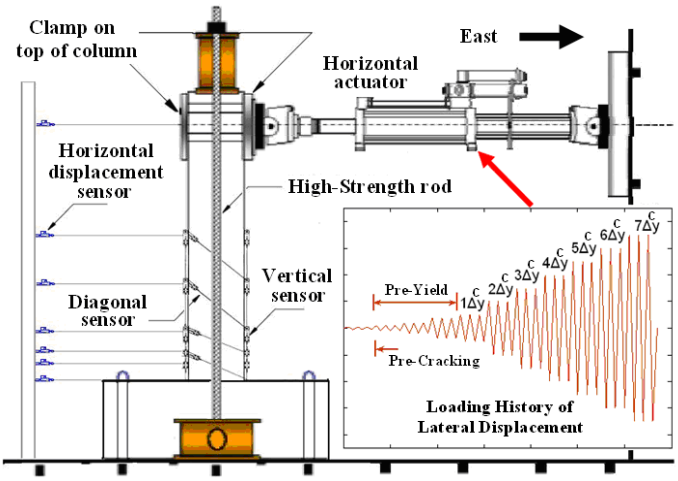


Fig.2 Overall Layout of Test Setup and Loading History of Lateral Displacement

Test Setup and Loading Mechanism

The test is carried out in State Key Laboratory of Bridge Structural Dynamics, Chongqing

Communications Research & Design Institute, the integral test setup is shown in Fig.2. In the test the vertical load is first of all applied as per the axial-load ratio and kept constant, and then the cyclically horizontal displacement load is applied. The loading history of displacement can be divided into three stages: pre-cracking, per-yield and after yield, and the curve of loading history of displacement is shown in Fig.2. Of them the displacement levels pre-cracking of column-bottom section of column are calculated based on elastic theory, the displacement levels per-yield include displacement level from two cracks to initial yield of longitudinal reinforcements and the displacement level of initial yield of longitudinal reinforcement, and the displacement levels after yield are determined based on multiples of theoretical yield displacement Δ^c_y. Three cycles are loaded for each displacement level. The mark for failure to specimens is that the lateral bearing capacity decreases below 85% of the maximum value, or longitudinal reinforcements are fractured, or core concrete is crushing in large area.

Before determination of loading history of lateral displacement, the moment-curvature analysis of section for various columns is carried out by XTRACT software so as to calculate corresponding initial yield of longitudinal reinforcement displacement Δ_{yfirst} and theoretical yield displacement Δ^c_y. The theoretical yield displacement Δ^c_y is composed of flexual deformation, shear deformation and bond-slip deformation caused by strain penetration effect, the calculation equations are as follows[3]:

$$\Delta_{yfirst} = \frac{L^2}{3}\phi'_y \qquad (1)$$

$$\Delta^c_y = \Delta_{flex} + \Delta_{shear} + \Delta_{slip} \qquad (2)$$

$$\Delta_{flex} = \frac{L^2}{3}\phi_y \qquad (3)$$

$$\Delta_{shear} = \frac{V_y}{k_v A G_{eff}} \times L \qquad (4)$$

$$\Delta_{slip} = \frac{d_b f_y \phi_y}{8u} \times L \qquad (5)$$

Where, L - height of column,φ_y' – curvature corresponding to the first yield of the longitudinal steel, φ_y – effective yield curvature, V_y – lateral force corresponding to equivalent yield bending moment M_y, k_v – shape factor (0.85 for circular column), A – gross section area of column, G_{eff} - effective shearing modulus, d_b – diameter of longitudinal reinforcement, f_y – yield strength of longitudinal reinforcement, u – average bonding strength of reinforcement-concrete (approximated into $0.8\sqrt{fc'}$ MPa).

The loading history of lateral displacement of specimens for various columns is determined as Table 3 in accordance with the idea of displacement loading based on three stages:

Table 3 Loading History of Lateral Displacement (Unit: mm)

Series	Column No.	Before crack	Before yield	Before yield	Initial yield	$1\Delta^c_y$	$2\Delta^c_y$	$3\Delta^c_y$	$4\Delta^c_y$	$5\Delta^c_y$	$6\Delta^c_y$	$7\Delta^c_y$
Col1	4508	1	3	6	9	12	24	36	48	60	72	84
	4515	1	3	5	8	14	28	42	56	70	84	98
	4524	1	3	6	9	17	34	51	68	85	102	119
Col2	5708	1	4	7	14	20	40	60	80	100	120	140
	5715	2	5	10	15	22	44	66	88	110	132	154
	5724	1	4	7	14	25	50	75	100	125	150	175
Col3	7008	3	8	16	24	32	64	96	128	160	192	224
	7015	2	4	9	19	29	58	87	116	145	174	203
	7024	2	6	13	24	36	72	108	144	180	216	252

Test Results and Evaluation

Damage state and hysteresis response of columns

All three series columns exhibit obvious flexual damage state, and the progression of damage for all columns are similar. Damage phenomena can be summarized as follows: (1) Concrete on the bottom of columns cracks first, flexual cracks increase, and space reduces. (2) Original cracks expand, width of cracks enlarges, flexual cracks development appears to be stable.(3) Flexual cracks on both sides show diagonal expansion and cover concrete start to spall. (4) Spalling area gradually enlarges along column height and arc surface, spiral reinforcements and longitudinal reinforcements become exposed. (5) Spiral reinforcements in bottom area of columns show obvious yield deformation, and longitudinal reinforcements begin to buckling. (6) Core concrete shows sign of crushing, and longitudinal reinforcements are fractured after serious buckling. Fig.3 gives progression of damage for Column5724, and Fig.6 shows force-displacement hysteresis curve and marks of various damage states for each column.

To progression of damage for columns, crushing and spalling of core concrete for columns with longitudinal reinforcements of 10mm diameter are not obvious, the failure of column is controlled by fracture of longitudinal steels. When columns with longitudinal reinforcements of 14mm or 18mm diameters are in the state of ultimate displacement, crushing and spalling of core concrete are rather obvious, the failure of column is controlled by fracture of longitudinal reinforcements and crushing of core concrete. From hysteresis curve of columns, the columns with low longitudinal reinforcement ratio and high axial load show obvious pinching, which indicates they are influenced by a certain shearing mechanism. And the columns with small diameter of longitudinal reinforcements can bear fewer numbers of cyclic loading when longitudinal steels fracture.

Fig.3 Progression of Damage for Column 5724

Displacement ductility of columns

The displacement ductility is an important measurement index for seismic performance of columns. The yield displacement of columns is estimated based on Park method[4] (as shown in Fig.4, C is the yield point), the ultimate displacement state corresponds to lateral bearing capacity on its force-displacement backbone curve decreasing below 85% of the maximum value (if not, the initial fracture of longitudinal reinforcement is taken as the ultimate displacement state), then the displacement ductility factor corresponding to the ultimate state of columns is shown in Table 4. It is observed that the columns under test exhibit excellent ductility performance (however, the displacement ductility of Column7024 is relatively bad since it is damaged earlier during test due to uncompacted vibration of concrete in bottom area of columns), and it is indicated that the relevant requirements and provisions in reference [2] are reasonable. The results of orthogonal analysis show that the axial-load ratio and longitudinal reinforcement ratio greatly influence the displacement ductility factor, which becomes smaller with increase of the axial-load ratio and longitudinal reinforcement ratio. In the mean time Table 4 also gives the comparison between theoretical yield

displacement Δ^c_y and experimental yield displacement Δ_y, the mean value and standard deviation of Δ^c_y/Δ_y are 1.01 and 0.25 respectively. Based on theoretical calculation and test results, the loading history of displacement stipulated is reasonable, and the estimate method for experimental yield displacement is reliable.

Table 4 Displacement Ductility Factor of Test Bridge Columns

Series	Column No.	Theoretical yield displacement Δ^c_y (mm)	Experimental yield displacement Δ_y (mm)	Experimental ultimate displacement Δ_u (mm)	Displacement ductility factor μ_Δ
Col1	4508	12.3	11.5	96.1	8.4
	4515	13.6	16.9	115.9	6.9
	4524	16.5	20.3	153.0	7.5
Col2	5708	20.1	17.5	129.5	7.4
	5715	22.1	23.1	161.7	7.0
	5724	24.5	31.5	220.9	7.0
Col3	7008	31.0	18.5	141.8	7.7
	7015	29.3	29.6	248.0	8.4
	7024	36.2	39.8	211.5	5.3

Properties of hysteretic energy

The properties of hysteretic energy reflect capability of columns dissipating external energy to resist seismic force and are quite important for seismic resistance of bridge ductility. The hysteresis energy dissipation of columns is defined as the area encompassed by enclosed hysteresis loop in force-displacement curve, the more the area of hysteresis loop, the stronger the capability of columns dissipating energy. The hysteresis energy dissipation of tested columns has the following feature: (1) At the same displacement level, the energy dissipated by the latter cycle is smaller than that dissipated by the former cycle due to accumulation of damages, this feature is more obvious especially in later period of loading. (2) When longitudinal reinforcements buckling, the energy dissipated by columns dramatically decreases and plunges especially when longitudinal reinforcements fracture.

Fig.5 gives normalized graphic representation of the accumulative dissipated energy to ultimate displacement state. From Fig.5, it can be seen that the accumulative dissipated energy is directly related to longitudinal reinforcement ratio (except Column7024, which is damaged too early due to uncompacted vibration in bottom area of columns), while is not obviously influenced by axial-load ratio and spiral reinforcement ratio.

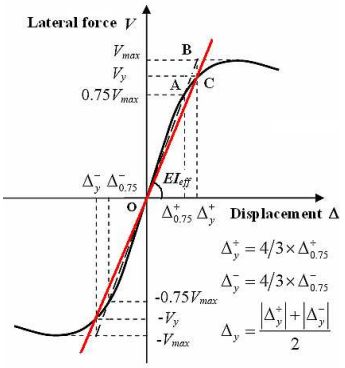


Fig.4 Definition of Yield Displacement

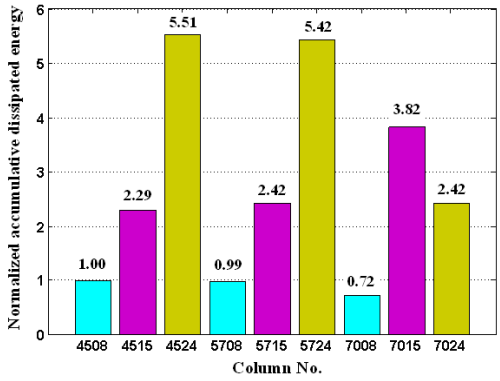
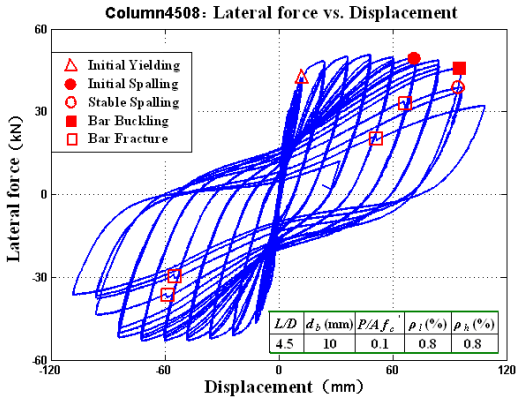
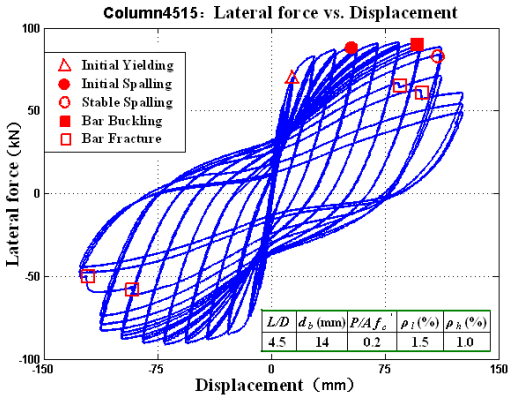
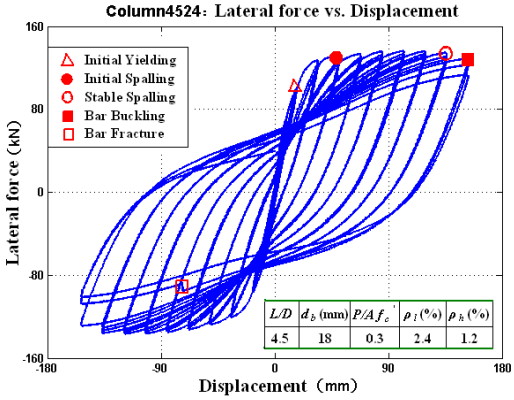
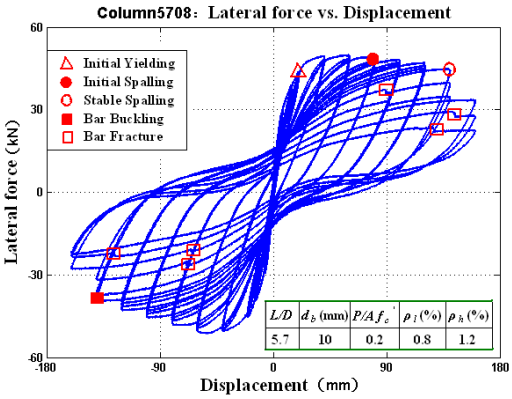


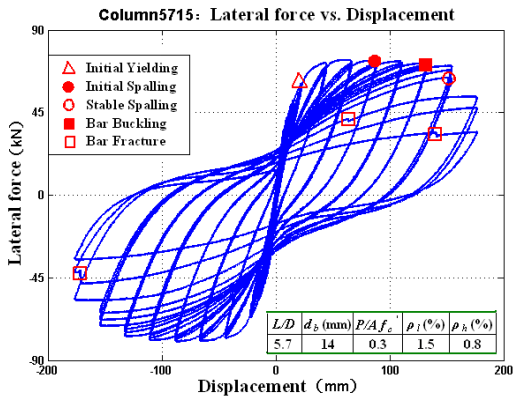
Fig.5 Normalized Accumulative Dissipated Energy

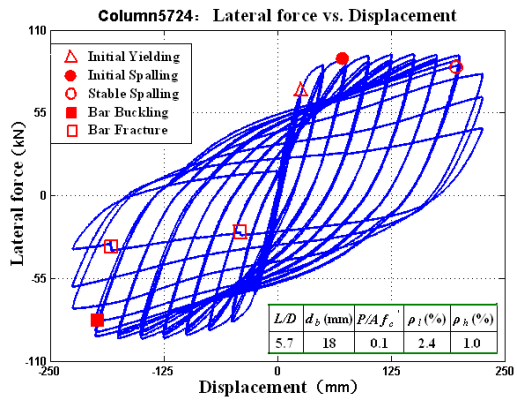


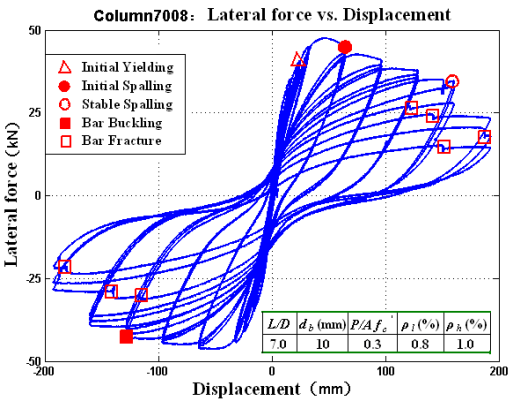


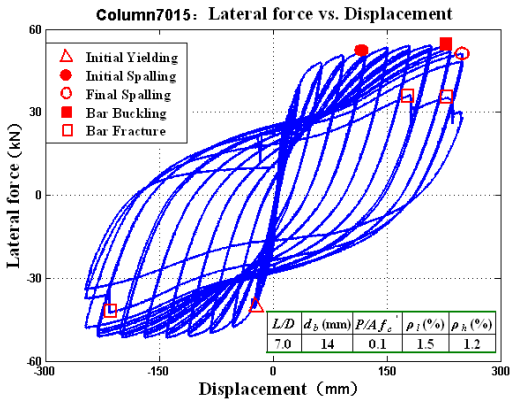












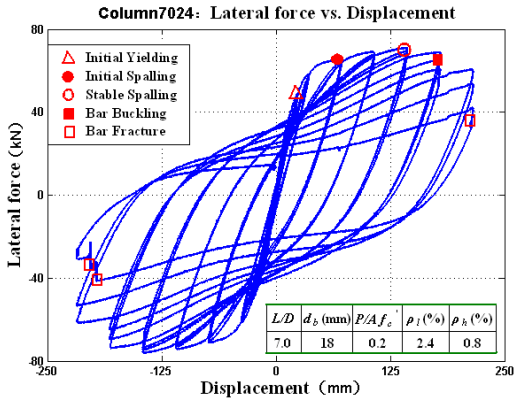


Fig.6 Force-Displacement Hysteretic Response of Bridge Columns

Conclusions

Based on the basic requirements in reference [2], the orthogonal test of four factors at three different levels for circular reinforced concrete bridge columns has been designed, and the influences of shear-span ratio, longitudinal reinforcement ratio and spiral reinforcement ratio, etc. on seismic performance of bridge columns have been investigated via quasi-static test. The test results show that:

(1) The progression of damage for all column specimens is similar: concrete cracks, flexual cracks expand diagonally, cover concrete spalling, spiral reinforcements yield and longitudinal reinforcements buckling, core concrete crushing and longitudinal reinforcements fracture. For columns with longitudinal reinforcements in small diameter, the failure of column is controlled by fracture of longitudinal steels, and which can bear fewer times of cyclic loading when longitudinal steels fracture. For columns with longitudinal reinforcements in big diameter, the failure of column is controlled by fracture of longitudinal steels and crushing of core concrete.

(2) The relevant requirements for ductility design of columns in reference [2] are reasonable, the displacement ductility of columns are greatly influenced by axial-load ratio and longitudinal reinforcement ratio, and becomes smaller with increase of axial-load ratio and longitudinal reinforcement ratio.

(3) Accumulative dissipated energy to ultimate displacement state of columns is mainly related to longitudinal reinforcement ratio.

(4) Displacement ductility and energy dissipated capability of Column7024 is relatively bad because of uncompacted vibration of concrete in bottom area of column, this leads to the adverse influence on ductility capability of columns.

References:

[1] L.C.Fan. Several Problems in Seismic Design of Bridges in Modernized Cities [J]. *Journal of Tongji University,* 1997, 25(2):147-153.

[2] National Standard of the People's Republic of China. JTG/T B02-01-2008. Guidelines for Seismic Design of Highway Bridges. Beijing. *China Communications Press*.2008.

[3] G.Q.Li. Experimental Study and Numerical Analysis on Seismic Performance of Reinforced Concrete Bridge Columns [D]. *Chongqing Jiaotong University*, Chongqing,2010.

[4] R.Park (1989). Evaluation of Ductilitry of Structures and Structural Assemblages from Laboratory Testing[J].*Bulletin of the New Zealand National Society for Earthquake Engineering*. Vol.22.No3.September, pp.155-166.

doi:10.4028/www.scientific.net/AMM.50-51.554

Fault statistical classification and Fault mode analysis of steam-induced vibration of steam turbine

Chengbing He[1,a], Yujiong Gu[1,b]

[1]Key Lab of Condition Monitoring and Control for Power Plant Equipment Ministry of Education (North China Electric Power University), Beijing, P, R. China

[a]hcbyy@126.com, [b]guyujiong@263.net

Key Words: steam-induced vibration; Failure Mode and Effect Analysis; steam turbine;fault mode

Abstract. Steam-induced vibration fault seriously affects the reliable operation of supercritical steam turbine. A lot of actual cases of steam-induced vibration of steam turbine were collected in this work. Based on the cases, steam-induced vibration fault was statistical and classified by fault severity, fault reason, occurrence load and happened time. The results show that steam-induced vibration usually occurs after turbine runs one year later and turbine is with a high load. Oil temperature change is the main running parameters which affect steam-induced vibration. Distribute steam mechanism, especially improper opening and order of adjustment valve, is the most important factor causing steam-induced vibration fault. Meanwhile, based on the FMEA method, fault mode of steam-induced vibration was analyzed in detail.

1 Introduction

Steam-induced vibration is a kind of self-excited vibration which is caused by inner flow-induced force of steam turbine. With steam parameters of steam turbine increased, the sensitivity that tangential steam forces act on the gap between rotor and stator and seal structure is improved. The problem will become more serious with the flow-induced force increased. It was proved by operation experience that steam-induced vibration occurs not only on subcritical steam turbines but also supercritical steam turbines. And the possibility occurred on supercritical turbines is much larger than that of subcritical turbines. Steam-induced vibration has become to one of the most important issues affecting the operation reliability of steam turbines[1-4]. We collected many actual cases of steam-induced vibration of steam turbine in this work. After studying on the cases, steam-induced vibration fault was statistical and classified from four aspects including fault severity, fault reason, occurrence load and happened time. And some valuable conclusions were achieved.

Fault statistical classification of steam-induced vibration

A large number of fault cases of steam-induced vibration were statistical and classified from four aspects: fault severity, fault reason, occurrence load and happened time.

2.1 Statistical classification according to fault severity

According to fault severity, steam-induced vibration can be divided into four categories: vibration near to alarm value, vibration up to about 80um, unit stop and check, forced shutdown. Statistical results were shown in Fig.1. It illustrated that the cases of vibration near to alarm value which were the main fault result were about 43%. Once steam-induced vibration was excited, there will be a sudden vibration on turbine rotor which led to vibration amplitude increase rapidly.

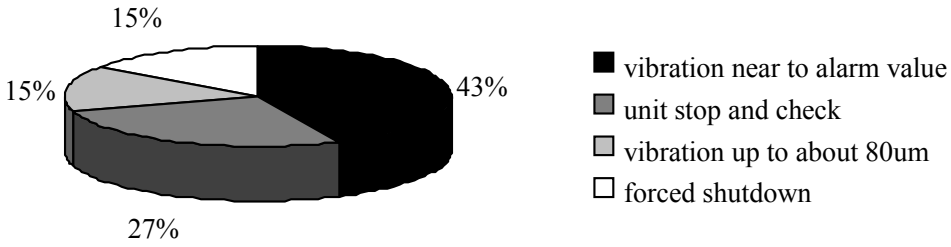


Fig.1. statistical results according to fault severity

2.2 Statistical classification according to fault reason

The fault reasons of steam-induced vibration included four aspects: steam distribution mechanism, uneven radial clearence, poor stability of high-pressure cylinder bearing, operating parameters changed. The statistical results were illustrated in Fig.2. It was obvious that the cases caused by steam distribution mechanism which was the main factor were 42%.

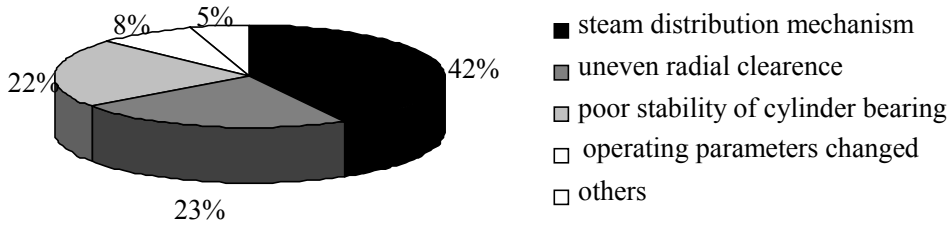


Fig.2. statistical results according to fault reason

Every one of reason above mentioned included several aspects. The analysis was as follows.

2.2.1 Steam distribution mechanism

Unsuitable operation modes usually consisted of the following: improper opening and order of adjustment valve, single-side jam of high pressure control valve, misalignment connected between main steam pipes and cylinder. The statistical results were illustrated in Fig.3. It could be found that improper opening and order of adjustment valve was the main factor causing steam-induced vibration.

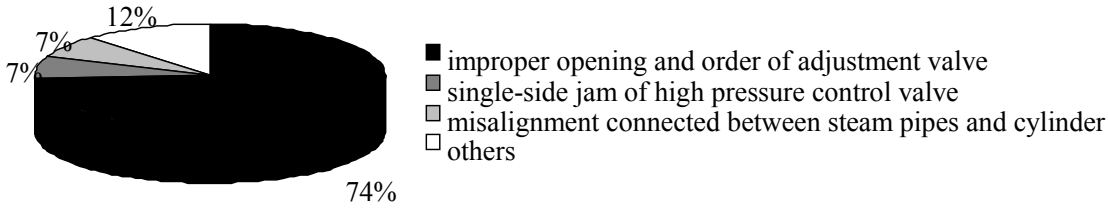


Fig.3. statistical results according to steam distribution mechanism

2.2.2 Uneven radial clearence

The reasons of uneven radial clearence contained several aspects: poor concentricity of the rotor and cylinder, rotor misalignment, improper labyrinth clearance. The statistical results were showed in Fig.4. Obviously, poor concentricity of the rotor and cylinder which amount was about 60% total number was the main factor leading to uneven radial clearence.

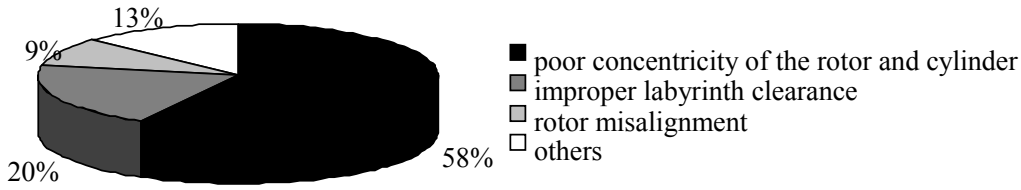


Fig.4. statistical results according to uneven radial clearence

2.2.3 Poor stability of high-pressure cylinder bearing

The reasons of poor stability of high-pressure cylinder bearing involved following aspects: bearing elevation changed, bearing headspace large, improper bearing type. The statistical results were illustrated in Fig.5. It could be found that the scale of bearing elevation changed was about 60%. And bearing elevation changed was the most important reason for poor stability of high-pressure cylinder bearing.

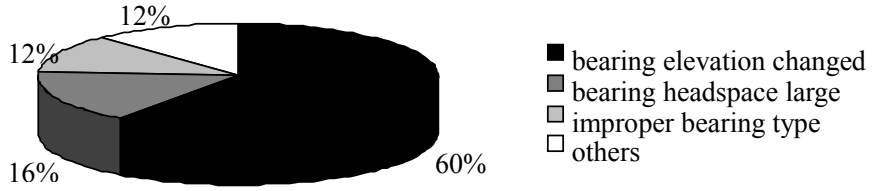


Fig.5. statistical results according to poor stability of high-pressure cylinder bearing

2.2.4 Operating parameters changed

Operating parameters changed mainly included oil temperature, main steam parameters, seal parameters and shaft imbalance. Fig.6 showed the statistical results which illustrated that oil temperature was the main factor leading to steam-induced vibration.

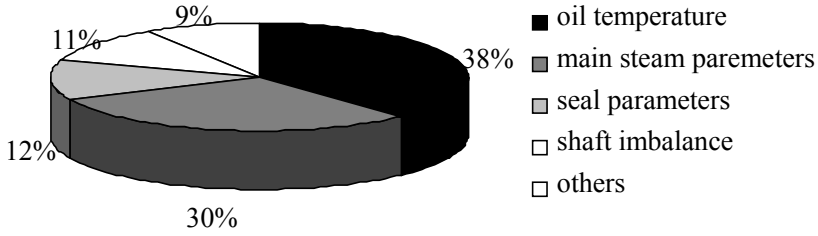


Fig.6. statistical results according to operating parameters changed

2.3 Statistical classification according to load

The statistical results according to load were illustrated in Fig.7. It was showed that most of steam-induced vibration occurred when the load was high. And low frequency vibration had a good reproducibility with load. With high load, steam flow and exciting force would increase. That was easier to bring about steam-induced vibration. In order to prevent or suppress steam-induced vibration, most of the units ran with limited load. That caused lower economy for power plant.

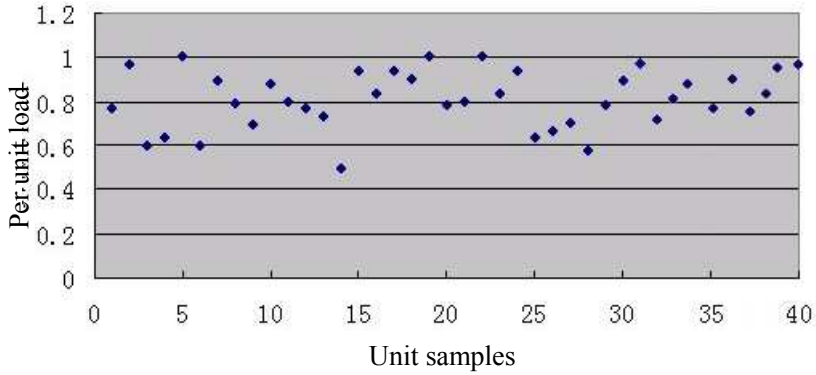


Fig.7. statistical results according to load

2.4 Statistical classification according to fault happened time

Happened time of steam-induced vibration can be divided into three categories: (1)new machine startup or early stage;(2)running one year later;(3)startup after overhaul.

The statistical results were illustrated in Fig.8. It could be found that steam-induced vibration occurred after one year operation usually. The reason was that shaft stability would become weak after the unit ran for some time, which led to thermal deformation, basement looseness and bearing fatigue damage. The cases caused by new machine and startup after overhaul accounted for nearly

half of the total. The reason was that some elements including misalignment between cylinder and rotor and uneven clerence caused by poor quality of installation and maintenance brought about steam-induced vibration. So the new units and the units after overhaul would pay more attention to monitor and prevent the occurrence of steam-induced vibration.

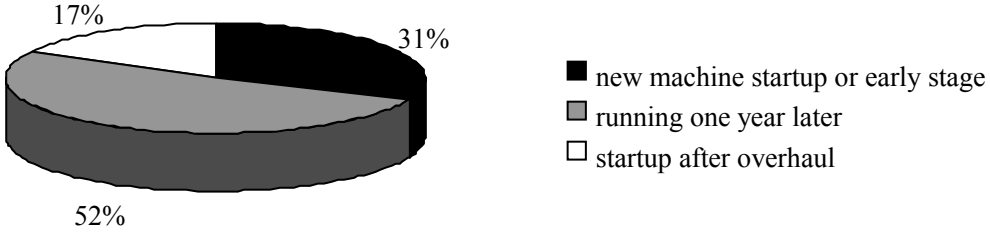


Fig.8. statistical results according to happened time

To sum up the above arguments, steam-induced vibration usually occurs after turbine runs one year later and turbine is with a high load. Oil temperature change is the main running parameters which affect steam-induced vibration. Distribute steam mechanism, especially improper opening and order of adjustment valve, is the most important factor causing steam-induced vibration fault.

3 Fault mode analysis of steam-induced vibration based on FMEA method

Failure Mode and Effect Analysis(FMEA) was an analysis method faced to system specific physical unit. Failure mode analysis used the bottom-up method. Based on the basic failure mode of system unit, potential fault reasons would be found out. Analyzing the effect by fault reasons, some actions were took to reduce and prevent the impact of fault on the system[5]. By using FMEA, we will at least get the fault symptoms, fault consequence and effect, fault reason, preventive actions and other information of steam-induced vibration.

Fault mode analysis of steam-induced vibration was shown in Tab.1.

Tab.1 Fault mode analysis of steam-induced vibration

<table>
<tr><td colspan="3">fault mode</td><td colspan="9">steam-induced vibration</td></tr>
<tr><td colspan="3">occurred position</td><td colspan="9">rotor-support system of high-pressure cylinder</td></tr>
<tr><td rowspan="4">fault symptoms</td><td rowspan="3">monitoring parameters symptoms</td><td>name</td><td>x、y bearing vibrations</td><td>bush vibration</td><td>frequency</td><td>load</td><td>main steam pressure</td><td>opening and order of adjustment valve</td><td>gap voltage</td><td>jacking oil pressure</td></tr>
<tr><td>value</td><td>amplitude sudden increased</td><td>amplitude sudden increased</td><td rowspan="2">$f \approx 0.5 f_P$ or $f \approx \frac{n_k}{60}$</td><td rowspan="2">happened with high load, strong representation</td><td rowspan="2">significant relationship between pressure and vibration</td><td rowspan="2">amplitude sudden increase with some opening</td><td>floating</td><td>fluctuating</td></tr>
<tr><td>threshold</td><td>125μm alarm</td><td>60μm alarm</td><td>floating value>20%</td><td>fluctuation value>0.2MPa</td></tr>
<tr><td colspan="2">visual signs of physical phenomena</td><td colspan="8">violent vibration; a humming noise in high-pressure cylinder bearing</td></tr>
<tr><td rowspan="4">fault consequence and effect</td><td colspan="2">current level</td><td colspan="8">rotor-support system of high-pressure cylinder</td></tr>
<tr><td colspan="2">effect described</td><td colspan="8">shaft stability decreased; induced rotor instability; oil pressure fluctuation; bush or shaft journal fatigue failure</td></tr>
<tr><td colspan="2">up level</td><td colspan="8">shafts of steam turbine</td></tr>
<tr><td colspan="2">effect described</td><td colspan="8">shafts or unit damage; rub-impact between rotor and cylinder; rotor bending or breaking; load output limited</td></tr>
<tr><td rowspan="2">fault reasons</td><td colspan="2">current level</td><td colspan="3">bearing headspace large</td><td colspan="2">improper bearing type</td><td colspan="2">improper oil temperature</td><td>bearing elevation changed</td></tr>
<tr><td colspan="2">judgment rules</td><td colspan="3">headspace＞ 1.5‰shaft journal</td><td colspan="2">stability of barrel bearing is poor than elliptical</td><td colspan="2">output oil temperature of oil cooler＜35℃</td><td>deviation from the specified value＞</td></tr>
</table>

			bearing		0.2mm
	up level	improper steam distribution mechanism	uneven radial clearence	improper steam distribution of gland	main steam pressure fluctuation
	judgment rules	check with steam distribution process	stop machine, maintenance record	stop machine, maintenance record	>0.1MPa
fault reson search		open cylinder and check			
preventive actions		1 strictly control bearing neck up-rising inclination, bush pre-tighten force and bearing load when steam turbine installation. 2 carefully adjust safety clearance of gland over shroud, diaphragm gland shaft external gland; prevent excessive clearance. 3 adjust bearing elevation of cold state should seriously considered the change character of bearing elevation when unit has different load. operation aspects： 1 potimization opening and order of adjustment valve. 2 prevention a large number of steam leakage, which caused bearing lost stability. 3 adjustment oil temperature.			

4 Conclusions

We collected many actual cases of steam-induced vibration of steam turbine in this work. The cases divided into four categories: fault severity, fault reason, occurrence load and happened time were statistical and classified. The results show that steam-induced vibration usually occurs after turbine runs one year later and turbine is with a high load. Oil temperature change is the main running parameters which affect steam-induced vibration. Distribute steam mechanism, especially improper opening and order of adjustment valve, is the most important factor causing steam-induced vibration fault.

Acknowledgment

The authors are grateful for the financial support from National Science Foundation of China (Grant No. 50905058), program for new century excellent talents in university(NCET-08-0769) and the Fundamental Research Funds for the Central Universities(Grant No. 10MG14).

References

[1] A.Muszynska. Whirl and whip-rotor/bearing stability problems[J]. *Journal of Sound and Vibration*,1986,110:3-7

[2] Alford J S. Protecting turbomachinery from self-excited whirl[J]. *ASME Journal of Engineering for Power*,1965(5):333～344

[3] Luo Mingwen, Ding Xuejun, Yang Yanlei. The Status Quo and Prospect Concerning a Study on Steam Flow Excitations in Large-sized Steam Turbines[J]. *Journal of Engineering for Thermal Energy and Power*, 2006,21(6):551-555

[4] Heung Soo Kim, Maenghyo Cho, Seung Jin Song. Stability analysis of a turbine rotor system with alford force[J]. *Journal of Sound and Vibration*,2003,260:167-182.

[5] Gu Yujiong. Theoretics and technology of state maintenance of power plant equipment[M].Bei Jing: *Chinese Electric Power Press*,2009

doi:10.4028/www.scientific.net/AMM.50-51.559

The control of stochastic resonance by harmonic signal

Meili Zhang, Min Lin

College of Metrology Technology and Engineering,China Jiliang University,Hangzhou, China

E-mail: linm@cjlu.edu.cn

Keywords: stochastic resonance; harmonic signal;control;

Abstract. The symmetry in a double-well potential of bistable systems subject to the action of a harmonic signal is broken. By variation of the amplitude of a harmonic signal, the asymmetry of bistable systems can be effectively controlled, the transitions probability between the two potential wells is increased and has influence on the effect of stochastic resonance in bistable systems. Numerical simulations show that adjusting the amplitude of a harmonic signal can produce stochastic resonance or make stochastic resonance more intense in bistable systems. It is a new method for stochastic resonance control.

Introduction

Stochastic resonance (SR) is a cooperative effect that is observed in nonlinear systems driven by a weak signal and noise. SR describes a nonlinear phenomenon that a weak input signal can be amplified and optimized by the assistance of the noise in nonlinear systems. The phenomenon happens when time scale of the weak signal and noise-induced probability transitions between the potential wells achieves stochastic synchronization. Since SR was proposed by Benzi et al in 1981[1], a large amount of numerical and experimental work has been devoted to the study of SR [2-3]. SR has been observed in a large variety of systems in physics, chemistry, engineering, biology and biomedical. In particular, SR has useful applications in engineering, signal processing, biology and biomedical science fields [4-5]. To quantitatively character SR, some measures to detect SR is introduced such as the spectral power, signal-to-noise ratio, and the residence-time distribution function. In many situations, however, the strength of the noise is not arbitrarily controllable. In order to enhance or suppress the spectral power, a control method based on a modulation of the height of the potential barrier is proposed [6]. Subsequently, a series of control methods such as harmonic mixing, adding external feedback and the control based on vibration resonance are proposed [7, 9]. Using these control methods, spectral power or signal-to- noise ratio of output signal of the bistable system can be effectively controlled.

This paper analyzes the characteristics of double-well potential of bistable system subject to the action of a harmonic signal. The effect of stochastic resonance in the bistable systems is quantitatively described by using spectral power. The influence of amplitude and frequency of a harmonic signal on SR is numerically analyzed. The stochastic resonance control method is proposed that can produce SR or enhance SR in bistable systems by adjusting the amplitude of a harmonic signal.

The Mode of Stochastic Resonance by Two Periodic Signals

The motion of a Brownian particle in a bistable potential in the presence of noise and two different period signals is given by

$$\dot{x} = -V'(x) + A\cos(\omega t) + \Gamma(t). \tag{1}$$

Where $V(x)$ is an asymmetry double-well potential,

$$V(x) = -\frac{a}{2}x^2 + \frac{b}{4}x^4 - xB\cos(\Omega t). \tag{2}$$

Where a,b is potential parameters of a nonlinear system, $A\cos(\omega t)$ is a basic frequency signal with an amplitude A and a frequency ω, $B\cos(\Omega t)$ is a harmonic signal with an amplitude B and a frequency $\Omega = n\omega$, n is integer with $n > 1$. $\Gamma(t)$ is a zero-mean, Gaussian white noise with autocorrelation function

$$<\Gamma(t)\Gamma(t')>= 2D\delta(t-t'). \tag{3}$$

Where D is the noise strength.

In the absence of period signal and noise, the double-well potential is symmetric. The potential minima are located at $x = \pm\sqrt{a/b}$ and are separated by a potential barrier with the height given by $\Delta V = a^2/4b$. The barrier top is located at $x = 0$.

In the absence of a harmonic signal, Eq.1 forms the archetypical model for SR. When the noise strength is small, particles stochastic fluctuate in the single potential well. Via reducing the height of the potential barrier, more particles in the single potential well can overcome the potential height, hop between wells and lead to exhibit SR. When the noise strength is large, particles disorderly switch between the two potential wells. Via increasing the height of the potential barrier, disordered switching can be transferred into ordered switching.

In the presence of a harmonic signal, the symmetry of the bistable system is broken and the double-well potential is tilted asymmetrically up and down, periodically raising and lowering the potential barrier, as displayed in Fig. 1. If particles stochastic fluctuate in the left potential well, the harmonic signal $B\cos(\Omega t) > 0$ pushes particles to the right potential well; if particles stochastic fluctuate in the right potential well, the harmonic signal $B\cos(\Omega t) < 0$ pushes particles to the left potential well. This periodically locks the potential back and forth, so particles can overcome the height of the potential barrier and hop between wells. Indeed, by adding a harmonic signal, the height of the potential barrier dynamically changes, therefore, the transitions of particles between potential wells can be controlled effectively and the effect of SR becomes controllable.

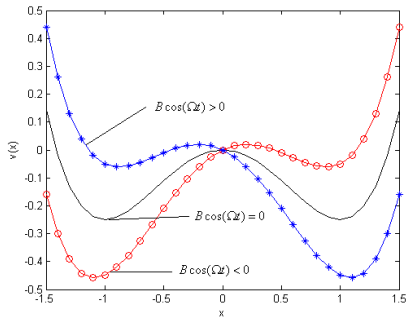


Figure 1. The bistable system potential by a harmonic signal

Results based on Numerical Simulation and Discussions

Stochastic Resonance by Harmonic Signal

Sets potential parameters $a = 1.0$, $b = 1.0$, the amplitude and frequency of basic frequency signal is respectively $A = 0.2$, $\omega = 0.01\pi$, the noise strength $D = 1.0$ and the sampling frequency is selected to be $f_s = 100Hz$. When the amplitude of harmonic signal is $B = 0$, the particles can not overcome the height of the potential barrier, only fluctuates in the single potential well. There is no SR effect in the bistable system. The spectral power of output at the signal frequency ω is $P(\omega) = 0.003434$, as shown in Fig. 2. When the amplitude and frequency of a harmonic signal is $B = 0.26$, $\Omega = 0.04\pi$ respectively, due to the action of the harmonic signal, double-well potential dynamically changes, particles hop between wells, SR in the bistable system happens. The spectral power of output at the signal frequency ω is $P(\omega) = 0.3751 > 0.003434$, as shown in Fig. 3.

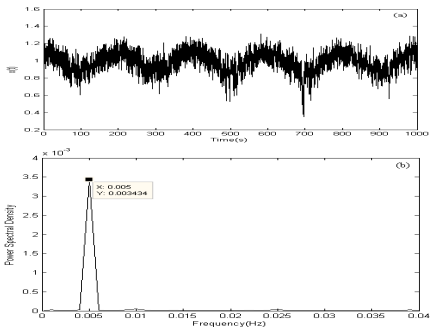


Figure 2. Time-domain waveform and power spectrum driven by single signal (a) Time-domain waveform (b) power spectrum of $x(t)$.

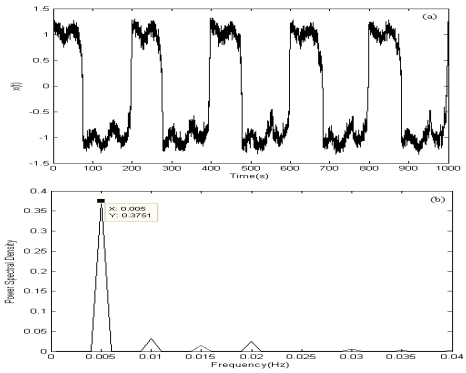


Figure 3. Time-domain waveform and power spectrum driven by double signals (a) Time-domain waveform (b) power spectrum of $x(t)$

The potential parameters $a=1.0$, $b=1.0$ are fixed, sets $A=0.25$, $\omega=0.01\pi$, $D=4.8$ respectively. When the amplitude of a harmonic signal is $B=0$, SR in the bistable system happens. The spectral power of output at the signal frequency ω is $P(\omega)=0.275$, as shown in Fig. 4. When the amplitude and frequency of a harmonic signal is $B=0.15$, $\Omega=0.06\pi$ respectively, the effect of stochastic resonance in bistable system is significantly reinforced, the spectral power of output at the signal frequency ω is $P(\omega)=0.3859>0.275$, as shown in Fig. 5.

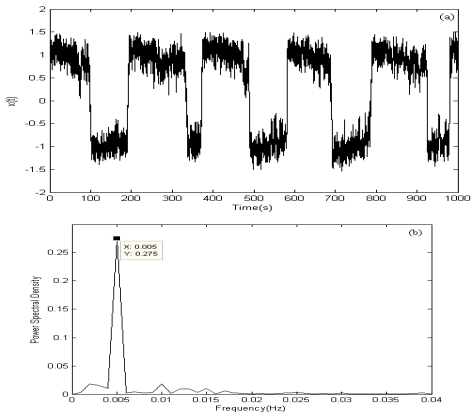


Figure 4. Time-domain waveform and power spectrum driven by single signal (a) Time-domain waveform (b) power spectrum of $x(t)$

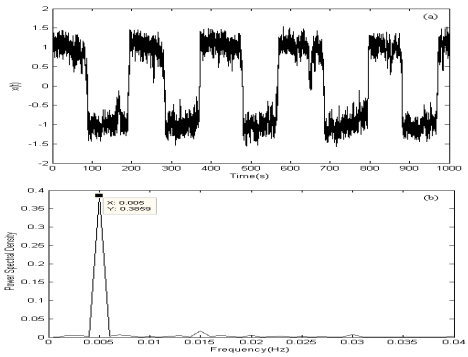


Figure 5. Time-domain waveform and power spectrum driven by double signals (a) Time-domain waveform (b) power spectrum of $x(t)$

The Influence of Parameters of Harmonica Signal on Stochastic Resonance

Set $a=1.0$, $b=1.0$, $A=0.1$, $\omega=0.01\pi$, $D=4$ respectively. From Fig. 6, when the value of Ω at $\Omega=0.04\pi$ is fixed, the spectral power of output at the signal frequency ω first increases with increasing the amplitude B of a harmonic signal, reaches a maximum, and then decreases again. The spectral power of output at the signal frequency ω is maximum at $B=0.27$. From Fig. 7, when the value of B at $B=0.27$ is fixed, the spectral power of output at the signal frequency ω changes with changing the frequency Ω of harmonic signal and has a peak at $\Omega=0.04\pi$.

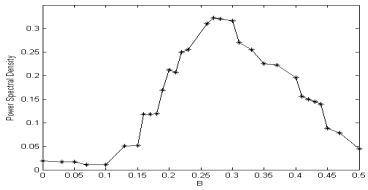


Figure 6. Power spectral density versus the amplitude B of a harmonic signal

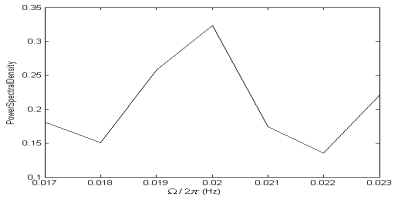


Figure 7. Power spectral density versus the frequency Ω of a harmonic signal

The potential parameters $a=1.0$, $b=1.0$ and noise strength $D=4$ are fixed. Set $A=0.25$, $\omega=0.01\pi$ respectively. From Fig. 8, when the value of Ω at $\Omega=0.02\pi$ is fixed, the spectral power of output at the signal frequency ω first increases with increasing the amplitude B of a harmonic signal, reaches a maximum, and then decreases again. The spectral power of output at the signal frequency ω is maximum at $B=0.317$. From Fig. 9, when the value of B at $B=0.317$ is fixed, the spectral power of output at the signal frequency ω changes with the change of the frequency Ω of a harmonic signal and has a peak at $\Omega=0.02\pi$.

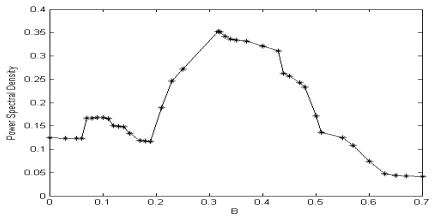


Figure 8. Power spectral density versus the amplitude B of a harmonic signal

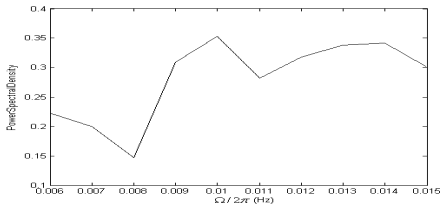


Figure 9. Power spectral density versus the frequency ω of a harmonic signal

Conclusions

The symmetry in a double-well potential subject to the action of a harmonic signal is broken. The height of the potential barrier of bistable systems can be dynamically adjusted by varying the amplitude of the harmonic signal. Numerical simulation shows that changes in the amplitude and frequency of a harmonic signal can affect the spectral power of bistable systems output at the basic frequency signal. Adjusting the amplitude of a harmonic signal can produce stochastic resonance or make stochastic resonance more intense in bistable systems so that SR in the bistable systems can be effectively controlled. A low frequency signal with small amplitude can produce SR using additional harmonic signal to control SR, and it widens the application range of SR, especially detecting weak signals in strong noise. A novel approach of detecting weak signals in strong noise using additional harmonic signal to control SR is presented.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Grant No. 10972207) and Natural Science Foundation of Zhejiang Province, China (Grant No. Y7080111) and Innovation Demonstration Base of Graduate Education of Zhejiang Province, China(YJ2008023). The authors awfully thanks for anonymous reviewers who made constructive comments.

References

[1] R. Benzi, A. Stutera, A. Vulpiani. The Mechanism of Stochastic Resonance, Phys A , vol.14(1981), p. L453.

[2] L. Gammaitoni, P. Hanggi, P. Jung, Fabio Marchesoni. Stochastic Resonance, Reviews of Modern Physics,vol. 70(1998), p. 223 - 287.

[3] X.L. Yang, B. Jiang, W.L. Wu, Q.Y. Tong, Adative Stochastic Resonance Technology in Detecting Weak Signal . SIGNAL PROCESSING, vol. 19(2003), p. 182-184.

[4] L. Zhang, A.G.Song. Development and Prospect of Stochastic Resonance in Signal Procssing, Acta Electronica Sinica, vol. 37(2009), p.812-817.

[5] N.Q. Hu, M. Chen, X.S. Wen. The Application of Stochastic Resonance Theory for Early Detecting Rub-Impact Fault of Rotor System, Chinese Journal of Mechanical Engineering, vol. 37(2001), p. 88-91.

[6] L. Gammaitonil, M. Löcher, A. Bulsara, P. Hänggi, J. Neff, K. Wiesenfeld et al. Controlling Stochastic Resonance, Phys. Rev. Lett, vol. 82(1999), p. 4574-4577.

[7] M. Lin, Y.M. Huang, L.M. Fang, The Feedback Control of Stochastic Resonance in Bistable System, Acta Physica Sinica, vol. 57(2008), p. 2041-2047.

[8] G. Schmid, P. Hanggi, Controlling Nonlinear Stochastic Resonance by Harmonic Mixing, Physica A, vol.351(2005), p. 95–105.

[9] M. Lin, Y.M. Huang, Stochastic Resonance Control Based on Vibration Resonance, Acta Physica Sinica, vol. 56(2007), p. 6173-6177.

doi:10.4028/www.scientific.net/AMM.50-51.564

Image Interpolation Algorithm Based on Edge Features

Yunfeng Yang[1,2,a] Xiaoguang Wei[1,b] Zhixun Su[2,c]

[1]School of Mathematica Science and Technology, Northeast Petroleum University, Daqing, 163318, Peoples Republic of China

[2]School of Mathematical Sciences, Dalian University of Technology, Dalian, 116024, Peoples Republic of China

[a]yfyang1080@gmail.com, [b]weixiaoguang1988@163.com, [c]zxsu@hotmail.com

Keywords: Image Interpolation, Edge Detection, B-spline, Bilinear Interpolation

Abstract. Image interpolation is used widely in the computer vision. Holding edge information is main problem in the image interpolation. By using bilinear and bicubic B-spline interpolation methods, a novel image interpolation approach was proposed in this paper. Firstly, inverse distance weighted average method was used to reduce image's noise. Secondly, edge detection operator was used to extract image's edges information. It can help us to select different interpolation methods in the image interpolation process. Finally, we selected bilinear interpolation approach at non-edge regions, and bicubic B-spline interpolation method was used near edges regions. Further more, control vertexes were computed from pixels with calculation formula which has been simplified in the B-spline interpolation process. Experiments showed the interpolated image by the proposed method had good vision results for it could hold image's edge information effectively.

Introduction

Image interpolation has been widely used in computer vision, such as still images zooming and signal resampling. The aim of Image interpolation is to predict unknown pixel through known image information. In the early years, nearest neighbor and bilinear interpolation methods were used in the processing of the image interpolation for their calculation is simple. But interpolation results can not reach people's requires for blocky artifacts. In the last few years, many image interpolation algorithms based on edges information were proposed because that edges are the most important manifest features of an image. So the treatment of image edges becomes the important goal in the process of an image interpolation[1]. Spline functions has good smoothness, and they have been widely used in data interpolation and fitting[2], [3]. They also were used in images interpolation because they can be used to solve blocky artifacts near image edges effectively. Yang[4] proposed an image interpolation algorithm based on Coons Surface. Durand[5] proposed a method based on B-spline function. An image's edges reflect it's high-frequency information, So there are also many interpolation methods based on image transform fields. Liu[6] proposed images interpolation algorithms based on wavelet transform respectively. Luo[7] proposed a new algorithm for image interpolation based on subdivision scheme for curve and surface design. B-Spline function is the important tool in an image interpolation for that their calculations are simple. But we should solve the blurry edges effectively in the process of images interpolation which based on B-spline function.

Presented in this paper is a novel images interpolation approach which is based on image's edges. In the image interpolation process, bicubic B-spline function, bilinear function and an image's edge detection were used.

Image Interpolation

Image Denoising. An image should be denoised before the interpolation for the reason that noise may be existed in the initial image. And noise can impact the interpolation results. Furthermore, the image edge detection operator will be impacted by the noise. The inverse distance weighted average method was used to reduce image's noise. It's a simple and effective denoising method. The inverse distance weighted average template is as follows.

$$\frac{1}{12}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 2 \\ 1 & 2 & 1 \end{pmatrix} \tag{1}$$

Edge Detection. In order to solve interpolation problems near image's edge pixels, we extract the image's edge with Roberts operator at first. Because Roberts operator has traits of higher precision and distinct image edges in one order image edge extraction operators. And Roberts operator is the simplest edge extraction method.

$$E(i,j) = | f(i,j) - f(i+1,j+1) | + | f(i+1,j) - f(i,j+1) | \tag{2}$$

Where f denotes an image, and E denotes it's edges, (i,j) expresses the position of image's pixel.

Through Extracting image's edges, we can get image's detail information. And it can help us to select different interpolation methods in the image interpolation process.

Bilinear Interpolation. After edge detection, we select different interpolation method based on the image's edges. Bilinear interpolation method is simple because of it's less computation, and it adapts to the interpolation of smooth regions in an image. So we select bilinear interpolation method at smooth regions of an image.

$$I(x,y) = x(v_{10} - v_{00}) + y(v_{01} - v_{00}) + xy(v_{11} + v_{00} - v_{10} - v_{01}) + v_{00} \tag{3}$$

where I denotes interpolated pixel, (x,y) denotes interpolated position, v_{00}, v_{01}, v_{10} and v_{11} which showed in Fig. 1 are four vertexes of an interpolation unit.

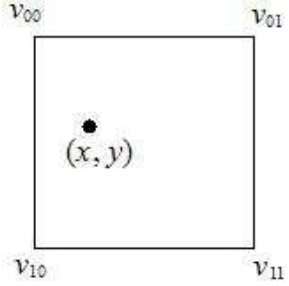


Fig. 1

Bicubic B-spline Interpolation The interpolation at edges of an image is important. B-spline is an effect data interpolation and fitting tool. We use bicubic B-spline Interpolation method near an image's edges. But if we choose pixel as B-spline function's control vertex directly, interpolated image's edges would become blurred. So we should calculate real control vertexes. Control vertexes computation formula is as follows[3].

$$b_i = \sqrt{3}\{P_i + \sum_{j=1}^{4}(P_{i+j} + P_{i-j})\lambda^j\} \tag{4}$$

where $\lambda = \sqrt{3} - 2 \approx -0.268$, P denotes pixel and b denotes control vertex.

But the work of calculations is complex when we use Eq. 4 to calculate control vertex. So we use simplified Eq. 5 in the proposed algorithm.

$$P_i^{'} = \alpha P_i \tag{5}$$

where $\alpha > 0$ is the weight coefficient that used to adjust pixel, and $P_i^{'}$ is the adjusted control vertex.

Results

In the image interpolation process, the mean value was calculated of pixels in an interpolation unit. And if a pixel in the unit was greater than the mean, we selected $\alpha = 1.05$ in Eq. 5. And Fig. 2 shows results of image interpolation, where (*a*) is the initial image, (*b*) is the edge image, (*c*) is the interpolated image with the proposed method, (*d*) is the interpolated image with B-spline function, and (*e*) is the interpolated image with bilinear function. Looks from interpolated results, image interpolated by the proposed method is better than results with bilinear and bicubic B-spline functions. There are blocky artifacts near edges in image interpolated by bilinear function. And image interpolated by bicubic B-spline function is too blur. The proposed method can solve those phenomena in interpolated images of (d) and (e) effectively.

(*a*) Initial Image

(*b*) Edge Image



(*c*) Proposed method (*d*) Bicubic B-spline method (*e*) bilinear method

Fig. 2 Image Interpolation

Conclusion

An important problem of the image interpolation is how to hold edge information. In order to solve blocky artifacts and blurring phenomena in the image interpolation, A novel method was proposed. bilinear interpolation method was selected at smooth regions, and bicubic B-spline interpolation method was used near image edges region. Experiments results show that the proposed algorithm is effect and simple in the image interpolation.

Acknowledgment

This work is supported in part by the National Natural Science Foundation of China(No. U0935004, 60873181) and the Fundamental Research Funds for the Central Universities.

References

[1] K.Jensen, D.Anastassion. Subpixel Edge Localization and the Interpolation of Still Images. IEEE Trans. on Image Processing, 1995,4,(3),p.285-295

[2] B.Su, D.Liu. Computational Geometry, Shanghai, Shanghai Scientific and Technical Publishers, 1980

[3] R.Zhou, Z.Su, Y.Xi, S.Shao. Curve and Surface in CAGD. Jilin University Press, 1993

[4] Y.Yang, Z.Su, J.Hu. A New Edge-holding Algorithm of Image Interpolation, Journal of Image and Graphics, 2005,10,(10),p.1248-1251

[5] C.X.Durand, D.Faguy. Rational zoom of bitmaps using B-spline interpolation in computerized 2-D animation, Computer Graphics Forum, 1990,9,(1),p.27-37

[6] Z.Liu, D.Liu. Reappraising about image magnification methods based on wavelet transformation. Journal of Image and Graphics, 2003,8,(4),p.403-408

[7] L. Luo, X.Yang. A Subdivision Approach to Image Interpolation. Journal of Computer-Aided Design & Computer Graphics,2006(9),p.1311-1316

doi:10.4028/www.scientific.net/AMM.50-51.568

Numerical simulation on failure mechanism of rock slope using RFPA

Nuwen Xu[1, a], Chunan Tang[1, b], Chun Sha[2, c], Rulin Zhang[3, d]

[1] Institute for Rock Instability and Seismicity Research, Dalian University of Technology, Dalian, Liaoning 116024, China

[2] Chengdu Hydroelectric Investigation and Design Institute, State Power Corporation of China, Chengdu, Sichuan 610072, China

[3] State Key Laboratory for Disaster Reduction in Civil Engineering, Tongji University, Shanghai 200092, China

[a]nwxu@yahoo.cn, [b] tang_chunan@yahoo.com, [c]shachun@chidi.com.cn, [d] zhangrulin@yahoo.cn

Key words: strength reduction method; rock slope; stability analysis; safety factor; RFPA

Abstract. This research applied a numerical code, RFPA^{2D} (Realistic Failure Process Analysis) to evaluate the stability and investigate the failure mode of the high rock slope during excavations based on Strength Reduction Method (SRM). The corresponding shapes and positions of the potential slip surfaces are rationally simulated in different stages, and the related safety coefficients are obtained, which agrees well with the allowable minimum safety factors of the slope. The numerical results show that the safety coefficient drops from 1.25 at the natural state to 1.09 after excavation, and then increases to 1.35 after slope reinforcement. Moreover, the potential slip surface of the left bank moves into deep rock mass after taking support measures, which demonstrates the reinforcement is reasonable and efficient. The study shows that cracks and faults will cause crucial influences on the slope stability, and RFPA^{2D} is a good tool to directly display the potential slip surface of the slope, which will offer valuable guidance for bolt support.

Introduction

Rock slope stability is one of the most pop and difficult problems in geotechnical engineering. How to effectively control deformation of high rock slope and ensure the safety of these engineering projects is a task that people always probe. In recent years, a variety of numerical techniques such as the methods of finite element, finite difference, infinite element, boundary element (for continuous geotechnical media), discrete element, block element, fractural element (for discontinuous geotechnical media) and their combination are widely adopted in the realm of sophisticated stability analysis for geotechnical and rock engineering. The numerical calculation methods can be divided into two groups [1]: the first is the numerical simulation combined with the optimum search method; and the second is the strength reduction method. The latter often combines with finite element programs for investigating the critical slip planes. It supposes some slip planes first, and then finds out the plane with the smallest safety parameter by stress-strain calculation results. For example, Lin et al. [1] investigated a new direct slip plane searching method in his paper, which mainly focused on the mechanism of slope deformation failure. By means of functional data analysis and the strength reduction technique, Zheng et al. [2] developed a technique for searching the critical slip surface. Yang et al. [3] put forward a modified Hoek-Brown failure criterion into the simulation and studied the effects of nonlinearity and dilation and associated or non-associated flow rules. These achievements above have great contributions to the stability research for high rock slope.

In this research, a numerical code RFPA^{2D} based on Strength Reduction Method was employed

to evaluate the stability of the left bank of Jinping first stage hydropower station. The failure modes of the high rock slope during excavations were investigated. Meanwhile, the safety factors in different stages were obtained from numerical simulation. The numerical results displayed the failure processes of the slope very well.

Engineering geological context

The Jinping first stage hydropower station is to be the largest double-curvature arch dam in the world, located more than 700 km southwest of Chengdu, Sichuan province, at the northeast edge of Xichang city. It is situated within the aslope transition zone from the Qinghai-Tibet Plateau to the Sichuan Basin. The slope of this project has lots of prominent characteristics such as large scale, higher stress levels, a strong rock unloading and deep cracks in the complex geological site. There are plenty of deep cracks and faults, such as f_{42}, f_5, f_8 and lamprophyres (X), which result in huge latent instable blocks in these areas. So the stability of huge block becomes an important problem under the condition of excavation, which directly leads this study to take an investigation on the stability of the slope. The geomorphic photograph of dam site before excavation and after reinforcement can be obtained from [4].

The principle of strength reduction FEM and corresponding program

The method of strength reduction FEM. For slopes, the safety factor F is traditionally defined as the ratio of the actual shear strength to the minimum shear strength required to prevent failure. As Duncan [5] points out, F is the factor by which the shear strength must be divided to bring the slope to the verge of failure. An obvious way of computing F with a finite element or a finite difference program is simply to reduce the shear strength until collapse occurs. The definition of safety factor based on the concept of strength reserve is that the slope will become critical equilibrium state when the shearing strength parameters c and φ will be replaced by the critical strength parameters c' and φ', and the corresponding formula is shown below:

$$\left.\begin{aligned} c' &= c / F_s \\ \tan\phi' &= \tan\phi / F_s \end{aligned}\right\} \tag{1}$$

We take the factor corresponding to the limit equilibrium state as the factor of safety of the slope theoretically. The safety coefficient of slope in the initial state can be obtained through the computing formula as following:

$$f_s = \frac{1}{1-\mu \times step} \tag{2}$$

Where μ represents the reduction coefficient, and $step$ is the current calculation step when the slope is critical equilibrium state.

The principle of RFPA code. Rock or rockmass is an obvious heterogeneity material and how to characterize the feature is of significance for the engineering. To consider the heterogeneity of rock or rockmass at the mesoscale level, in RFPA^{2D} code [6], the solid of structure is assumed to be composed of many mesoscopic elements with the same size, and the mechanical properties of these elements are assumed to conform to a given distribution law, such as uniform distribution, normal distribution or Weibull distribution. Hereby, Weibull distribution is employed to define the material properties, and it can be described as following:

$$f(u)=\frac{m}{u_0}\cdot\left(\frac{u}{u_0}\right)^{m-1}\cdot e^{-\left(\frac{u}{u_0}\right)^m} \tag{3}$$

where u is the mechanical parameter of the element, such as strength, elastic modulus, etc; the scale parameter u_0 is related to the average of the element parameters and the parameter m defines the shape of the distribution function. In RFPA^{2D} code, m is defined as the homogeneity index of the material. A larger value of m implies a more heterogeneous material and vice versa. The computationally produced heterogeneous medium is analogous to a real specimen tested in the laboratory, so in this investigation it is referred to as a numerical specimen.

Stability analysis of rock slope under different stages

Numerical models of rock slope. A typical cross-section V-V' of the slope can be got from [7]. Numerical models are set up in plane strain mode by RFPA-SRM for a heterogeneous rock slope, with 1100 m height, and 2000 m length, as shown in Fig.1. Each numerical model is discretized to 137,500 (275×500) elements. The numerical models are fixed in both horizontal and vertical directions on bottom boundary, in horizontal direction on both sides, and the upper boundary is free. A Mohr-Coulomb criterion is applied in these models and the factor of strength reduction is 0.01 for each step. For simplicity, only main joints, fractures and rockmass have been considered in this study, and the parameters of the model can be obtained from [7].

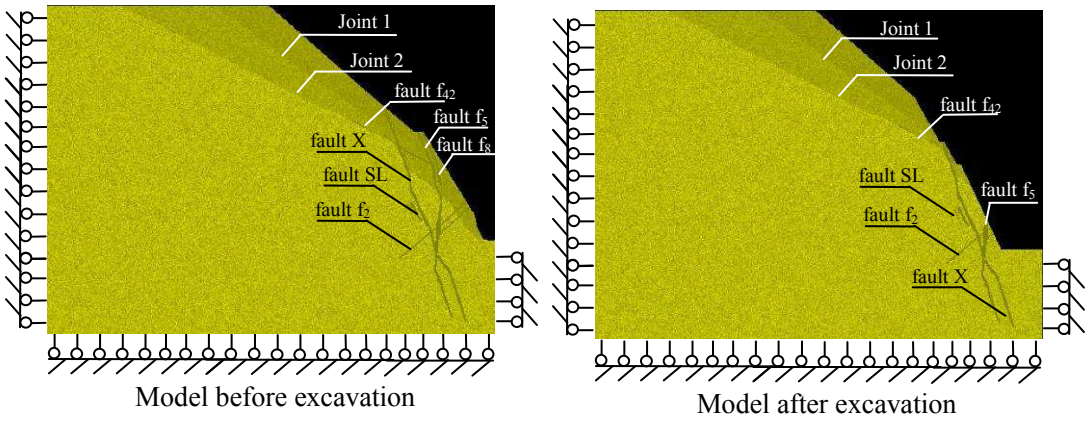


Fig. 1 FEM models geometry (The brightness of rock mass in the figures expresses the mechanical parameters such as elastic modulus and compression strength of rock mass, the brighter the color, the bigger the strength parameter).

Simulation results and analysis. Applied with RFPA-SRM, both safety coefficients and the corresponding potential slip surface during different stages have been obtained. In addition, slope instability phenomena were thus displayed directly through the failure processes of the slope.

Fig.2 shows the failure modes of the left slope in Jinping first stage hydropower station. From Fig. 2(a) it can be found that the safety factor of the natural slope is 1.25, which is calculated from formula (2). The acoustics emission figures show that many circles distribute along lamprophyre X and faults f_2, f_5, f_8 and f_{42}, in which the red circles indicate tensile fractures and the white ones compressive fractures. The potential slip surface also continually distribute along such faults and fractures, which indicates that the yielding and deformation of the slope mainly concentrated in the controlled structure direction such as faults and fractures.

Fig. 2(b) illustrates the failure model of the left bank after excavation without reinforcement. The AE figure shows that the left bank will be broken along lamprophyre X after excavation, and a

plenty of red and white circle also occur along faults f_2, f_5 and X which means compressive and tensile wreck exist in such faults and fractures. The safety factor drops to 1.09 from 1.25 in the natural slope, which indicates it will be very hazardous if there is no anchoring measure after excavation of the left bank.

Fig.2(c) shows the failure mode of the left slope after reinforcement. To compare with Fig. 2(b), it can be seen that the potential slip surface and safety coefficient and acoustic emission have significant changes. Before bolt support, the potential slip surface mainly occurs along faults f_5 and f_8, and some zones of fracture distribute along with faults f_2 and f_{42}. Then, after reinforcement from altitude 1700m to 1930m with anchorage cables, the potential slip surface of the left bank transfers to deep rockmass, and the safety factor increases to 1.35, which demonstrates the current completed support measures of the slope could improve local and overall stability of the slope effectively, and the safety factor of the slope after reinforcement has already met engineering design requirements.

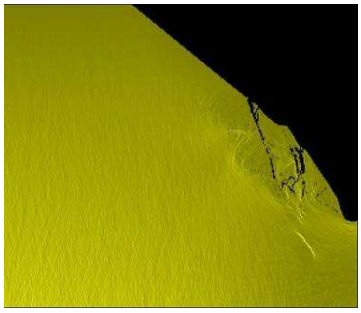
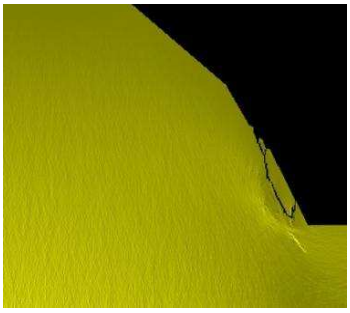
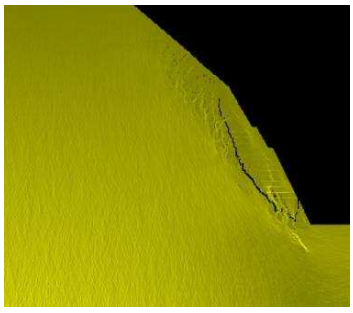
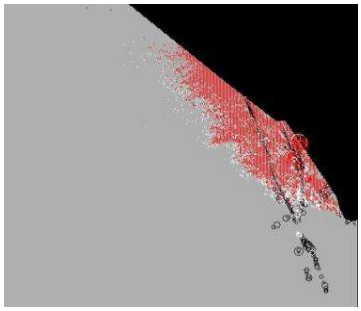
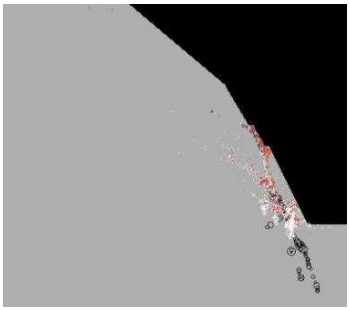
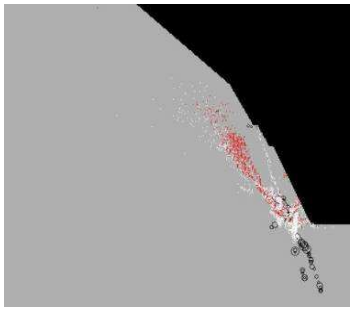
Terms	The failure mode of the natural slope	The failure mode of the slope after excavation	The failure mode of the slope after reinforcement
The maximum shear stress			
AE			
F	1.25	1.09	1.35
	(a)	(b)	(c)

Fig. 2 Simulation results of slope under various conditions (The red circles in the AE figures express tensile fractures, while the white circles express compressive fractures.)

Discussion. The failure mechanisms of slope have been demonstrated in the previous section. The results show that the safety parameter of the slope drops from 1.25 at the initial state to 1.09 after excavations, and the potential slip surface indicates that the instability of slope will be affected not only by excavations, but also faults and fractures. Namely the slope failure will occur easily when there are many faults and fractures in the deep rock mass. Safety factor increases to 1.35 after reinforcement measures through bolt anchorage, and the measure have changed the failure mode of the slope distinctly, it can be obviously seen that the potential slip surface transfers to the deep rookmass, which verifies anchor reinforcement in Jinping first project is a reasonable and reliable approach to prevent slope instability. In addition, the simulation results in this paper are agreement with other references' conclusions. It thus can be concluded that faults and fractures are one of the most significant factors on the slope stability.

Referring to *Hydraulic and Hydropower Engineering Slope Design Specifications* (DL/T5353-2006, China National Standard), the allowable minimum safety factor of the slope of Jinping I hydropower station is shown in Table 1, from which it can be seen that the safety factors obtained by numerical simulation above can satisfy the allowable minimum safety factor very well.

Table 1 The allowable minimum safety factor of the slope in Jinping I hydropower station [7]

Slope types	Slope level	Regular running	Work progress	Special condition I (consideration on rainfall)	Special condition II (consideration on earthquake)
spandrel groove high slope	1	1.3	1.15	1.20	1.10
slope at cable crane platform	1	1.25	1.15	1.15	1.05

Conclusions

(1) Incorporating with strength reduction method, RFPA^{2D} code can be adopted to estimate the stability and investigate the failure mechanism of the rock slope during different stages.
(2) The numerical results display the potential slip surfaces of the slope during different conditions. The safety factors of the slope at the natural state, as well as after excavation and reinforcement, have been rationally obtained. The safety coefficient drops from 1.25 at the natural state to 1.09 after excavation, and then it increases to 1.35 after the slope reinforcement. It is indicated that the support measures can improve local and overall stability of the slope reasonably and effectively.
(3) Cracks and fractures play an important role on the occurrence of the potential slip surfaces. It is advisable to conduct the latter site tracing monitoring and feedback analysis in connection with such potential instable blocks in the left bank.

Acknowledgements

Financial supports from the National Natural Science Foundation of China (Grant Nos. 10672028, 50820125405, and 40638040) and National Basic Research Program of China (973 Program, Grant No. 2007CB209400) are greatly appreciated.

References

[1] H.Lin, P.Cao and F.Q.Gong etc: J. *Central South Uni*.Tech Vol.16 -1(2009), p. 131−135

[2] H.Zheng, G.H.Sun and D.F.Liu: *Computers and Geotechnics* Vol. 36-1(2009), p. 1-5

[3] X.L.Yang, L.LI and J.H.Yin: Int. J. *Numerical and Analytical Methods in Geomechanics* Vol.28-2(2004), p. 181−190

[4] N.W.Xu, C.A.Tang and C.Sha etc: Chinese J. Rock. *Mech.Eng* Vol. 29-5(2010), p. 915–925(in Chinese)

[5] J.M.Duncan: *Journal of Geotech. Geoenviron Eng* (ASCE) Vol, 122-7(1996), p. 577-596

[6] C.A.Tang. Int. J. Rock Mech and Min Sci Vol. 34(1997), p. 249-261

[7] Chengdu Hydroelectric Investigation and Design Institute, State Power Corporation of China: *Special report on deformation and stability analysis of the left bank in Jinping First Stage Hydropower Station* (2008) (in Chinese)

doi:10.4028/www.scientific.net/AMM.50-51.573

Networking Register System for Web-press Machine

Li Meihong[1,a], Du Ye[1,b], Liu Jiqiang[1,c], Zheng Xin[1,d]

[1]School of computer and information technology, Beijing jiaotong university, Beijing, China

[a]mhli1@bjtu.edu.cn, [b]ydu@bjtu.edu.cn, [c]jqliu@bjtu.edu.cn, [d]jutlmh@163.com

Keywords：Color register; Mark; Image processing; Iteration

Abstract：To resolve the color-register problem on web-press machine, a kind of networking register system was presented. Firstly, a special diamond mark was designed, which is easier to be embedded. Secondly, the high-speed cameras with external trigger signal were adopted to capture mark images under the control of synchronous encoder unit, at the same time the real-time deviation of color was obtained by analyzing the mark image. Lastly, the data of color deviation were transmitted to the press system over Ethernet. Thus, a close-loop and real-time system was built. Especially, an Ethernet solution for the press machine with many towers was proposed. According to the actual customer's reports, waste of newspaper decreases greatly, in addition, print quality and print efficiency are both promoted.

Introduction

Issue speed and print quality are very important for the advertisement customer that is the key profit source for web-press newspaper. In general, color press uses four kinds of basic colors such as black, red, blue and yellow colors, and other color can be registered from the above colors. Actually web press has a high-speed rotation with a maximum of 180 thousands pieces per hour, color-register error would not only result in a huge waste of newspaper and print ink, print efficiency and quality but also decrease. In addition, this problem is more serious in the press with many towers. For these reasons, it is imperative to adopt automatic color-register system (CRS).

It is statistic that there are not color-register systems used on the most of web-press machines in China, and color register depends on a worker's manual control according to his experience. With the development of print technology, the studies on web press are more popular. Especially, a kind of auto-color register system for gravure machine in Ref. [1], Ref. [2] presents a key register technology with no marks, and cross mark is analyzed using image processing in Ref. [3] and [4].

In this paper a kind of intelligent register system for web press is presented. Firstly, a special diamond mark was designed, which is easier to be embedded in newspaper. Secondly, the high-speed cameras with external trigger signal were adopted to capture mark images under the control of synchronous encoder unit, then a kind of pre-process method is adopted for the special yellow mark that is not easy to be identified. Lastly, a set of total image-process scheme is presented to compute the color deviation. Especially, an Ethernet solution for the press machine with many towers was proposed. A set of networking register system based on the above scheme is put into practice in a newspaper office of China.

Schematic Structure

Automatic CRS is composed of supervising unit, camera unit and encoder unit. Supervising unit (Register center) is used to manage and supervising the whole register system, after the configuration, the register system would work automatically without any interferences, and real-time color deviation, mark image, rotation speed would be displayed on the touch screen. The high-speed dual-camera with external trigger signal is adopted to capture mark images on two-side newspaper under the control of synchronous encoder unit, at the same time the current color deviation is obtained by analyzing the mark images. Lastly, the data of color deviation were transmitted to the press system over Ethernet. Thus, a real-time and close-loop system was built. Encoder unit is used to obtain

rotation speed and synchronous signal, which is installed on the main axis of press machine or folder machine. The connection between CRS and press machine is shown in figure 1.

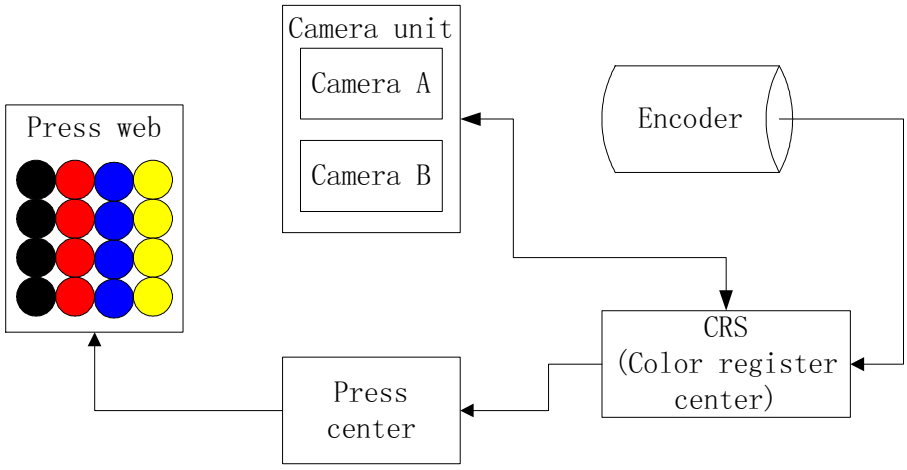


Figure 1. Connection of register system

Deviation Computation

A. *Register mark*

CRS depends on the register marks totally, thus a group of marks are designed for each color in four-color press machine. There are two black marks that act as the measure base, and others have only one. The actual register mark is shown in figure 2, the diamond with the side length between 0.3 and 1 millimeter is designed, which is easier to be embedded and identified in the newspaper. In general, the register mark is embedded on the middle top or bottom of the newspaper.

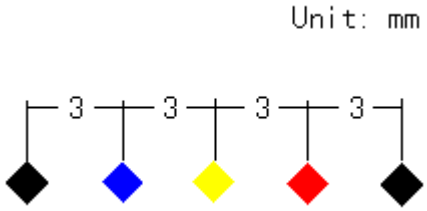
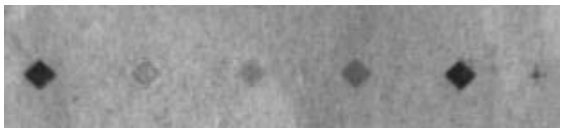


Figure 2. Register mark

B. *Image preprocess*

There may be the below problem in mark image, only the yellow mark is not very clear, which is shown in figure 3(a), it is the reason that yellow color is often printed gently or non-white paper is adopted. The following is the individual channel method for RGB-format image.

(a) Original image



(b) Blue channel reserved

(c) Green channel reserved

Figure 3. Mark image of individual channel

The original color image is shown in figure 3(a), the yellow mark under light-red background is not easy to detect. For single-CCD camera, the Bayer-format image is captured, which would be converted into RGB-format image composed by red, green and blue channel. To display the yellow mark clearly, only the blue channel is reserved shown in figure 3(b), in which the blue mark almost disappears. At the same reason, only the green channel is reserved shown in figure 3(c), in which the yellow mark would not be detected.

C. *Diamond identification*

After the above preprocess, the RGB-format image is transformed into gray-scale one, which is used to identify five pieces of marks. Finally the center of gravity for each diamond mark is to be obtained, the flowchart of mark detection is shown in figure 4, in which it refers to the phases such as image preprocess, iterative threshold computation, canny detection, contour tracing, polygon imitation and gravity computation as in Ref. [5] and [6].

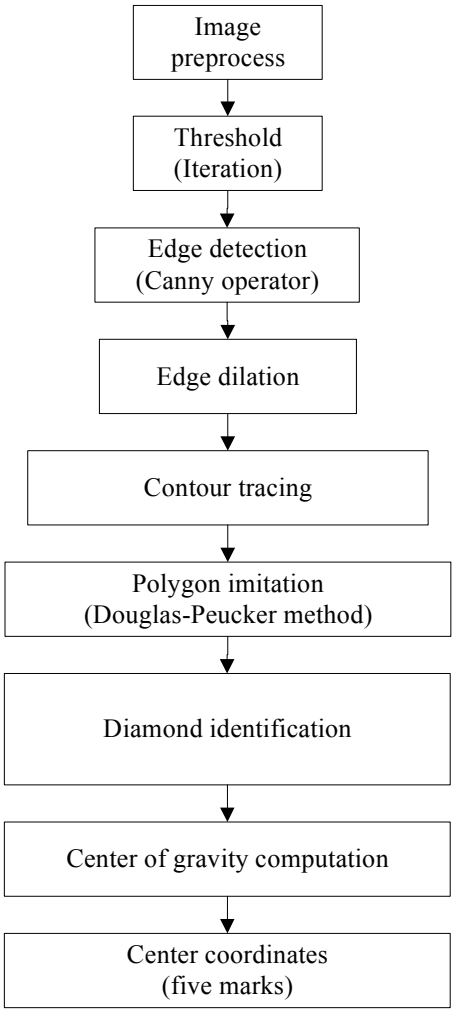


Figure 4. Process of mark identification

D. *Error calibration*

It is supposed that left-top corner of image is the coordinate center, (X_i, Y_i) denotes the coordinate of each diamond mark (i = 0~4). Here (X_0, Y_0) and (X_R, Y_R) are the coordinates of the left and right black mark respectively, it is to obtain the deviation relative to the left black mark (X_0, Y_0).

$$\begin{cases} SX_i = X_i - X_0 \\ SY_i = Y_i - Y_0 \end{cases} \tag{1}$$

The following transformation of coordinate is adopted:

$$triside = \sqrt{(SX_R - SX_0)^2 + (SY_R - SY_0)^2}$$

$$\begin{cases} \cos ine = (SX_R - SX_0)/triside \\ \sin ine = (SY_R - SY_0)/triside \end{cases} \quad (2)$$

Hence the factual deviation of each mark is as follows:

$$\begin{cases} mmx_i = SX_i * \cos ine + SY_i * \sin ine \\ mmy_i = -SX_i * \sin ine + SY_i * \cos ine \end{cases} \quad (3)$$

Here the two black marks are used as measure base, the distance between the twos equals to 12 millimeters, then the horizontal and vertical deviation of other 3 kinds of marks is obtained relative to the left black mark. The horizontal deviation means that other colors is deflected along the direction of press axis relative to the left black mark, and the vertical deviation is the one along the forward direction of newspaper. It is noted that the base refers to the one for measurement and control, the measure base must be black mark, and the control base can be variant. For example, when the red color is a current base, CRS would drive the black motors automatically if the black deviation exists.

After the error calibration is completed, the automatic correction control is very simple. The correction data is transferred to the executive unit of press machine automatically over Ethernet, hence a real-time and close-loop system is built.

Networking Structure

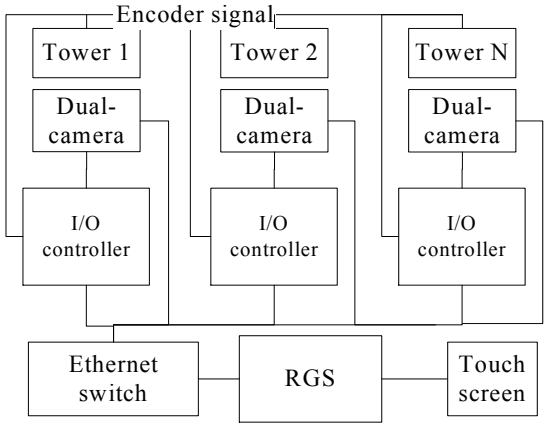


Figure 5. Networking structure for many towers

It is compulsory that CRS is used on the press machine with many towers for reason of huge waste. Here a kind of scheme on networking register is presented, which is shown in figure 5. In this figure, there are two cameras in each tower for two-side newspaper. In addition, each tower is connected with an IO controller that is used to control the capturing action of cameras, at the other hand, the controller is wired with the encoder for speed and synchronous signal. All cameras, IO controllers and supervising unit are united together over Ethernet, hence theoretically over 80 towers are supported in a IP segment if one tower uses 3 IP addresses. When CRS is put into practice, much touch screens are adopted for reliability and convenience, work of any cameras and any controllers could be observed on any touch screens.

Waste Evaluation

The above scheme-based CRS is put into use in a newspaper office of China. The following is waste statistics from the actual customer. In the newspaper office, a set of folder machine is connected the press machine with 3 towers, each tower can work independently or together, and a maximum of 48-plate newspaper would be printed in one time. With no consideration of waste of machine failure, the

waste of newspaper is 500~1000 pieces after using a set of new printing plates. If new plates are used for three times at one night, the waste would reach to 2000~2500 pieces one night, after one month the number is from 60 to 75 thousands pieces.

After CRS is adopted, it is unavoidable because CRS could work well after capturing clear images. It is statistical that one-time waste after new plate is 50~150 pieces, the number is not over 10 thousands pieces in one month, the actual waste is only 1/8~1/6 of the before. In fact, print quality and print efficiency are both promoted except for the save of papers and inks.

Conclusions

In this paper, a kind of special register mark is designed, which is easier to be embedded in newspaper, and to be identified. A kind of pre-process method is adopted for the special yellow mark that is not easy to be identified. In addition, a set of total image-process scheme is presented to obtain the color deviation, and the deviation data is transferred to the press machine automatically. Hence a real-time and close-loop system is built.

Especially, a kind of solution on Ethernet for the press machine with many towers was proposed. A set of networking register system based on the above solution is put into practice in a newspaper office of China. It is statistical that the actual waste is only 1/8~1/6 of the before. Finally, the method is suitable for the two-color and four-color web machine, and CRS would work well on the machine with normal or double roll, which supports the speed of a maximum of 200 thousands pieces per hour.

Acknowledgement

The work described in this paper has been supported in part by the National Key Basic Research Program (No.2007CB307101), China NSF(No.K09A300150), China PCSIRT(No.IRT0707), the Discipline Construction and Postgraduate

References

[1] D.C. Chen; H. Zhang; A multi-axis synchronous register control system for color rotogravure printer[C], Technology and Innovation Conference. ITIC 2006. 2006, Page(s): 1913 – 1918

[2] J. Gao; B.H. Lu; Research on an Overprinting Accuracy Analysis Algorithm without Color Marks of Multi-color Printings[C], Computational Intelligence and Security Workshops. CISW 2007. 2007, Page(s): 914 – 917

[3] Y.Z. Wang, J.H. Liu, D.S. Li, Automatic detection method for register deviation, Journal of mechanical engineering [J] Vol.45 No.12, 242-246

[4] H.X. Liu; W.J. Yang; M. Huang; B. Wu; Y.F. Xu; Detection and Control Algorithm of Multi-Color Printing Registration Based on Computer Vision[C], Image and Signal Processing. CISP '09. 2009, Page(s): 1 – 4

[5] Q. Cui. Image process and analysis [M]. Beijing: Science press, 2000

[6] Y.J. Zhang. Image segmentation[M]. Beijing: Science press, 2001.

doi:10.4028/www.scientific.net/AMM.50-51.578

A method of network forensics analysis based on frequent sequence mining

Zhong Xiu-yu

School of Computer Science, Jiaying University, Meizhou, Guangdong, China

e-mail: wch@jyu.edu.cn

Keywords: network forensics; data analysis; frequent sequence mining; Jpcap

Abstract. For the mistaken report and false alarm occurring frequently in intrusion detection system (IDS), the evidence based on forensics system of IDS is inefficient and low credibility. Frequent sequence mining based on Jpcap is proposed for network forensics analysis. After fetching and filtering network data package, the system mines data with frequent sequence according to the evidence relevance to build and update signature database of offense, and judges whether the current user's behavior is legal in the network forensics analysis stage or not. Simulation results show that the algorithm of frequent sequence mining can identify the new crime behavior and improve the credibility and efficiency of evidence in network forensics analysis.

1.Introduction

Computer forensics usually uses the intrusion detection system (IDS) to recognize illegal behavior and monitor network. In order to guarantee the confidentiality and usability of system resources, IDS monitors network or/and system activities to detect malicious activities or policy violations and produces to the management stations [1]. Because of the technology problem, the mistaken report and false alarm occur frequently in IDS, which causes the credibility of network evidence based on IDS forensics lower. Joo D and Hong T designed the anomaly intrusion detection method based on neural network [2]. M.I. Cohen designed PyFlag for network forensics to trace the network traffic generated by a system [3]. The magnanimous network data need to analyze in network forensics. For each kind of intrusion, there is often association in crime time, crime tool and crime technology. Using data mining technology, the system can discover the time relations of events and the related data of specific crime. Each kind of information is connected before the final judgment.

Many scholars have made research on the improvement and the application of frequent sequence algorithm. M.Y. Su proposed a real-time NIDS with incremental mining for fuzzy association rules [4]. J. Pei proposed mining sequential patterns by patter-growth [5]. C. Lei explored how to efficiently maintain closed sequential patterns in a dynamic sequence database environment [6]. F. Wu and S.W. Chang proposed item-transformation methods to mine frequent patterns [7]. Syed Khairuzzaman Tanbeer presented a compact pattern tree that captured database information with one scan and provided the same mining performance as the FP growth method [8]. En Tzu Wang and Guan-ling Lee proposed sanitization algorithm to modify databases for hiding sensitive patterns [9]. The frequent sequence mining algorithm applying in the network forensics analysis can improve the credibility and efficiency of evidence.

2. Development platform

The prime tasks of network forensics are fetching data and data analysis. There are many tools to fetch network data, such as BPF technology. The system is developed on the Linux+Libpcap+MySql platform. Jpcap is java classes for gathering and filtrating network data. The basic structure of Jpcap is shown as figure 1[10]. Jpcap provides java API to access the bottom layer network data on windows or UNIX system. Jpcap relies on the WinPcap or libpcap. The development environment of Jpcap must be installed if Jpcap classes are used in java.

A. The basic structure of Jpcap

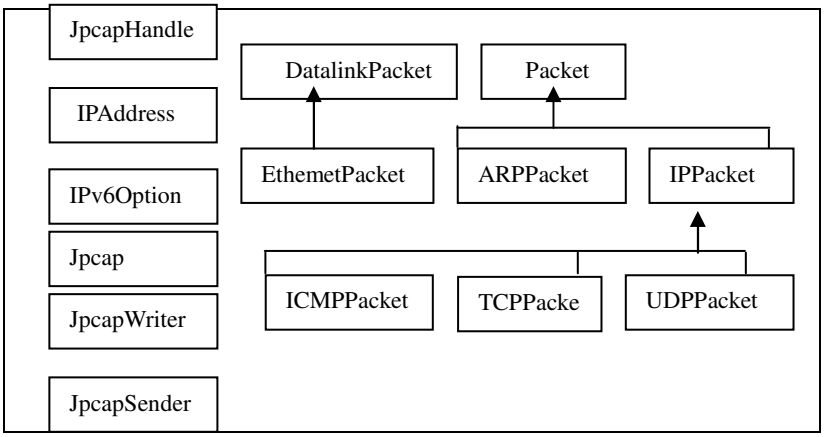


Figure 1. The basic structure of Jpcap

B. Basic class and subclass of Jpcap

Class Packet is the basic class for the captured data packages, which provides the basic information of data package, such as length, time and so on. The captured data packages are divided into two kinds, which correspond to two subclasses of Packet: ARPPacket and IPPacket. ARPPacket takes out its various segment data according to the ARP data content, IPPacket divides the data in more detail. These classes closely correlate with the data link layer, and they express correlation information with the MAC address by the class EthemetPacket. EthemetPacket is subclass of DatalinkPacket. IPPacket has three subclasses: ICMPPacket, TCPPacket and UDPPacket.

3. Fetching data and pretreatment based on Jpcap

Jpcap provides functions of fetching and processing data conveniently. With setting the monitor network card as “the promiscuous pattern”, the system receives the data package directly from data link layer, and monitors each kind of network data packet continuously, finally saves all network data stream into the MySql database in accordance with time.

A. Fetching data

Jpcap uses an event model to process the package, the concrete process is as follows:

(1) An execution association class jpcap.JpcapHandler is defined as follows:

```
public class Jpcaphandler implements JpcapHandler {
        public void handlePacket(Packet packet)
            {System.out.println(packet);}                    }
```

(2) The monitor network equipment is established by the method getDeviceList( ) that returns strings;

(3) One equipment is selected from name list of monitor devices as following: String deviceName = devices[0];

(4) The monitor is turned on by the method Jpcap.openDevice().

(5) Monitor is started by processPacket() or loopPacket().

(6) Data packages are send. Jpcap provides class JpcapSender to transmit the data package, which is used to transmit IPPacket and its subclass. Data packages can be send by the method sendPacket() after defined.

The captured data packages are analyzed by class Packet and its subclass to obtain the detailed information.

B. Pretreatment

(1) The function setFilter(java.lang.String condition, boolean optimize) of Jpcap filtrates data package, its parameter “condition” means filtration condition

(2)The useful information is fetched, for example, the MAC address is got from data package as following:

EthernetPacket ethernetPacket=(EthernetPacket)packet.datalink;

System.out.print("source MAC address" +ethernetPacket.getSourceAddress()+"destination MAC address" + ethernetPacket.getDestinationAddress() +"\n");

(3) The type of Data packages is judged

The captured data package is classified and judged whether it is TCP packet, UDP packet, ICMP packet, IP packet or the ARP packet and so on. The detail information is got directly from subclass IPPacket and ARPPacket of Packet.

The example distinction is carried on and the package type is judged by the base class Packet as following:

if (packet instance of ***Packet)

{ ***Packet p=(***Packet)packet; }

(4) According to the index deposit, the various fields of captured data packages are saved into hash-table for the next analysis step. The filed of hash-table as table 1:

TABLE I. FILED OF HASH-TABLE

pid	timestamp	length	protocol	srcMAC	srcIP	dstMAC	dstIP	srcPORT	dstPORT	state	data	descrip

C. State analysis of TCP package

After the determination of data package type, the state analysis method to each kind of data package is similar. The state analysis of TCP data package is taken as the example, the function server state (TCPPacket tcp) is called which returns state string as following:

(1) state=""; (2) if(tcp.ack=false) goto（4）; (3) state=state+'A';

(4) if (tcp.fin=false) goto (6); (5) state=state+'F'; (6) if(tcp.psh=false) goto (8);

(7) state=state+'P'; (8) if (tcp.rst=false) goto (10); (9) state=state+'R';

(10) if(tcp.syn=false) goto（12）; (11) state=state+'S'; (12)if(tcp.urg=false) goto（14）;

(13) state=state+'U'; (14) return state, the end.

To each data package or the suspicious data package, its characteristic analysis is carried on. If the data package is quite suspicious, it may be mined characteristic.

4. Mining frequent sequence

For each kind of computer crime, there is often association in crime time, crime tool and crime technology. Association rule analysis technology discovers the useful dependence or the related knowledge from the massive data. An exception model is build utilizing the association, and the model is added into the crime signature database which is renewed according to the real-time network data. The exception model is utilized in the data analysis of the crime detection to enhance the accuracy and validity of data analysis.

Using the association rule, the system mines criminality characteristics among different crime forms, and different evidence relations among identical events. The invasion characteristics is set, the corresponding event pattern is built and stored into the knowledge database. For example, ARP deceit, SYN FLOOD, ICMP invasion characteristics may be established as table 2 according to table 1.

TABLE II. PARTIAL INTRUSION CHARACTERISTIC

Intrusion type	Signature string
ARP deceit	3=arp,4 or 3=arp,5
SYN FLOOD	3,10=A,5
ICMP attack	3=icmp,2>255

At the stage of data analysis, current user behavior is judged legal or not through match results of user behavior and association rules which are stored in databases. From judging that the behavior has the crime signature or is related to some crime, crime evidence which may be crime is withdrawn. The application of mining frequent sequence in the data analysis is shown in Figure 2.

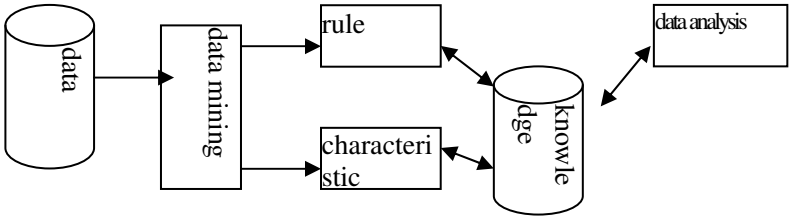


Figure 2. Application flow of frequent sequence mining

A. Apriori algorithm

The Apriori algorithm is one kind of most influential mining Boolean association rule algorithm, and the rule is expressed by frequent item collection. The association rule has two important attributes: Support level P(A∪B), namely the probability of the two items of collections A and B which simultaneously appear in business collection D. Confidence level P(B|A), namely probability that collection A appears in items of in business collection D, items of collection B also simultaneously display. The rules which simultaneously satisfy the smallest support threshold value and the smallest confidence level threshold value are called the strong rule. Giving business collection D, the association rule mining creates the rules whose support and confidence level is bigger separately than the smallest support and confidence level which the user assigns. The Apriori core algorithm has used the recursion method in order to produce all frequency collections [11].

B. Frequent sequence mining algorithm

The first step of the algorithm is to discover all frequency collection. These appearing frequency of items collections are not smaller than the minimum pre-definition support level. The second step is to generate the strong association rule by the frequency collection. The algorithm mainly contains the four functions: function creatC1() obtains the first candidate collection, function getL() obtains the frequent collection, function getC() obtains the candidate collection, function count() calculates the record number of the candidate collection. The function getL() obtains frequent collection as follows:

(1) begin (2) if (i>=list.size) then goto (8); （3）q=list.get(i+k-1);temp=q.count/ length;
(4) if (temp>=0.3) then goto (6); (5) i=i+k, goto (2); (6) if (n>=k) then goto (2);
(7) list.remove(i), goto (6); (8) end.

The algorithm final outcomes the frequent record compendium, which corresponds to an instance of class Node. The class diagram of Node is shown in Figure 3. The class attribute significance is shown as follows: (1) Count means the quantity of the record appearance. (2)Column is a Vector set, which indicates the set of dimension and stores the record field. (3)Data is a Vector set which stored corresponding value set of field. The frequent record set produces the strong association rules, which are stored in the signature database.

C. The case

In the SynFlood attack forensics, an example of frequent sequence mining application is given. Criminal sends massive SYN connection requests to the destination host in short time, but each connection is only half connection in direction. That will cause the connection request buffer of the destination host occupied completely, thus make the destination host unable to handle new connection requests, and the criminal achieves the purpose that the host will refuse to serve. This kind of situation will display in connecting records, in short period of time, some ports on a host can receive a number of “50” connecting state, which means half connection. The server is taken as the axis attribute and the destination host is taken as the quotation attribute to discover the frequent sequence service pattern of the same destination host in two seconds window time. Using frequent sequence mining method obtains patterns of implicit crimes, which indicates that the HTTP service of the host has encountered the SynFlood attack. There are not “50” patterns in normal data because half connection cannot frequently occurred in a short time.

5. Experimental result and analysis

A. Test environment

The experiments are made on a simplified network using the internal laboratory in campus net and the accommodations area network.

B. Test result

Using the Aladdin flood attack to simulate the Flood attack, the ARP deceit and the ping attack, the examination rate of system can reaches 100%, and the crime process is recorded completely. In addition, the system can also examine and record the suspicious crime. An interface of forensics data analysis is shown in Figure 4.

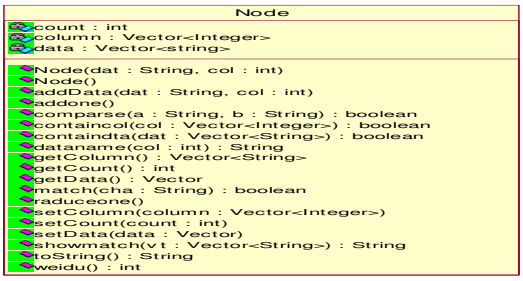


Figure 3 Class diagram of Node.

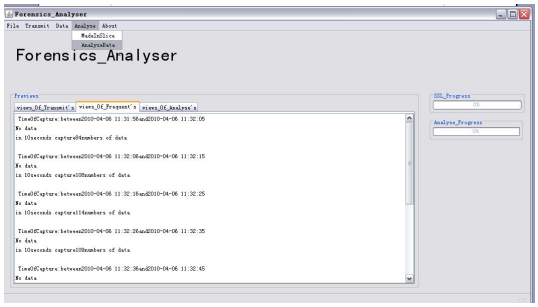


Figure 4. An interface of forensics data analysis

6. Conclusion

The application of frequent sequence mining algorithm in network forensics analysis was studied. The analogous system has realized functions: crime detection, crime characteristic mining and evidence inquiry. After fetching the network data package, the data is pretreated and mined the frequent sequence excavation to obtain the crime characteristic patterns. The patterns applying in network forensics data analysis enhance reliability of the evidence.

Acknowledgment

The authors want to thank the Natural Science Foundation of Guangdong Province and the Science and Technology Plan Project of Guangdong Province for their general support for the research (with grant NO. 9151009001000043 and 0911050400004 respectively).

References

[1] D. Li, Z.T. Li, J. Ma. Processing intrusion detection alerts in large-scale network. International Symposium on Electronic Commerce Security (ISECS'08), 2008.
[2] Joo D, Hong T, Han I. The neural network models for IDS based on the asymmetric costs of false negative errors and false positive errors.Expert Systems withApplications,2003,25(1):69-75.
[3] M.I. Cohen. PyFlag–An advanced network forensic framework. digital investigation 5 (2008) :112 – S120.
[4] M.Y. Sua, G.J. Yub, C.Y. Lin. A real-time network intrusion detection system for large-scale attacks based on an incremental mining approach, Computers & security, 28(2009) :301–309.
[5] J. Pei, J.W. Han. Mining Sequential Patterns by Patter-growth: The Prefix Span Approach. IEEE Transactions on Knowledge and Data Engineering. 2004,6(10):1-17
[6] C. Lei, T.J. Wang. Efficient algorithms for incremental maintenance of closed sequential patterns in large databases. Data & Knowledge Engineering. 2009,68(1): 68-106
[7] F. Wu, S.W. Chang. A new approach to mine frequent patterns using item-transformation methods. Information Systems. 32 (2007) :1056–1072.
[8] Syed Khairuzzaman Tanbeer, Chowdhury Farhan Ahmed. Efficient single-pass frequent pattern mining using a prefix-tree. Information Sciences 179 (2008) :559–583.
[9] En Tzu Wang, Guanling Lee. An efficient sanitization algorithm for balancing information privacy and knowledge discovery in association patterns mining. Data & Knowledge Engineering 65 (2008): 463–484.
[10] http://netresearch.ics.uci.edu/kfujii/jpcap/doc/
[11] J.W. Han, Micheline Kamber. Data Mining: Concepts and Techniques.2003:152-155

doi:10.4028/www.scientific.net/AMM.50-51.583

Application of Harmonic Balance Method to Numerical Model Unsteady Viscous Flow around Oscillating Cascade

Yongqiang Shi [1, a], Qingzhen Yang [2,b] and Xinhai Zhou [3,c]

[11] School of Engine & Energy, Northwestern Polytechnical University, Xi'an 710072, China

[a]yqshi@nwpu.edu.cn, [b]qzyang@nwpu.edu.cn, [c]zhouxh@nwpu.edu.cn

Keywords: Oscillating cascade; harmonic balance method; unsteady flow; dual-time stepping method; unsteady aerodynamic force

Abstract. A harmonic balance approach has been developed to compute nonlinear viscous unsteady flows around oscillating blades. The computed results using two orders harmonic balance method are compared with those by conventional dual-time stepping method. Results obtained with the present method agree well with those from dual-time stepping method, which demonstrate the ability of the present analysis method to model accurately the unsteady flow. Furthermore， the present method is highly efficient. It is about 36 times fast than conventional dual-time stepping method in the present computation. Then the effects of oscillation amplitude and reduced frequency on unsteadiness of flows are studied. The analysis exploits the fact that, (1) the hysteresis effect of unsteady flow is hardly affected by oscillation amplitude, but the first harmonic unsteady pressure across the blade is proportional to oscillation amplitude; (2) the higher the reduced frequency, the wider the range of unsteady aerodynamic forces, the more intense the hysteresis effect.

Introduction

The flutter and forced response of blade are two limiting factor in the development of turbomachines, and a great deal of blade failures are caused by vibration. Modern turbomachines are designed with higher power, higher aerodynamic loading, higher performance and lighter weight. To meet these requirements, turbomachinery are changed from traditional low-aspect-ratio, midspan dampered blades toward high-aspect-ratio, swept, damperless configurations, leading to vibration problems more serious. In order to deeply study and reveal flutter mechanism in turbomachinery, it is necessary to model unsteady flow around cascades and predict periodic unsteady aerodynamic force loading on blade accurately and efficiently.

At present, numerical modeling unsteady flow around oscillating cascade and analyzing aerodynamic force loading on blade by CFD technique, analyzing the interaction between vibrating blade and flow, has been an important mean and method to study flutter and forced vibration. Time-marching method is a traditional approach in numerical modeling unsteady flow. However, it needs excessively large computational time. Afterward, Jameson developed the dual-time stepping method to enhance the computational efficiency[1]. It is 7~9 times faster than traditional approach, and it has been widely used in computation of unsteady flow [2].

Although dual-time stepping method is more efficient, it is also a huge of computational cost for modeling the complicated unsteady flow. So a great many researched have done by several investigators, and they developed efficient frequency-domain or time-linearized technique to model unsteady flows[3-7], and widely used for modeling unsteady flow in turbomachinery. But, the disadvantage of time-linearized method can not be ignored. It can not model dynamic nonlinearities. Therefore, Kenneth C. Hall[8] developed a harmonic balance technique to compute unsteady nonlinear flows. And the conclusion points out that this method at least one to two orders of magnitude faster than conventional nonlinear time-domain CFD simulations. Hereafter, Kenneth C. Hall[9-12] applied this technique to analyze other unsteady flows, and accomplished a series of outstanding progresses and achievements.

In this paper, the harmonic balance approach for computing unsteady flow around oscillating cascades is presented. The computational results from the present harmonic balance Navier-Stokes method are compared to a time-domain method Navier-Stokes analysis. And then, the effects of oscillation amplitude and reduced frequency on unsteady flows and aerodynamic forces are studied.

Methodology

Governing Equations. An integral form of the two-dimensional unsteady Navier-Stokes equations over a moving finite area ΔA is

$$\frac{\partial}{\partial t}\int_{\Delta A} U dA + \int_{\Delta S}\left[(E - E_v - Uu_{mg})\vec{i} + (F - F_v - Uv_{mg})\vec{j}\right]\cdot d\vec{S} = 0 \tag{1}$$

where $U = [\rho, \rho u, \rho v, e]^T$, E, and F are inviscid flux vectors, and E_v and F_v are viscous flux vectors. u_{mg} and v_{mg} are the x and y components of grid moving velocity, and the extra inviscid flux term Uu_{mg} and Uv_{mg} account for the contribution to the fluxes due to moving grids.

The Harmonic Balance Technique. Because the flow is temporally periodic, the flow variables can be written as a Fourier series at a given frequency ω.

$$U(x,y,t) \approx A_0(x,y) + \sum_{n=1}^{N}\left[A_n(x,y)\cos(n\omega t) + B_n(x,y)\sin(n\omega t)\right] \tag{2}$$

where, A_0, A_n and B_n are the Fourier coefficients of the variable U.

Selecting $2N$+1 points over one temporal period, Eq. 2 can be written in matix form as

$$U^* = E^{-1}\tilde{U} \tag{3}$$

where U^* is the vector of conservation variables at $2N$+1 points in time over one temporal period, E^{-1} is matrix that is the inverse discrete Fourier transform operator, and $\tilde{U}$ is the vector of Fourier coefficients of the flow variables.

Conversely, one can obtain the Fourier cofficients of the conservation variables by Fourier transform.

$$\tilde{U} = EU^* \tag{4}$$

where E is the discrete Fourier transform operator written in matrix form.

Then, we write the Navier-Stokes equations at all sub-time levels simultaneously, so that

$$\int_{\Delta A}\frac{\partial U^*}{\partial t}dA + \int_{\Delta S}\left[E^*\vec{i} + F^*\vec{j}\right]\cdot d\vec{S} = 0 \tag{5}$$

where E^* and F^* are the vectors of x- and y-fluxes evaluated at U^*. Note that the vectors E^* and F^* here include inviscid and viscous terms.

Eq. 7 includes $2N$+1 sets of conservation equations, and the equations are coupled only through the time derivative term, which can be expressed in matrix form

$$\frac{\partial U^*}{\partial t} = \frac{\partial E}{\partial t}\tilde{U} = \frac{\partial E}{\partial t}E^{-1}U^* = DU^* \tag{6}$$

where D is a spectral operator.

Finally, subtitution of Eq. 6 into Eq. 5 gives the desired semidiscrere harmonic balance equations

$$DU^*\Delta A + \sum\left(E^*\Delta S_x + F^*\Delta S_y\right) = 0 \tag{7}$$

where the summation is over the boundary lines of a mesh cell, with the normal projected areas in the two directions being ΔS_x and ΔS_y, respectively.

Numerical Solution Technique. Eq. 7 includes $2N$+1 sets of conservation equations relativing to 2N+1 points in time over one temporal period, and has $4\times(2N+1)$ equations. To solve the harmonic balance equations, a "peseudo-time" term is introduce so that the equations can be solve using a conventional time-marching technique. Thus, Eq. 7 becomes

$$\frac{\partial U^*}{\partial \tau} + DU^* + \sum\left(E^*\Delta S_x + F^*\Delta S_y\right)/\Delta A = 0 \tag{8}$$

where τ is pseudo time, used only to march Eq.8 to steady state, driving the pseudo-time term to zero. Thereby the flow variables on the $2N+1$ points in time is obtained. And then, the Fourier coefficients are solved by applying Fourier transform. Thus, we can obtain flow variables over one temporal period by inverse Fourier transform.

In this work, the time-marching solutions are performed by a four-stage Runge-Kutta integration in the pseudo time. The marching process is accelerated by the local-stepping and implicit residual smoothing techniques.

Validation of Harmonic Balance Technique

A cascade form by the NACA0012 airfoils has been chosen for testing the harmonic balance approach to solve unsteady flow. The computational results by using harmonic balance method compare to those from dual-time stepping approach. The geometric parameters of the NACA0012 cascade are: pitch/chord ratio P/C=1; stagger angle γ=0°. The inlet flow Mach number is 0.593, average incidence angle is 0° and Re=3.21×1 0^6. Each airfoil is subject to torsion mode around its leading edge and the torsion amplitude α=2°. The reduced frequency ($K=\omega c/u_\infty$) is 0.2. Two inter-blade phase angle σ=0°is chosen.

The unsteady flow around oscillating cascade is modeled respectively by two orders harmonic balance approach and dual-time steeping method and the results are given in Fig.1. Here, the unsteady pressure coefficient is presented in the form of $Cp = p^1/[0.5\rho_\infty u_\infty^2 A_m]$, where p^1 is the first harmonic pressure across the airfoil; ρ_∞ and u_∞ are density and velocity in inlet and A_m is the torsion amplitude.

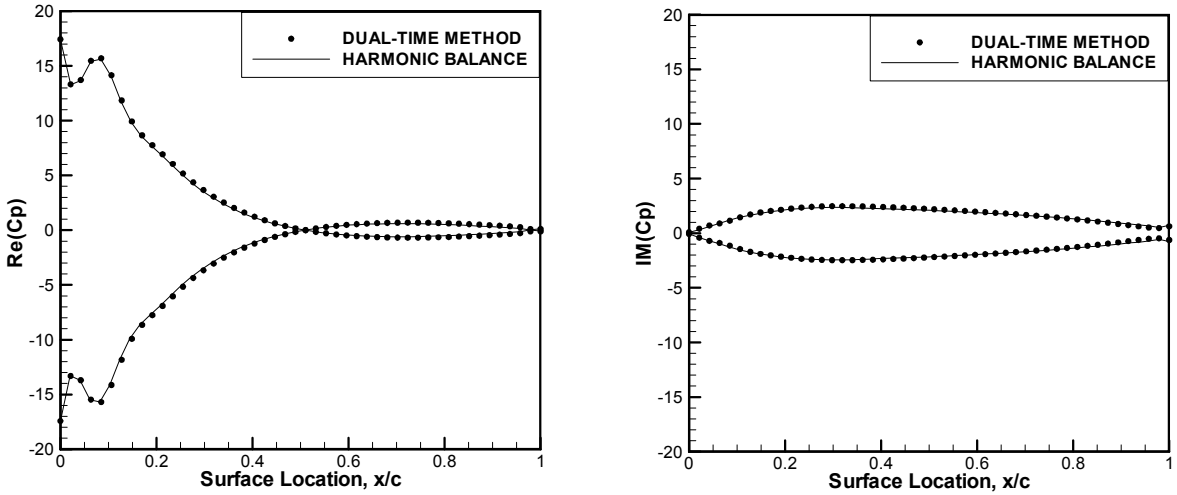


Fig. 1 Comparison of calculated real and imaginary parts of unsteady pressure coefficient on airfoil surface

It can be seen from the figures that the results obtained with the present method and the previous dual-time steeping method are in virtually exact agreement, indicating that the flow around oscillating cascade can correctly predict by present method. More importantly, the computational cost of present method is only one thirty-sixteenth of that by dual-time stepping method. It is obvious that, the harmonic balance technique have the same precision as the dual-time stepping method in solving the periodic unsteady flow, and it requiring much less CPU time, is a efficient computational method.

Numerical Results and Discussions

In order to deeply study the flow characteristics, the unsteady flow around oscillating cascade with various torsion amplitudes and reduced frequencies are numerical solved and analyzed.

The Effect of Oscillation Amplitude. Fig. 2 ~ Fig. 4 shows the computational results of unsteady flow when torsion amplitudes are 2°, 4°, 6° and 8°, respectively. The unsteady pressure jump coefficient is defined as:

$$\Delta Cp = \Delta p^1/[0.5\rho_\infty u_\infty^2 A_m] \tag{9}$$

where Δp^1 is the first harmonic pressure jump across the airfoil.

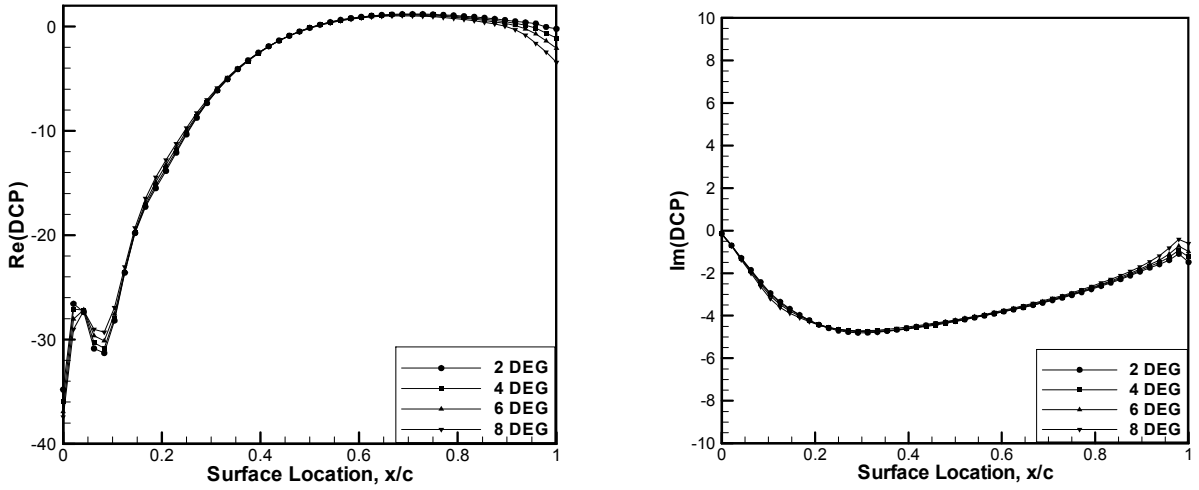


Fig.2 Real and imaginary part of ΔCp on airfoil surface at α=2°, 4°, 6° and 8° 8°

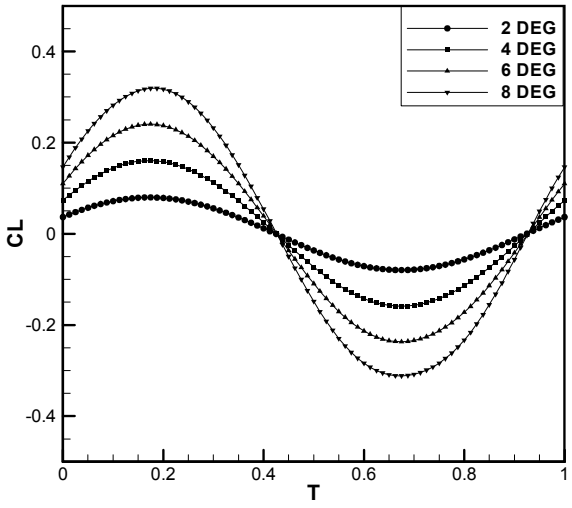


Fig. 3 Unsteady lift coefficient in a period at α=2°, 4°, 6° and 8°

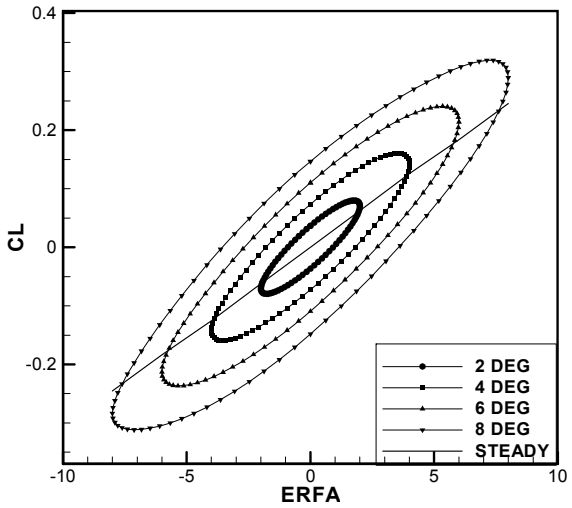


Fig. 4 Unsteady lift coefficient versus angle of attack atα=2°, 4°, 6° and 8°

Fig. 2 shows the real and imaginary parts of the unsteady pressure jump coefficient. It can be seen that ΔCp almost coincide at various torsion amplitudes, only different near the leading edge and trailing edge. So the vibration amplitude has little effect on the unsteady pressure jump coefficient. But we can learn from Eq. 9 that ΔCp is in inversely proportion to torsion amplitude. Thus, the first harmonic pressure jump is almost proportional to the vibration amplitude, only at the leading and trailing edges the relationship between them is nonlinear.

Shown in Fig. 3 are the lift coefficients vary with time in one period at various torsion amplitudes. It shows that when the vibration amplitude increases, the unsteadiness of the lift coefficient is more intense, but the phase seems to keep unchanged. So, the vibration amplitude is no effect on the hysteresis effect of unsteady flow at the same vibration frequency and other conditions.

Similarly, the relationships between unsteady lift coefficient and incidence angle at various vibration amplitudes are shown in Fig. 4. It also shows the relationship at steady state. We see that increasing the torsion amplitude, the range of incidence angle increases, and the range of lift coefficient extends. Conversely, the vibration amplitude of cascade will be large if the unsteadiness of the flow is intense. So, in the work of turbomachinery design, to weaken the forced vibration or avoid the blade flutter, we should weaken the unsteadiness of the flow at full steam.

The Effects of the Reduced Frequency. Keeping torsion amplitude α=2° and other conditions unchanged, the unsteady flow was modeled at K=0.04, 0.2, 0.36, 0.6 and 0.8, respectively. The computational results are shown in Fig. 5 ~ Fig. 7.

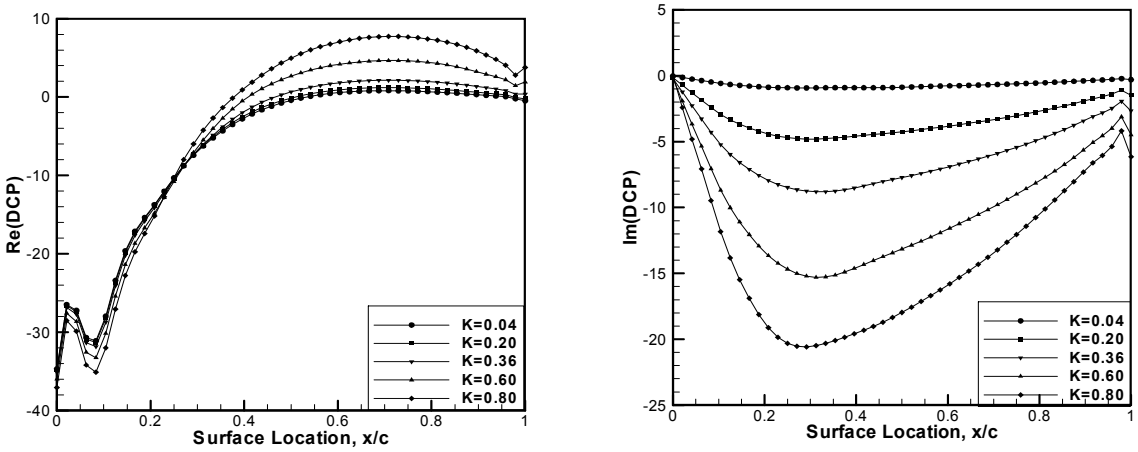


Fig. 5 Real and imaginary part of ΔCp on airfoil surface at K=0.04, 0.2, 0.36, 0.6 and 0.8

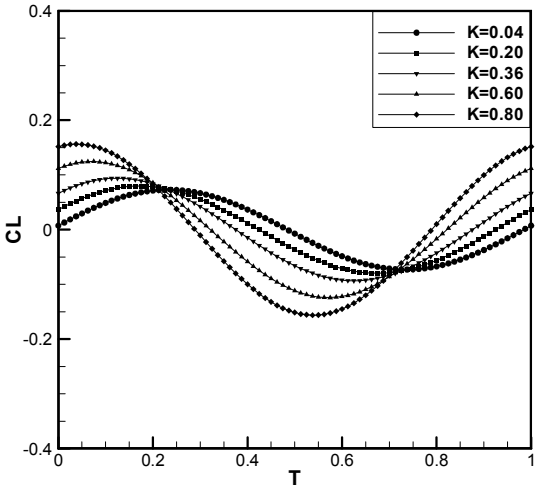


Fig. 6 Unsteady lift coefficient in a period at K=0.04, 0.2, 0.36, 0.6 and 0.8

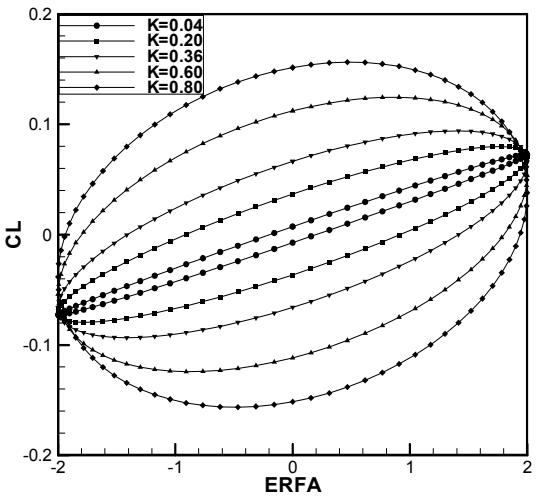


Fig. 7 Unsteady lift coefficient versus angle of attack at K=0.04, 0.2, 0.36, 0.6 and 0.8

Fig. 5 illustrates that the reduced frequency has markedly effect on the unsteady pressure jump coefficient. With the increase of the reduced frequency, the range of the unsteady aerodynamic force on the airfoil surface extends. In Fig. 6, increasing reduced frequency, the amplitude of the lift coefficient increases, and unsteadiness of the aerodynamic force is more intense. It agrees with the conclusion from Fig. 5. It also shows that the phase between the lift coefficient and vibration movement increases and the hysteresis effect is more obvious as the frequency increasing. Fig. 7 shows that the higher the reduced frequency, the more extensive the range of the lift coefficient.

As can be seen from the above analysis that, reduced frequency influence not only on the intensity of the unsteadiness, but also on the hysteresis effect of the unsteady flow. As the frequency increasing, the unsteady fluctuating intensity of the flow increases, the hysteresis effect is more obvious indicating the unsteady flow is more nonlinear. Thus imply that, the frequency of the unsteady fluctuating of the flow have a significant effect on endurance strength and service life of blades. It also can be seen that, at a vibration frequency, the phase between the aerodynamic force and the vibration movement can be to 90°, if the work by aerodynamic force overcome the mechanical damping, the cascade will definitely be in aeroelastic instability—flutter.

Summary

In the present work, a harmonic balance approach and corresponding code has been developed for analyzing the unsteady flow around oscillating cascades. The results from present harmonic balance

method compare with predictions obtained with the dual-time stepping method, indicating that the present method correctly predict the unsteady flow around oscillating cascades. Comparing the predicted results with the two approaches proves that harmonic balance method is efficient to compute unsteady flow. The example shows 36 times computational efficiency by present method in comparison with that by dual-time stepping method. So it can be widely used in computation of periodic unsteady flows.

Furthermore, the influences of oscillation amplitude and frequency on unsteady flow are studied. Firstly, the hysteresis effect of unsteady flow is hardly affected by oscillation amplitude, but the first harmonic pressure across the airfoil is in proportion to vibration amplitude. Secondly, frequency has significant effects on unsteady aerodynamic force and hysteresis effect. Increasing reduced frequency, the range of the aerodynamic force extends, and the hysteresis effect is more obvious.

Acknowledgments

This work is sponsored by the National Natural Science Foundation of China (No. 50376053).

References

[1] A. Jameson: Time Dependent Calculations Using Multigrid, with Applications to Unsteady Flows Past Airfoils and Wings. AIAA Paper 91-1596 (1991).

[2] Y. C. Hu and X. H. Zhou: Application of Dual-Time stepping Method to Calculating Unsteady Viscous Flow around Oscillating Cascade. Vol.21(2003), p. 269-272.

[3] J. M. Verdon and J. R. Capar: Development of a Linear Unsteady Aerodynamic Analysis for finite-Deflection Subsonic Cascades. AIAA Journal, Vol.20(1982), p.1259-1267.

[4] D.S. Whitehead: A Finite-Element Solution of Unsteady Two-Dimensional Flow in Cascades, International Journal for Numerical Methods in Fluids, Vol. 10(1990), p. 13-34.

[5] K. C. Hall and W. S. Clark: Linearized Euler Predictions of Unsteady Aerodynamic Loads in Cascades, AIAA Journal, Vol. 31(1993), p. 540-550.

[6] K. C. Hall and C. B. Lorence: Calculation of Three-Demension Unsteady Flows in Turbomachinery Using the Linearized Harmonic Euler Equations. Journal of Turbomachinery, Vol. 115(1993), p.213-230.

[7] K. Ekici, D. M. Voytovych and K. C. Hall: Time-Linearized Navier-Stokes Analysis of Flutter in Multistage Turbomachines, AIAA 2005-836.

[8] K. C. Hall, J. P. Thomas, W. S. Clark: Computation of Unsteady Nonlinear Flows in Cascades Using a Harmonic Balance Technique. AIAA Journal. Vol. 40(2002), pp. 879-886.

[9] K. Ekici and K. C. Hall: Nonlinear Analysis of Unsteady Flows in Multistage Turbomachines Using the Harmonic Balance Technique. 44th AIAA Aerospace Sciences Meeting and Exhibit; Reno, NV; USA; 9-12 Jan. 2006. p. 1-14.

[10] K. Ekici and K. C. Hall: Nonlinear analysis of unsteady flows in multistage turbomachines using harmonic balance. AIAA Journal. Vol.45(2007), p. 1047-1057.

[11] K. Ekici and K. C. Hall: Computationally Fast Harmonic Balance Methods for Unsteady Aerodynamic Predictions of Helicopter Rotors. AIAA 2008-1439.

[12] K. Ekici, E. K. Robert and K. C. Hall: The Effect of Aerodynamic Asymmetries on Turbomachinery Flutter. AIAA 2009-893.

©(2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.589

Deposition of fluorocarbon films by RF magnetron sputtering at varying target-substrate distance

Z.J.Liu [1, 2, a], Q.Ji [1, 2, b], Y.H.Zhang [1, c], Y.Z. Xia [1, 2, d*]

1 College of chemistry and engineering, Qingdao University, Qingdao 266071

2 State Key Laboratory Cultivating Base of New Fiber Materials and Modern Textile, Qingdao University, Qingdao 266071

[a]liuji198511@163.com, [b]jiquan@188.com, [c]qduzhyh@163.com, [d]polymer@188.com

Keywords: Target-substrate distance; RF magnetron sputtering; SEM; Relative deposition rate

Abstract. The target-substrate distance has a significant effect on the morphology and the relative deposition rate of fluorocarbon films deposited by RF magnetron sputtering. The films were deposited on regenerated cellulose substrate using a polytetrafluoroethylene (PTFE) target, with argon as the working gas. The surface and fracture cross-section morphologies of the deposited films were examined by scanning electron microscopy (SEM). The average thickness of the fluorocarbon films and the relative deposition rates were calculated for varying target-substrate distances. The films had island-like structures composed of nanoscale particles, and the surfaces were not flat. The density of the particles increased and their size decreased with increase of the target-substrate distance in the range 30 to 80 mm. The relative deposition rate increased at first and then decreased with increased target-substrate distance: the optimal target-substrate distance was 50-60 mm.

Introduction

PTFE has a low dielectric constant [1, 2], low friction coefficient [3, 4], ultra-hydrophobic properties [5] and high corrosion resistance [6], and has attracted much attention from researchers. The surface morphology has a significant impact on the properties of fluorocarbon (e.g., PTFE) films. There are many methods for changing the surface morphology, such as varying deposition time, pressure and power in the sputtering process. However, while there are many reports on the effects of those variables on surface morphology [7-9], there have been no reports on the effects of changing the target-substrate distance. We have found that the target-substrate distance has a significant impact on the surface morphology of fluorocarbon films. In experiments on preparation of films on regenerated cellulose substrates by RF magnetron sputtering, we varied the target-substrate distance and determined the surface morphology and relative deposition rate of the films. The optimal range of target-substrate distances for best film quality was determined.

Experiment

Fluorocarbon films were deposited on regenerated cellulose substrates using an RF magnetron sputtering apparatus with 13.56 MHz power supply. The regenerated cellulose substrates were commercially available films (60 g cm^{-2}). They were ultrasonically cleaned in acetone, rinsed in alcohol then dried in hot air. A commercial-grade PTFE sheet bonded to the cathode was used as the target, with argon as the working gas. Both electrodes were cooled using tap water. The chamber was evacuated to a pressure of 5×10^{-2} Pa then working gas was introduced into the chamber to a pressure of 2×10^{-1} Pa. The sputtering target was cleaned for 10 minutes at 100 W power. The discharge parameters that were used are given in Table 1.

The surface and fracture cross-section morphologies of the fluorocarbon layers were studied with a JSM-6390 scanning electron microscope (SEM). The average thickness of the films and the relative deposition rate were calculated as a function of target-substrate distance.

Table 1 Preparation conditions of fluorocarbon films and composite films.

Parameter	
Deposition time [min]	45
Deposition power [W]	150
Deposition Pressure [Pa]	2
Target-substrate distance [mm]	30-80

Results and Discussion

SEM analysis

Conventional SEM is the most common technique for the study of defect morphology. Fig. 1 shows the cross-sectional SEM image of a film formed with target-substrate distance 80 mm.

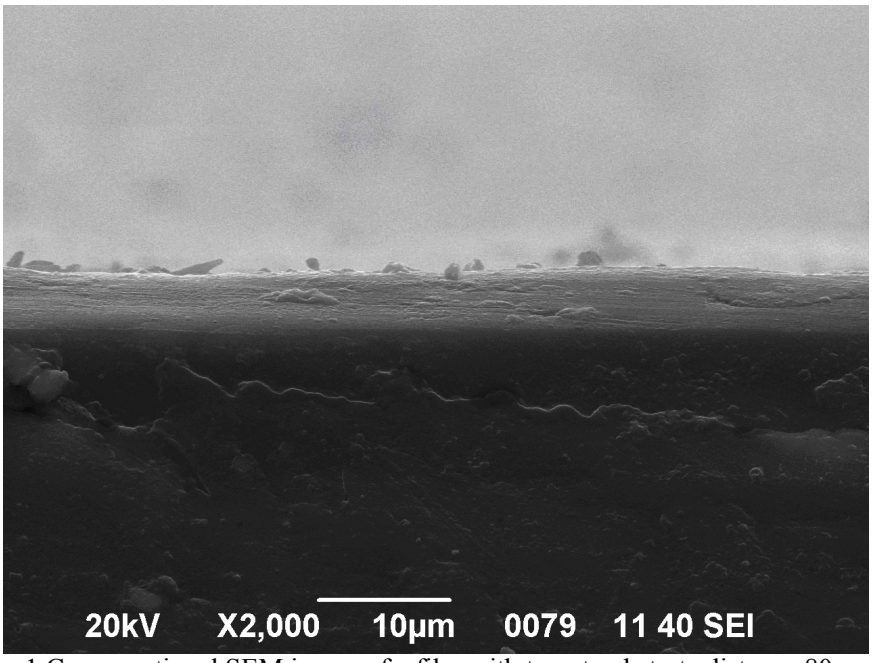


Fig. 1 Cross-sectional SEM image of a film with target-substrate distance 80 mm.

Fig. 1 shows that this combination of substrate and fluorocarbon film is uniform with no holes or cracks, and relatively high compactness. The occurrence of cracking in a sputtered film is mostly caused by differences in sputtering conditions and annealing temperature.

Fig. 2 shows the surface morphologies of the fluorocarbon films. It is apparent that change of target-substrate distance had little effect on the uniformity of the films, all of which were homogeneous. However, it affected the average particle size and compactness of the films. Under certain sputtering conditions, these nanoparticles aggregated to form island-like structures but not bulky congeries, and the island-like structures combined in arrays of polygonal meshes. That is to say, the fluorocarbon particles were not a disordered array but formed a regular island-like structure, as previously reported [10].

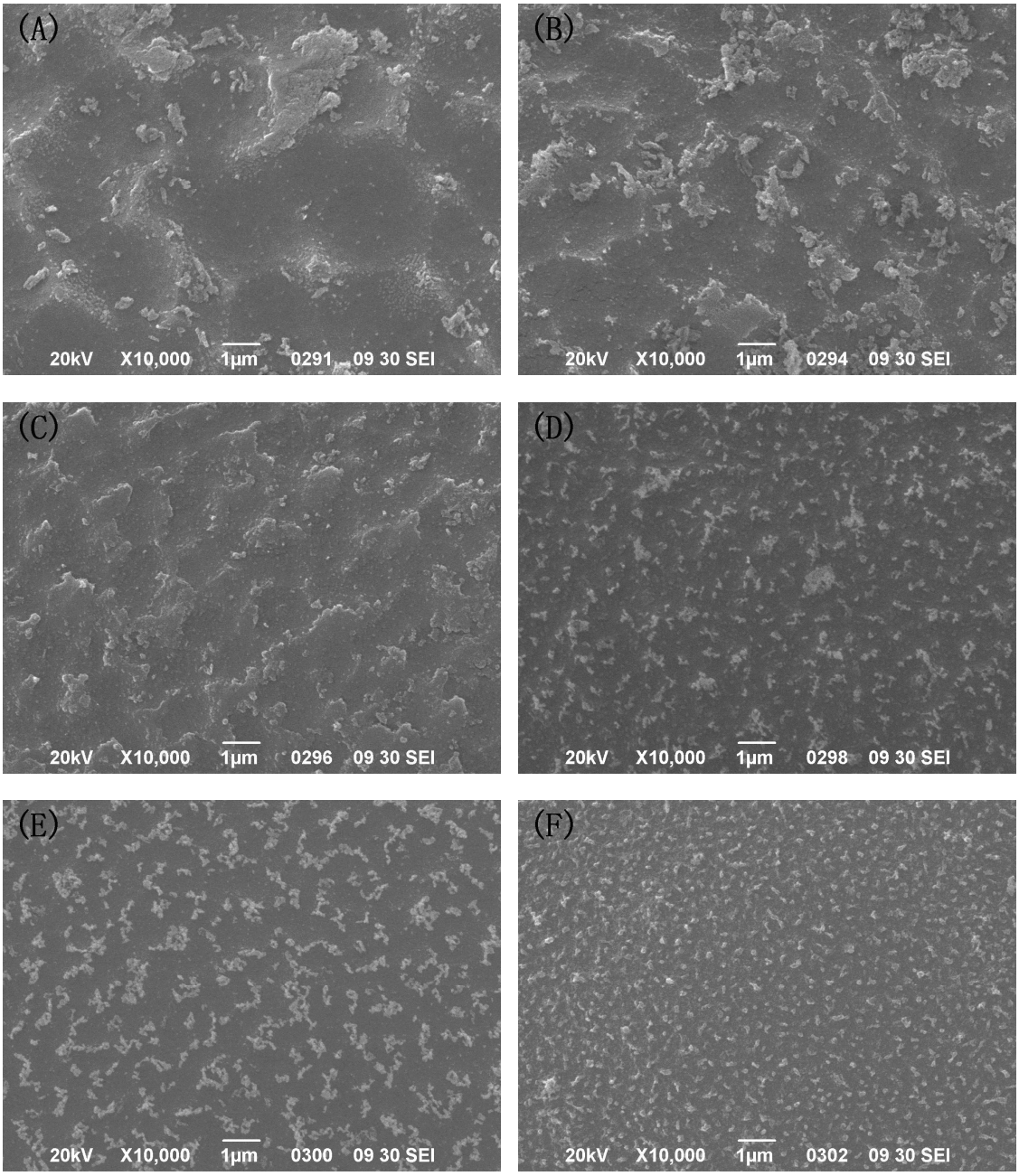


Fig. 2 SEM images of fluorocarbon thin films on regenerated cellulose substrate. The target-substrate distances were (A), 30, (B) 40, (C) 50, (D) 60, (E) 70 and (F) 80 mm.

In images A to C of Fig. 2 the grain size varies from 0.5 to 1 μm, indicating that smaller grains agglomerated to form larger clusters. Coalescence of agglomerated grains or clusters can form continuous films with low particle density. The size of the particles decreased to 50 nm at target-substrate distance 80 mm. Images D to F of Fig. 2 show uniform, compact distributions of grains with sizes in the range 50–200 nm. This phenomenon can be attributed to nucleation and island formation [11]. Because the deposition power was kept constant at 150 W the sputtering ion energy was invariant, so when the target-substrate distance was small, adatoms did not have sufficient time to move on the surface of the films. Thus, only relatively large grains were formed to agglomerate for nucleation. When the target-substrate distance increased, sufficient time was available for the adatoms to have significant surface mobility and form nanoparticles.

Deposition rate analysis

Table 2 shows the relative deposition rate of films corresponding to varying target-substrate distances, for constant deposition time (45 min), power (150 W) and pressure (2 Pa).

Table 2 Relative deposition rates of films at different target-substrate distances.

sample	target-substrate distance [mm]	average thickness [μm]	relative deposition rate [nm s^{-1}]
	L_2	H_m	R_m
A	30	1.74	0.644
B	40	1.82	0.674
C	50	7.31	2.707
D	60	6.95	2.574
E	70	4.68	1.733
F	80	3.60	1.333

From Table 2 it is clear that the relative deposition rate increased at first and then decreased with increasing target-substrate distance. Moreover, when the target-substrate distance increased from 40 to 50 mm, there was a very large increase in relative deposition rate. This phenomenon was modeled as follows.

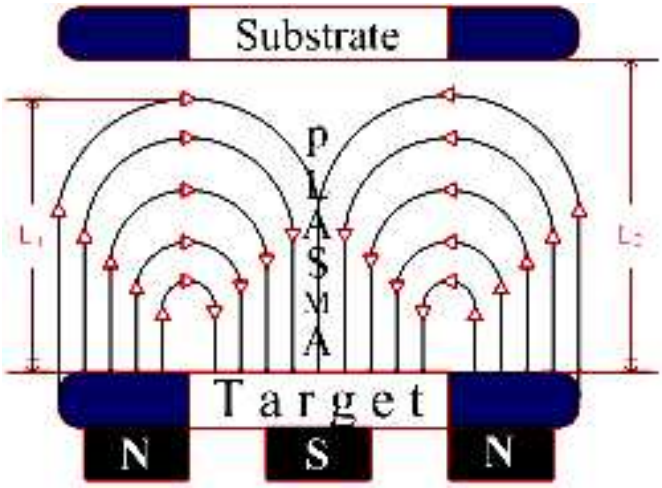


Fig. 3 Schematic diagram of magnetron sputtering.

Fig. 3 shows that magnets below the target create a magnetic field at the surface of the target. The magnetic field can constrain secondary electron motion on the surface of the target. The secondary electrons undergo a cycloid-spiral motion in this region because of the Lorentz force of the magnetic field, and bombard argon atoms, converting them to Ar^+ ions, until the kinetic energy of the secondary electrons is reduced to zero. Ar^+ ions fly at high speed to the cathode in the electric field, bombard the target surface and initiate sputtering. In the sputtering process, the relative deposition rates of fluorocarbon film increase with increase of Ar^+ concentration. With reference to Fig. 3, L1 is the scope of the strongest magnetic field lines loop, and because the magnetic field is an intrinsic attribute of the magnet L_1 was constant. L_2 is the target-substrate distance. On the basis of the data in Table 2, we analyze the deposition rate data as follows.

When $L_1 > L_2$, there are many magnetic field lines that cannot form a loop between the target and substrate, hence most secondary electrons do not have a cycloid-spiral motion, so those secondary electron directly bombard the substrate instead of argon atoms. Consequently, the ionization efficiency in the plasma is low, and the Ar^+ concentration becomes small. As shown in Table II when the target-substrate distance was 30 or 40 mm, there was a low relative deposition rate of fluorocarbon film, and there were high substrate heating effects due to the secondary electrons directly bombarding the substrate [12].

When $L_1 = L_2$, the strongest magnetic field lines just form loops in this closed field region, the secondary electrons undergo cycloid-spiral motion that can sustain bombarding argon atoms, and little Ar+ bombards the chamber walls so there is a high density plasma in the vacuum chamber, giving the highest relative deposition rate of fluorocarbon film. Thus as shown in Table 2 when the target-substrate distance was 50 mm, the deposition rates of fluorocarbon films were 3-4 times greater than when the target-substrate distance was 30 or 40 mm.

When $L_1 < L_2$, the magnetic field lines cannot give good coverage of the entire region between the target and substrate, more Ar^+ fly to the chamber walls and are lost, with the increased target-substrate distance. This situation leads to the Ar^+ concentration becoming smaller in the vacuum chamber, so as shown in Table II when the target-substrate distance increased from 60 mm to 80 mm, the relative deposition rate of the fluorocarbon films became smaller.

Conclusions

Fluorocarbon films were deposited onto regenerated cellulose substrate by RF magnetron sputtering. The surface of the films was relatively rough. The density of the particles increased and the size of the particles decreased with increase of the target-substrate distance from 30 to 80 mm. When the target-substrate distance is optimal, the relative deposition rate of the fluorocarbon films has its maximum value, because the plasma had its highest density under these conditions. This analysis gives an optimal target-substrate distance in the range 50-60 mm.

Acknowledgements

The authors are grateful for financial support from the National Natural Science Foundation of China (No.50673046) and the Program for Taishan Scholar and the Program for Innovative Research Team of Shandong Province, China.

References

[1] D.J. Morrison and T. Robertson: Thin Solid Films Vol. 15 (1973), p. 87

[2] Y. Matsubara, K. Kishimoto and K. Endo: IEDM Vol. 98 (1998), p. 841-842

[3] H. Biederman, S.M. Ojha and L. Holland: Thin Solid Films Vol. 98 (1977), p. 329

[4] C.Z. Liu, R.G. Fairhurst, L. Ren, S.M. Green, J. Tong and R.D. Arnell: Surf.Coat.Technol. Vol. 149 (2002), p. 143

[5] Y. Matsumoto and M. Ishida: Sensors and Actuators Vol. A83 (2000), p. 179

[6] Y. Matsumoto, K. Yoshida and M. Ishida:Sensors and Actuators Vol. A66 (1998), p. 308

[7] Y.H. Zhang and H.J. Qi: Surf.Coat.Technol. Vol. 202 (2008), p. 2612-2615

[8] C. Charton: Vacuum Vol. 68 (2003), p. 65-73

[9] S. Dhananjay, A.B. Bodas and S. Mandale A. Gangal: Appl.Surf.Sci. Vol. 245 (2005), p. 202-207

[10] H. Qi, Y. Fu and D. Wang: Surf.Coat.Technol. Vol. 131 (2000), p. 177

[11] A.N. Banerjee, C.K. Ghosh, K. K. Chattopadhyay, H. Minoura, K.S. Ajay, A. Akiba, A. Kamiya and T. Endo: Thin Solid Films Vol. 496 (2006), p. 112-116

[12] P.J. Kelly and R.D. Arnell: Vacuum Vol. 56 (2000), p. 160-161

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.594

Research on Intelligent Oracle Bone Fragments Rejoining Technology

Aiming Wang[1, 2,a], Jipeng Wang[1,b], Yanqiang Ge[1,c],
Guoying Liu[1,d], Wenying Ge[1,e]

[1] Computer Science Department, Anyang Normal University, Anyang, Henan, 455002, China

[2]School of Computer Science, Wuhan University of Technology, Wuhan, Hubei, 430070, China

[a] wam508@126.com, [b] wangjipeng2002@yahoo.com.cn, [c] lafo2004@sohu.com,

[d] nickyrui@163.com, [e] wh_wtl@126.com

Keywords: oracle bone rejoining; intelligent; contour matching; feature extraction; trend feature.

Abstract. An dynamic suspected target fragment database based on three elements (bone block, fragment and feature) is designed to store pending oracle bone fragments to be rejoined. Oracle bone fragments images are preprocessed on the basis of texture and target/background difference, and then extracted the contour feature, regional feature and trend feature based on actual demand for oracle bone rejoining. An intelligent oracle bone fragments rejoining system prototype is implemented.

Introduction

Oracle bone inscription is a writing system that is found in China, which is the irreplaceable first-hand material to study Chinese ancient history. According to the recent statistics, the number of oracle bones that have been unearthed is as many as 150,000, the number of new findings is unpredictable. Oracle bone inscription is mainly recorded in the oracle bones, the tortoise and shell bones are becoming thinner and increase the possibility of fracture. When the Shang Dynasty perished, some of oracle bones were abandoned in the Yin Ruins (Anyang, Henan Province), which had already folded and been in great disorder. The broken oracle bones when unearthing are resold several times, so the bones being part of a version dispersed to different place. Only rejoin together these fragments as much as possible, and then we can better understand the rhetoric, location and syntax rules of oracle inscriptions, and make a more comprehensive study of inscriptions content.

The broken oracle bone rejoining is an important work in oracle bone inscription research. However, the traditional oracle bones rejoining method that entirely depends on human to organize need heavy workload. Can a new approach be designed to emancipate oracle bones scientists from this heavy work? This is a common academics concern. With the development of computer technology, the method of using computer technology support oracle bones rejoining has become a new topic of oracle bone inscription study, textual research and application. In 1973, Professor Zhou Hongxiang, University of California, Los Angeles, tried to use computer technology in the oracle bones rejoining,. In 1974, this research was also engaged in China, 1/4 fragments of a bone version can be rejoined. Because of the theory, technology, method deficiencies, these study has to be further improved, still has a big gap comparing with the need of oracle bones inscription research.

In August, 2001, the Computer-aided Oracle Bone Fragments Rejoining Research Group is established to study oracle bones rejoining based on fragment, edge, bone version, inscription and handwriting. After nearly 10 years of efforts, important initial results in theory and technology have been made, some results win very popular praise of oracle bones inscription experts through practical application.

Basic Thinking of Intelligent Oracle Bone Fragments Rejoining

Based on the experience of those who rejoin oracle bone fragments, artificial oracle bone fragments are rejoined on the basis of the characteristics, broken trace, radian, color, font style, inscription

content, text strokes, marks of drilling and digging, the direction of divinations, etc. According to the research data owned, this paper's study is based on cherished rubbings, the rejoining conditions have six items of the era, handwriting, bone version, fragment, inscriptions and edge. In addition to the first item, the five items are shown with graphics, and use computers to input images to achieve automatic processing of basic information.

Use technologies of image denoising, segmentation, edge detection, contour tracking, pattern recognition and artificial intelligence, preprocess images of oracle bones which are to be rejoined, get an accurate outline of oracle bones. Through comparing contours piecewise, the system searches candidate oracle bones that can match successfully from the oracle bones database. The oracle bones database stores the image code and contour of all oracle bones.

In this paper, the oracle bones database (database 1, 2) is on the basis of the fragments that need to be rejoined in seven oracle bone inscription works, such as "Oracle Collection", " Oracle Bones at Xiaotun Southland " and so on. All oracle bone fragments selected have been re-copied and added collation, interpretation by oracle bone inscription researchers. Database 1 (Oracular Shells) now contains 9829 fragments to be rejoined, database 2 (Diviner Bones) contains 2622 fragments. The oracle bones database mainly consists of three tables.

Oracle Bone Fragments Rejoining Process

Oracle Bone Image Preprocessing. To improve the patching precision of oracle bone fragments, it is an essential to preprocess oracle bone images, which includes image scanning, image denoising, image segmentation and contour tracing and contour extraction.

Image denoising is the foundation of image segmentation, feature extraction, image matching, etc. The denoising result directly impacts on oracle bone fragments rejoining. Due to the limitations of Shooting devices and transmission channels, some noise is included inevitably in oracle bone images capturing and transferring to computers. Two image denoising methods of the spatial domain and transform domain are uesd is paper: the spatial domain processing method based on the similarity between adjacent pixels uses the mean filtering, median filtering, Wiener filtering and Markov random field model.

Image segmentation separates images into multiple disjoint regions with independent characteristics. To identify and analyze the outline goals, the edges need to be extracted from the images. Here the edge-based and region-based image segmentation methods are used to segment different types of oracle bone fragments. The positive side of tortoise shells choices the region-based image segmentation method, the method extracts a number of same or similar features pixel points to form region from the divided image spatial domain regions with the code of meeting the homogeneity. The reverse side of tortoise shells and other bone pieces choice the edge-based image segmentation method, studying the gray change of each image pixel in some neighborhood, and uses edge operator method to detect the edge.

Uses edge extraction to obtain edge information of fragments. Determination of the contour shape has a significant impact on feature description, fragment matching. The worm-following algorithm is selected to extract oracle bone fragment contour.The operating principle of worm-following algorithm is as follows: a hypothetical worm from advances from a white background pixel area to black background pixel area, the black background pixel area indicates a occlusive contour; the worm turns left and continues moving to the next pixel when the worm enters into a black pixel, the worm turns left again if the next pixel is black, otherwise, the worm turns right if the next pixel is white; This process continues until the worm reaches its starting point.This contour tracking method need repeat many times to get result, sometimes the worm maybe track the same area repeatly to cause the program into an infinite loop.

Edge Segment Extraction Algorithm of Oracle Bone Image. Computer-aided oracle bone fragment rejoining adopts the way of searching the matching contour segments piecewise from the database. Therefore, how to select a suitable edge segments from the contours to be patched becomes

the technology key to the whole system.The existing approach is to set a fixed contour fragment length 'Comp_len', directly intercept 'Comp_len' length contour fragment of a candidate contour from the contour database to be patching to match. However, this interception approach will bring greater problem: a) if Comp_len is too small, it is difficult to express the features of contour segments, more matching oracle bone fragments are identified from the database; b) if Comp_len is too big, the selected contour segment may span multiple matching contour areas of oracle bone fragments. To solve the problem, this paper presents an improved strategy: start matching from smaller Comp_len and set a certain length, increase the length of extraction contour with the success of matching.

Determine the edge segments,only by expressing the characteristics of contour segments reasonably can we conduct the follow-up matching operation. Freeman chain code is used to express contour line here [1,2,3], then calculate the Feature Fourier Descriptors [4,5], statistical moment [6,7], shape function [8] of contour segments. Fourier descriptors assume that the coordinate of each contour position is a complex, the discrete sequence of a complex can be obtained when tracking contour in certain direction. After processing the discrete sequence by using the Fourier transform, use the Fourier coefficients as the shape feature of contours. One-dimensional edge shape is depicted quantitatively by using statistical moment, it can obtain spindle by connecting two vertices of the edge, then normalize the edge curve into a unit area to be dealt with as a histogram, and calculate the mean, variance and high-order moment. Regard the contour segment as a set of points, compute the shape feature value of each edge point, then define a variable associating with each edge point.

Feature Matching Algorithm of Oracle Bone Contour Segments. Have a clear shape description function, also this function has invariant features of translation, optional, scale, etc., and then a well-designed shape matching algorithm will match oracle bone contour segments well. In addition to carrying on characteristic calculation fragment based on edge points related to contour segments in actual calculation, the overall shape features of candidate patching oracle bones is also considered, such as contour center of gravity, radius and other relevant information.

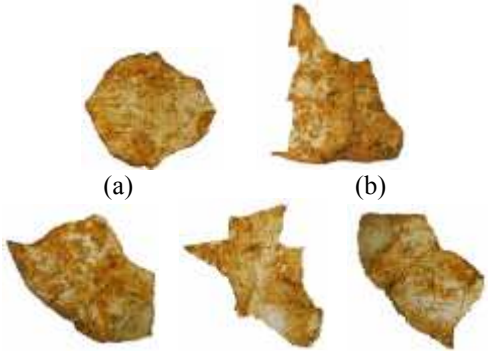
When calculating the features of oracle bone fragment contour segment, the contour segment need to be rotated to couple with contour segments that wait to be rejoined. The first step is to determine the matching criterion when carrying on feature matching for oracle bone fragment contour segment. Matching criterion is the measure to judge two images matching degree. The measure of registration criteria is targeted; there is not a single standard to judge. In practical application, the appropriate registration criteria is selected on the basis of different practical methods.

Experiment

Experiment Description. To test and verify the effectiveness of the proposed oracle bone fragments rejoining method, an rejoining experiment has been conducted.

A oracle bone fragment image that was rejoined by oracle bone experts(shown in Fig. 1) has been scanned from the work of "Oracle Collection", then artificially separate the image into pieces (shown in Fig. 2), select the first separate piece image as the oracle bone fragment to be rejoined, rotate other oracle bone fragment images with a random angle and store each image into the database. In the rejoining process, we selected the simplest contour segment extraction strategy, from the beginning position of contour, extract the contour segment of a given length 'Comp_len' to search from the database, using the pixel-by-pixel moving strategy when searching on contour lines of each candidate oracle bone fragment.

Figure 1. The experimental oracle bone fragment image (already rejoining)



(a) (b) (c) (d) (e)

Figure 2. Oracle bone fragment images to be rejoined: image (a)-(e) respectively corresponding to the label 1-5 region in Fig. 3

To indicate the contour information and take into account the whole information of oracle bone fragment contour which contour segments consist in, the feature calculation method designed is shown in Fig. 3. The red dotted line represents contour segment, O represents the contour center, d represents the average distance of red contour segment to the center O, then the feature of the i th pixel features on the contour segment can be expressed:

$$F_s(i) = \|P_i - O\| - d \tag{2}$$

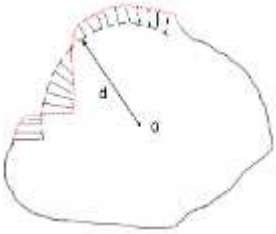


Figure 3. The schematic diagram of feature calculation

P_i is the coordinate of the i th contour point. Calculate the feature of each edge position by using the above feature calculation method. After calculating each pixel position of the red contour segment, we can obtain the feature vector F of the contour segment; its size is Comp_len$\times$1.This paper proposes this edge representation method as the edge histogram.

Process and Result Analysis. In this experiment, we set the binarization threshold 255, the length of contour segment to be matched 75 pixels, the minimum similarity threshold 0.8, and use the edge histogram to represent the contour segment feature. Use the red color to represent contour segments that is matched, the sub step result represents the contour information of oracle bone fragment to be rejoined. The rest four contour images are read from the database, and strongly couple with the contour segment to be matched. The coupling contour segments are displayed in red. Along the contour line of oracle bone fragment to be rejoined, search oracle bone fragment image that can be coupled piecewise in the database. If some coupling oracle bone fragment images are found, then use the red line to represent in the corresponding contour figure.

The experiment result shows that the contour segment to be matched has found a matching contour segment from the first candidate oracle bone fragment image, the similarity is 0.82986. The result is in accord with the real coupling result obtained. Several experiments on oracle bone fragments rejoining have been conducted with this system. Several expected results are initially realized.

Conclusion

The result shows that the system can automatically find objective oracle bone fragments when they exit in the database. In fact, the edges of different oracle bone fragments may be same. That is, for the given oracle bone fragments to be rejoining, there are usually multiple suspect target oracle bone fragments existing in the database, the fragments automatically form the dynamic alternative oracle bone fragments database. Judgment of suspect target oracle bone fragments can achieve through human-computer interaction, according to "time, handwriting and inscriptions".

When oracle bone fragments to be rejoining are selected, the system can automatically generate a suspect target oracle bone fragments dynamic database. Oracle bone inscription experts can rejoin the oracle bone fragments through human-computer interaction, based on the alternative oracle bone fragments database. This will save a lot of time for oracle bone fragments rejoining persons. The future work will study the automatically ultimate rejoining method of suspect target oracle bone fragments, based on artificial intelligence and "age, handwriting, inscriptions".

Acknowledgment

It is a project granted from the National Natural Science Funds of China (NO: 60973051) and Key Science and Technology Project of Henan Province (NO: 092102210112).

References

[1] Freeman H. On the encoding of arbitary geometric configurations. *IRE Trans*, Vol. 10 (1961), p. 260-268.

[2] Freeman H. Shape description via the use of critical potions. *Pattern Recognition*, Vol. 10 (1978), p. 159-166.

[3] Freeman H. Computer processing of line-drawing images. *ACM Computing Surveys*, Vol. 6 (1974), p. 57-97.

[4] Zhan CT, Roskies RZ, Fourier descriptors for plane closed curves. *IEEE Trans. Computer,* Vol. 21 (1972), p. 269-281.

[5] Persoon E, Fu K S. Shape descrimination using Fourier descriptors. *IEEE Trans. System, Man, Cybernetics,* Vol. 7 (1977), p. 170-179.

[6] Teague M R. Image analysis via the general theory of moments. *J. Opt Soc Amer,* Vol. 70 (1980), p. 920-930.

[7] H.S.Liu, Q.F.Shape description function determine methods in the contour shape matching. *Huazhong University of Science & Technology (Natural Science Version)*, Vol. 33 (2005), p. 13-16.

doi:10.4028/www.scientific.net/AMM.50-51.599

A data-driven process for estimating nonlinear material models

X.Y. Kou[1,2,a], S.T. Tan[1,b] and Hod Lipson[2,c]

[1]Department of Mechanical Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong

[2]Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY, USA

[a]kouxy@hku.hk, [b]sttan@hku.hk, [c]hod.lipson@cornell.edu

Keywords: Material model, Nonlinear elastic model, Elasto-plastic, Virtual tensile test, FEA, COMSOL Multiphysics

Abstract. Driven by the wide range of new material properties offered by multi-material 3D printing, there is emerging need to create predictive material models for these materials. A data driven process for estimating nonlinear material model is presented in this paper. In contrast with classical methods which derive the engineering stress-strain relationship assuming constant cross-section area and fixed length of a specimen, the proposed approach takes full advantage of 3D geometry of the specimen to estimate the material models. Give a hypothetical material model, virtual tensile tests are performed using Finite Element Analysis (FEA) method, and the parameters of the material model are estimated by minimizing the discrepancies of the virtual responses and the experimental results. The detailed material models, numerical algorithms as well as the optimization approaches are presented and finally preliminary results are offered.

Introduction

Driven by the wide range of new material properties offered by multi-material 3D printing, there is emerging need to create predictive material models for these materials. Tensile tests are one of the most widely used mechanical tests to derive such material properties as Young's modulus and yield stress. The output of the tensile test is a force-elongation curve, which is usually converted to the stress-strain curve for estimation of material properties. In such a conversion, it is usually assumed that the cross-section area and the length of the specimen are constant in the entire tensile test. The derived stress/strain is referred to as *engineering* stress/strain. Because of its simplicity and satisfactory accuracy in modeling metallic materials, it has been widely used in the past.

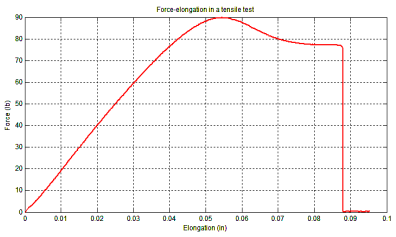
For metallic materials, such engineering stress-strain based parameter estimation proves to be effective, since metals are significantly stiff and the change in the cross-section of the specimen is usually negligible. For polymer materials, such simplifications may introduce significant errors in the estimated parameters. In lieu of such classical approaches, there have been many targeted approaches that tries to get the *true stress-strain* from tensile tests. Arriaga *et al* [1] proposed empirical formulae to convert engineering stress-strain to true Cauchy stress. Grytten et al [2] used Digital Image Correlation (DIC) to measure local strain values in the tensile test of ductile thermoplastic materials.

In this paper, a different data-driven approach is proposed to estimate the material models. We propose different hypothetical material models, and for each material model, the parameters are estimated by minimizing an objective function which relates the model parameters to the discrepancies of the virtual and experimental tensile test data.

Problem Formulation

Fig.1 (a) shows a force-elongation curve derived from a tensile test on a resin-type material FullCure720 used by a rapid prototyping machine Objet PolyJet [3]. As is seen that there exists monotonically decreasing region in this curve where the elongation still increments even though the applied force decreases. In classical plasticity theory, such a phenomenon is generally referred to

unstable behavior of a material [4]. As such, the region of interest in this research is confined in the stable range of this curve, as shown in Fig.1 (b).

(a) 

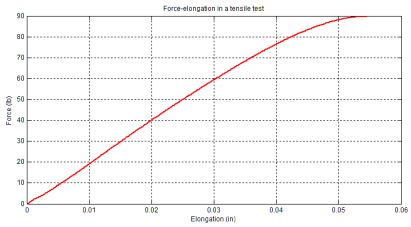
(b) 

Fig. 1 Nonlinear elastic behavior of polymer in a tensile test. (a) The entire *force-elongation* profile in a tensile test; (b) The data in stable range only.

In Fig.1 (b), it is clear that there is certain degree of nonlinearities which can hardly be explained using classical elastic model. To derive a material model for the FullCure720 material, we propose several hypothetical models. For each hypothetical material model, virtual tensile tests are performed using 3D Finite Element Analysis (FEA) method, and the results are compared against tensile test experiments. The parameters of the material model are then estimated by minimizing the discrepancies of the virtual responses and the experimental results.

Let the parameters of a material model be $p = [p_1, p_2, ..., p_m]$, where m is the total number of parameters, then the parameter estimation of a material model is formulated as an optimization problem:

$$\arg\min f(p) = f(p_1, p_2, ..., p_m) = \left| C_e - C_v \right| \quad (1)$$

where C_e and C_v are the experimental and virtual elongation-force curve, $|\bullet|$ is some metric describing the distance/consistency of the two curves.

Material Models and Specimen Geometries

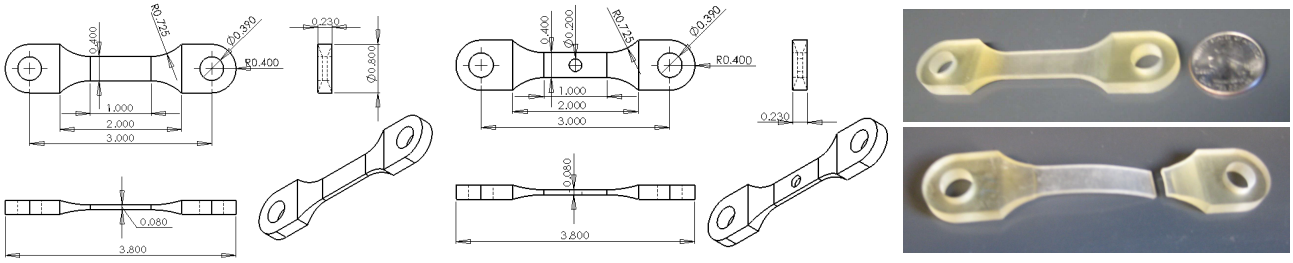
In this paper, two hypothetical material models are proposed. The first *bilinear material model* is characterized with three parameters: the Young's modulus E_0 describes the elastic modulus in the elastic range where the stress is proportional to the strain; the yield stress defines the proportional limit; the tangent modulus E_{tan} is used for modeling the *plastic* hardening behavior. The second *strain-dependent plastic hardening model* is a generalized case of Model (I) in that the plastic hardening behavior is not represented with a constant tangent modulus, but with a general function of plastic strain. In this research, this function is represented by a spline-interpolation function using k interpolation points on the plastic strain-stress plane. For comparative reasons, the most widely used *linear elastic model* is also included as a reference to indicate the relative fitness of each model, as shown in Table 1.

Two specimens with different geometries as shown in Fig. 2 are fabricated with Objet PolyJet 3D printer [3] using the FullCure720 material. Tensile tests are performed on these specimens using the same tensile test machine (MTS 858 Mini Bionix). For each specimen, six tensile tests are repeated with the same strain rate (0.0005 in/sec) applied in all tests.

Virtual Tensile Test and Parameter Estimation Algorithm

The virtual tensile tests are performed using commercial FEA package COMSOL Multiphysics [5]. To simulate the tensile test process, a number of points on the elongation-force plane must be determined. To accomplish this, prescribed displacements are imposed on given geometric faces, and the reaction forces are regarded as output, as shown in Fig. 3 (a). Given a specific set of parameter values (e.g. E_0 and E_{tan}) of the material model and a virtual elongation e, the responses (e.g. reaction forces and von-Mises stress) of the specimen can be evaluated using FEA method, as shown in Fig. 3

(b). By sequentially inputting a collection of virtual elongations to the FEA module, the virtual force vectors can be calculated; by connecting these virtual (elongation, force) points, a virtual tensile test curve can be obtained.

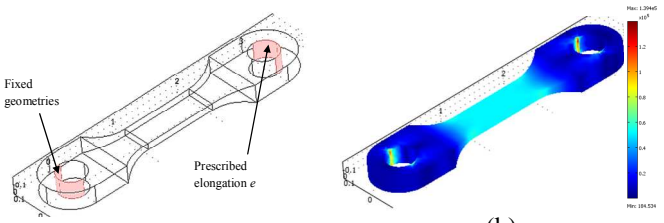


(a) Geometry (I) (b) Geometry (II) (c) Images of Specimen (I)

Fig. 2 Specimen geometries. (a) Specimen (I) with no hole; (b) Specimen (II) with a circular hole. (c) Specimens before and after tensile test.

Table 1 Material models and their parameters.

Material models		Model parameters
(I)	Bilinear material model	Young's modulus E_0; Yield stress σ_{ys} Tangent modulus E_{tan}
(II)	Strain-dependent plastic hardening model	Young's modulus E_0; Yield stress σ_{ys} Spline based plastic hardening function
(III)	Linear Elastic material model	Young's modulus E_0



(a) (b)

Fig. 3 Finite element analysis of the specimen with no holes in the middle. (a) Boundary and loading conditions of the finite element model; (b) von-Mises stress distribution of the specimen.

Let the *virtual* reaction force evaluated at the virtual elongation e_i be $f_{vir}(e_i)$, and the corresponding applied force in the *experimental* tensile test be $f_{exp}(e_i)$, then the discrepancies of the virtual and experimental tensile test curve can be calculated with Eq.(2):

$$\left|C_v - C_e\right| = \sqrt{\frac{\sum_{i=1}^{N}(f_{\exp}(e_i) - f_{vir}(e_i))^2}{N}} \tag{2}$$

Eq. (2) is essentially a Root Mean Squared (RMS) error representing the distance/consistency of the experimental and virtual tensile test curves. By substitution of Eq. (2) into Eq.(1), the objective function of the optimization problem can be then determined.

Fig. 4 shows a schematic flowchart of the proposed numerical optimization method using a *hill climbing* algorithm. Since the analytic gradient of the objective function is unavailable in this optimization, finite difference method is used instead.

Results

The optimization module is implemented with COMSOL Multiphysics with Matlab script. For each specimen geometry, the *averaged* elongation-force curve of six experiments is used as the final curve that a material model seeks to match. In this paper such averaged curves are called *training curves*. By changing the material model parameters iteratively as illustrated in Fig. 4 , the virtual elongation-force curve gradually converged to the training curve, as shown in Fig. 5 (a). Fig. 5 (b) plots the iterative update of objective function values. Here only the results of applying material model (I) on geometry

(I) are presented and similar curve convergence and objective function plots for other material models are omitted here.

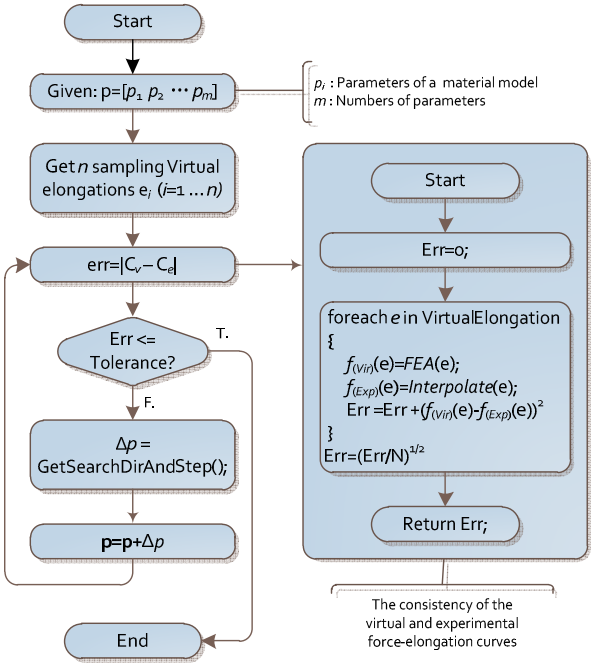


Fig. 4 Flowchart of the proposed numerical optimization algorithm

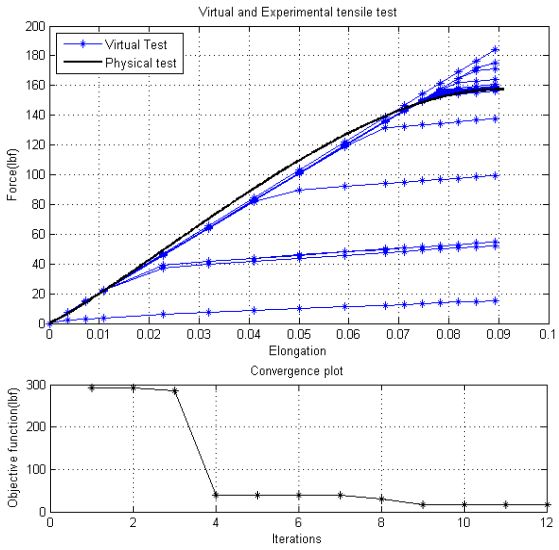


Fig. 5 (a) Iterative update of convergence plot of the virtual curve to the experimental elongation-force curve. (b) Iterative update of objective function values in the optimization process. The geometry of the specimen is shown in Fig. 2 (a).

Table 2 and Fig. 6 show the detailed material parameters derived from the proposed approach. Using the optimized material model parameters, the fitness of these models are further validated against experiment data. The results of using material model (I), (II) and (III) to evaluate the responses of specimen (I) are shown in Fig. 7 (a), (b) and (c) respectively. Due to limited space, the results for specimen (II) are not provided. For each experimental tensile test curve $C_e^{(i)}$, we calculate the RMS error $\Delta_{RMS}^{(i)}$ and maximum error $\Delta_{MAX}^{(i)}$ of the virtual and applied load evaluated at N elongations e_j using Eq. (3) and Eq. (4):

$$\Delta_{RMS}^{(i)} = \left| C_e^{(i)} - C_v \right| = \sqrt{\frac{\sum_{j=1}^{N} (f_{\exp}^{(i)}(e_j) - f_{vir}(e_j))^2}{N}}, 1 \le j \le N \tag{3}$$

$$\Delta_{MAX}^{(i)} = \max(f_{\exp}^{(i)}(e_1) - f_{vir}(e_1),\ f_{\exp}^{(i)}(e_2) - f_{vir}(e_2),\ \ldots,\ f_{\exp}^{(i)}(e_N) - f_{vir}(e_N)) \tag{4}$$

Table 2 Material models and the optimization parameter values.

Material Model	(I)	(II)	(III)
Iterations	12	16	10
Function evaluations	86	360	101
Time (Second)	4029.3	24240	399.3
Optimized parameters (psi)	E_0=130000.1 σ_{ys}=4109.9 E_{Tan}=10003.1	E_0=129998.4 σ_{ys}=3956.9	E_0= 117705

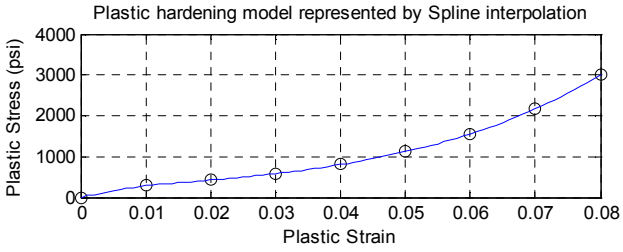
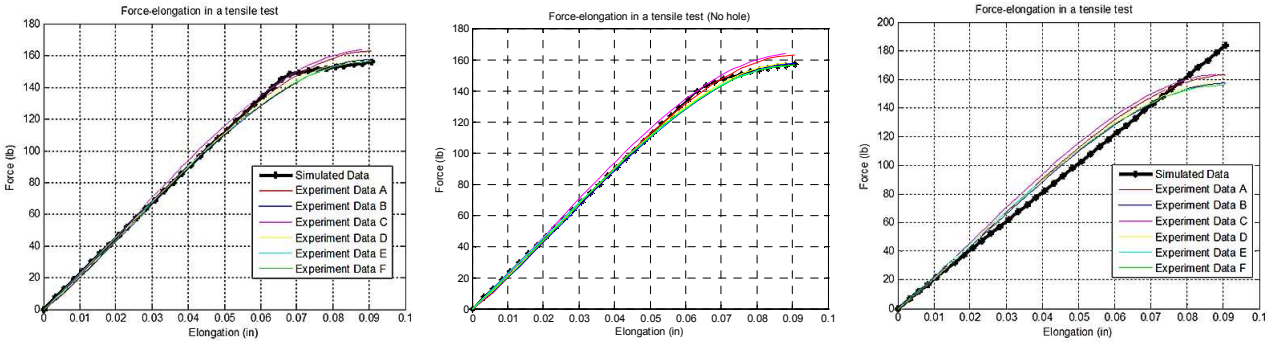


Fig. 6 Plastic hardening functions represented by spline interpolations. The 2nd to 8th ordinates of the data points are parts of the parameters of Material model (II).

For the six evaluated RMS error $\Delta_{RMS}^{(i)}$, the maximum of these errors are defined as the *RMS validation* errors of the material model, and the *maximum validation* errors can be similarly calculated. These validation errors are listed in Table 3.



(a) Material model (I) (b) Material model (II) (c) Material model (III)

Fig. 7 Validation of different material models.

Table 3 Validation errors of different material models.

	Specimen (I) shown in Fig. 2 (a)		Specimen (II) shown in Fig. 2 (b)	
Validation Error (lbf)	RMS	MAX	RMS	MAX
Material Model (I)	25.228	8.680	26.642	12.408
Material Model (II)	24.009	7.648	27.485	13.631
Material Model (III)	60.637	27.487	21.321	10.805

From Table 3, it can be seen that in terms of consistency with experiment results, material Model (II) demonstrated satisfactory performance for both specimens (I) and (II). Material model (III) though slightly outperforms Material (II) for specimen (II), however, the RMS and maximum errors for specimen (I) are much larger than those calculated from material model (II).

Conclusions

An new data driven approach based on virtual tensile tests and numerical optimization is proposed to estimate the parameters of a material model. In the virtual tensile test, 3D geometric information of the specimen is fully exploited, which differs from conventional approaches where only two dimensions (i.e. cross-section area and length of a specimen) were used. Experiments show that this is an effective approach for parameter estimation of a given material model.

Acknowledgement

This research has been supported in part by Objet Geometries. The 1st and 2nd authors would like to thank the University of Hong Kong for the support and the granted professional leave during the 1st author's attachment at Cornell University. Jon Hiller has provided many constructive comments at the initial stage of this investigation. Daniel J. Brooks has helped much in the tensile tests. Their efforts and kind help are appreciated.

References

1. Arriaga, A., et al., *Finite-element analysis of quasi-static characterisation tests in thermoplastic materials: Experimental and numerical analysis results correlation with ANSYS.* Polymer Testing, 2007. **26**(3): p. 284-305.
2. Grytten, F., et al., *Use of digital image correlation to measure large-strain tensile properties of ductile thermoplastics.* Polymer Testing, 2009. **28**(6): p. 653-660.
3. *Objet PolyJet, http://www.interpro-rtc.com/objetPolyJet.html.*

4. Kaliszky, S., *Plasticity: theory and engineering applications* 1990: Elsevier.
5. *COMSOL Multuphysics, http://www.comsol.com/*.

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.605

A New Heuristic Attribute Reduction Method Based on Boolean Matrix

Xie Yan-hong

Department of Computer Science and Technology,Dezhou University,Dezhou,China

dzxieyh@163.com

Keywords: rough set; information system; attribute reduction; boolean matrix; heuristic algorithm

Abstract. Attribute reduction of rough set is a very important research topic. The shortcoming of Reference [8,9]'s method is got a superset of a true reduction sometimes, and the disadvantage of Reference [10]'s algorithm is could not get a right attribute reduction sometimes. To overcome the above shortcomings, a new heuristic attribute reduction method based on Boolean matrix is put forward. Finally, the method's feasibility and validity are proved by examples.

Preface

Rough set theory, proposed by Z.Pawlak in 1982, is a tool for dealing with uncertainty and vagueness aspects of knowledge model. The main idea of rough sets corresponds to the lower and upper approximations based on equivalence relations[1]. Thus, attribute reduction is a very important research topic[2]. At present , there are many attribute reduction algorithms[2-6], and own respective advantages and disadvantages.

Boolean matrix has the performances of simple operation and storage convenience, and the appearing frequency of attribute as heuristic information has the advantage of intuition and easy to accept. So the attribute reduction method of rough sets based on Boolean matrix is a good way. In Reference [7], Rongxing Li put forward the Boolean matrix representation mode of concepts and computation of rough sets and described the relationship among attribute sets, Boolean matrix and linear logical equation group. Based on Reference [7], Reference [8] and [9] adopted linear equation group as attribute reduction model and put forward heuristic attribute reduction method based on Boolean matrix. However, the common disadvantage of the two algorithms is that the result is the superset of a true reduction sometimes. In Reference [10], Aizeng Qian has proved that the solution of linear equation group is only a subset of the attribute reduction set of information system and proposed an improved heuristic algorithm. But the improved algorithm could not get a right attribute reduction when there is no solution to the corresponding linear equation. In the paper, a new heuristic attribute reduction method is put forward. The method starts from the conjunctive normal form of attributes and remove the non-core attribute of lowest significance in the conjunctive normal form expressed by Boolean matrix to get a reduction. The method not only overcomes the shortcoming of algorithm of Reference [8-10] but also take attribute significance as heuristic information and the minimum reduction is got for most conditions.

Preliminaries

To facilitate our discussion, some main concepts of attribute reduction and boolean matrix of information system will be given in this section.

Definition 1 An information system is defined as a family of sets S=<U,A,V,f>, where U is a non-empty universe of objects , A is a finite non-empty set of attributes, V is the value set of A and f:U×A→V is information function.

Definition 2 Let S=<U,A,V,f> be an information system. Given a set of attributes P⊆A, binary relation on attribute set P called indiscernibility relation, where ind(P)={(x,y)∈U×U| $\forall a \in P, a(x) = a(y)$} .

Definition 3 Let S=<U,A,V,f> be an information system. $a \in A$, if ind(A-{a})=ind(A) , then a is called as superfluous attribute, otherwise, a is called as necessary attribute.

Definition 4 Let S=<U,A,V,f> be an information system. If $\forall a \in A$ is necessary, then A is called as independence.

Definition 5 Let S=<U,A,V,f> be an information system, B⊆A, if（1）ind(B)=ind(A);（2）B is independence; Then B is called as a reduction of A.

Definition 6 Let S=<U,A,V,f> be an information system, suppose B_1, B_2, …, B_r are all reductions, then $\bigcap_{i=1}^{r} B_i$ is called as core of A. The attribute of core is called as absolutely necessary attribute; the attribute of $\bigcup_{i=1}^{r} B_i - \bigcap_{i=1}^{r} B_i$ is called as relatively necessary attribute; the attribute of $A - \bigcup_{i=1}^{r} B_i$ is called as absolutely superfluous attribute.

Absolutely necessary attribute is also known as core attribute, and relatively necessary attribute and absolutely superfluous attribute are also known as non-core attribute.

Deficiency of the existing algorithms

The algorithms' main idea in Reference [8] and [9] is adopted linear equation group as attribute reduction model of information system, and remove the column which has the most number of 1 and the rows whose value is 1 in the removed column of Boolean matrix. However, the common disadvantage of the two algorithms is that the result is the superset of a true reduction sometimes.

Example 1 Let S_1 be a information system, attribute set A={a_1,a_2,a_3,a_4,a_5,a_6,a_7}, the corresponding minimum Boolean matrix M1 is shown in Fig 1.

$$M_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Figure 1. The minimum Boolean matrix corresponding to S_1

The result of algorithms in Reference [8] and [9] is {a_1,a_2,a_3,a_4}. In fact, its subset {a_2,a_3,a_4} is a real reduction.

Reference [10] has proved that the solution of linear equation group is only a subset of the attribute reduction set of information system improved the heuristic algorithm in Reference [8] and [9] and proposed an attribute reduction solution based on the solution of linear equation group. But the improved algorithm could not get a right attribute reduction when there is no solution to the corresponding linear equation.

TABLE I. INFORMATION SYSTEM S_2

U	a_1	a_2	a_3	a_4
x_1	1	0	2	1
x_2	1	3	4	1
x_3	2	3	1	1
x_4	3	0	1	1

Example 2 Information System S_2 is shown in Table 1.

{a_1} is obtained by the algorithm in Reference [10]. Obviously, {a_1} is not a true reduction.

In order to overcome the shortcoming of the algorithms in Reference [8-10], the paper put forward a new heuristic attribute reduction method based on Boolean matrix which remove the non-

core attribute of lowest significance in the conjunctive normal form expressed by Boolean matrix to get a reduction

Definition of Boolean matrix and attribute significance

Definition 7[3] Let S=<U,A,V,f> be an information system, where $U=\{x_1,x_2,\ldots,x_m\}$, $A=\{a_1,a_2,\ldots,a_n\}$. Discemible matrix is defined as $M(S)=[c_{ij}]_{m\times m}$, where

$c_{ij}=\{a \mid a\in A,\ a(x_i)\neq a(x_j),\ x_i\in U,\ x_j\in U,\ i,\ j=0,2,\cdots,\ m-1\}$

if C_{ij} is the disjunction of elements of $[c_{ij}]$, then the unique discembile function corresponding to the discemible matrix M(S) is $f_{M(S)}=\bigwedge_{i,j=0}^{m-1} C_{ij}$.

Definition 8 Let S=<U,A,V,f> be an information system, universe U is divided into $X=\{X_1,X_2,\ldots,X_m\}$ under the attribute set $A=\{a_1,a_2,\ldots,a_n\}$, and constructing a boolean matrix $C_A=(c_{ij})_{C_m^2\times n}$, where

$$c_{ij}=\begin{cases}0 & f(X_p,a_j)=f(X_q,a_j)\\ 1 & f(X_p,a_j)\neq f(X_q,a_j)\end{cases}\quad (1\le p<q\le m)$$

Where, $i=(q-1)(n-1)-\dfrac{p(p-1)}{2}+(q-p)$.

The meaning of this boolean matrix is: $c_{ij}=0$ denote that attribute a_j can not distinct class X_p and X_q, and $c_{ij}=1$ denote that attribute a_j can distinct class X_p and X_q.

Definition 9 The following two transformations are equivalent transformation of Boolean matrix: (1) swap the two rows of Boolean matrix; (2) if the vector of ist row is substring [9] of jst one, then delete the jst row.

According to the commutative law and absorption law, the conjunctive normal form corresponding to the Boolean matrix after equivalent transformation is equivalence to the original conjunctive normal form. So the Boolean matrix after transformation is equivalence to the original Boolean matrix.

Definition 10 The Boolean matrix which satisfies any row is not another's substring is called minimum Boolean matrix, denoted by SC_A.

Theorem 1[9] Attribute a_i is the core attribute holds if and only if there exists a row vector a=(0,…,0,1,0,…,0) in minimum Boolean matrix, where, the unique 1 is located in the ist column.

Theorem 2[9] Attribute a_i is the absolute unnecessary holds only and only if the ist column of minimum Boolean matrix SC_D are all zero.

Definition 11 Let S=<U,A,V,f> be an information system, $E\subseteq A$. $\forall a\in E$, the significance of attribute a in E, denoted by SIG(a,E), is defined as the number of 1 in the column corresponding to the attribute a in SC_E.

Heuristic method for attribute reduction based on Boolean matrix

The principle of the algorithm is: under the premise of guaranteeing every disjunctive term of conjunctive normal form of discemible function is true, remove the non-core attribute of lowest significance. So it works in this way until it could not remove further. The remaining attribute set is an attribute reduction. It is noteworthy that when some relative necessary attributes are removed, the remaining attributes' property may changed because the change in their set. For example, some relative necessary attribute have the property of core attribute and some relative necessary attribute have the property of absolutely superfluous attribute. So, this reduction algorithm includes two steps: (1) Generate the minimum Boolean matrix. This step adopt the idea of sampling the Boolean matrix when generating it by referencing Reference [10], and the difference is that the core attribute set is found when the Boolean matrix is generated. (2) Remove the columns corresponding to the non-core attribute of lowest significance in the minimum Boolean matrix and then obtain the new minimum Boolean matrix by equivalent transformation. Repeat this step until the attributes in Boolean matrix are all core attribute.

The description of the heuristic algorithm for attribute reduction based on Boolean matrix as following:

Input: Information System S

Output: An attribute reduction set E

Step 1 Set Boolean matrix M and attribute set E NULL.

Step 2 By referencing the generation idea of Boolean matrix in Reference [10], simple the Boolean matrix M when generate it. If only the component corresponding to attribute *a* in the new generated row is 1 and $a \notin E$, then execute $E = E \cup \{a\}$.

Step 3 Repeat Step 4 to Step 6 until the attributes corresponding to Boolean matrix are all in set E.

Step 4 Find a column m which corresponding attribute is not in set E and the SIG value is minimum in Boolean matrix M. If the attributes satisfying above condition is more than one, then choose one randomly.

Step 5 Remove column m.

Step 6 If the minimum SIG value is not zero (the corresponding attribute is relative necessary attribute), then scan the Boolean matrix M. If there is only one 1 in row n and the corresponding attribute is not in set E, then execute $E = E \cup \{a\}$ and compare the row n and other rows in Boolean matrix M. In the comparing process, if a certain row is a substring of the row n, then delete the certain row.

Step 7 Output set E.

The process of finding the attribute reduction of Information System S_2 in example 2 by this algorithm is:

(1) The minimum Boolean matrix is got after executing the Step 1 to Step 2 as shown in Fig 2(a). This time, the attribute reduction set E is NULL.

(2) Chose the fourth column whose SIG value is minimal and its corresponding attribute is a_4. This column's SIG value is 0 and implies a_4 is a absolutely unnecessary attribute. So delete this column directly and get the new minimum Boolean matrix as shown in Fig 2(b).

(3) Chose the first column whose SIG value is minimal and its corresponding attribute is a_1. This column's SIG value is not 0 and a_1 is not in set E, and implies that a_1 is a relatively necessary attribute. Delete this column and get a_3 and a_2 are all the core attribute after scaning Boolean matrix M and put them into set E. At last, the Boolean matrix after equivalent transformation is shown in Fig 2 (c).

(4) This time, the attributes corresponding to the Boolean matrix M are all in set e. So E={a_2,a_3} is a reduction and it is a true reduction.

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \xrightarrow{MIN\ (SIG)=0,\ E=\{\}} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \xrightarrow{MIN\ (SIG)=2,\ E=\{a_2,\ a_3\}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(a) (b) (c)

Figure 2. the minimum Boolean matrix transformation diagram

In the same way, the reduction {a_2,a_3,a_4} is obtained by the algorithm and it is a true minimum reduction.

Conclusions

The shortcomings of Reference [8-10]'s method are pointed out, and a new heuristic attribute reduction method based on Booolean matrix is put forward. The new method not only makeup the shortcoming of the algorithms Reference [8-10], but also improve the algorithm's efficiency further.

Acknowledgement

This work is partly supported by the Foundation of science and technology of Dezhou city (20090162-8). The author is grateful to Prof. Fengsheng Xu, and the members of project team for helpful discussions, but will take full responsibility for the views expressed in the paper.

References

[1] Pawlak Z. “Rough Sets”. International Journal of Computer and Information Sciences. Vol. 11, pp. 341-356.(1982)

[2] F.S.Xu. “An Attribute and Value Reduction and Rule Extraction Algorithm”, Computer Engineering & Science,Vol. 30, pp. 61-63.(2008)

[3] L.J.Huang, Z.M.Zhang,C.Y.Zhou,M.H.Huang, “Method of Data Reduction Based on Boolean Matrix”. Computer Engineering and Applications, Vol. 34, pp.160-161:181.(2006)

[4] F.S.Xu,T.Z.Li,H.J.Li. “An Improved Discernibility Matrix and the Computation of the Core”. Computer Engineering & Science, Vol. 31, pp. 53-55.(2009)

[5] T.Gao,W.Q.Liu. “A New Method of Complete Data Reduction Based on the Boolean Matrix”. Computer Engineering & Science, Vol. 31, pp. 60-62:98,.(2009)

[6] Y.S.Liu,M.D.Wu,D.Wang. “A Fast Algorithm for All the Minimum Attribute Reductions in Rough Sets”, Computer Engineering & Science, Vol. 29(1):97-100.)(2007)

[7] L.X.Li,S.W.Yun,B.R.Yang. “Boolean Matrix Presentation of Rough Set Concept and Operation”, Computer Engineering, Vol 31,pp. 16-17. (2005)

[8] L.X.Li,S.W.Yun,B.R.Yang. “Attribute Reduction Heuristic Algorithm in Rough Sets Based on Boolean Matrix”, Computer Engineering, Vol. 33, pp. 205-206.(2007)

[9] D.L.Wang. “Knowledge reduction algorithm based on elementary row transformation of Boolean matrix.Computer Applications”, Vol 27,pp. 2267-2269. (2007)

[10] A.Z.Qian. “An Improved Attribute Reduction Heuristic Algorithm Based on Boolean Matrix”, Los Angeles:CSIE 2009, V:769-772. (2009)

©2014 Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.610

The Prediction Research on Ports of Coasts Based on Grey Theory in Fangcheng port

Wen-Fang Liu[1, a] Zhi-Gao Liao[2,b] Chen-Yu Lu[3,c]

[1]Guangxi University of Technology Department of Management Liuzhou City, China
[a] E-mail :wenfang168@163.com [b]E-mail:liaoxie_27@sina.com [c]E-mail:gxutlu@163.com

Keywords-Prediction; Gray theory; GM(1,1);Ports of Coasts

Abstract. This paper use grey theory to establish prediction model to forecast the volume of ports of coasts in Fangcheng Port in Guangxi. Since the access to statistical data of ports of coasts in Fangcheng Port in Guangxi, it is difficult to forecast using traditional method. Through establish the GM(1,1) model , this paper gives the predation on ports of coasts in Fangcheng Port in Guangxi, The predicted results show that the throughput growth faster in the next few years, and will achieve Historical breakthrough in 2014**.**

Introduction

Logistics demand involves goods transportation, distribution, warehousing, packaging, handling loading and unloading, distribution processing, and related logistics information links, when commodity, and waste goods and material etc create demand on space , time and costs of goods. The index system that shows logistics demand involves quantified indicators such as volume of freight, rotation volume of freight transport ,inventory and amount of finish, value terms indicators such as logistics costs, logistics income value-added supply chain. At present, the logistics statistics information issued by Guangxi statistical bureau are composed of 8 main indicators, their are total costs of social logistics, value-added of logistics , total value of social logistics, goods total transporting amount, rotation volume of freight transport, cargo Handled at major ports ,Container Handled at Ports, fixed assets investment related logistics industry.

Guangxi Fangcheng Port is a key area of western development and asean open cooperation, it will play important role to implement master strategy for regional development and a win-win strategy of opening up. It is necessary to construct Guangxi in Fangcheng Port economic region to become open cooperation between China and asean's logistics base, processing and manufacturing base and exchange center of information; become leading, support of western development strategy and openness, high radiation highland, economic prosperity, and a harmonious society, the important ecological good international regional economic zone.

Fangcheng Port coastal ports are in line with the international market and domestic markets and resources. Along with the rapid development of the Fangcheng Port economic zone, the coastal port will realize great-leap-forward development. In order to know future vision about the Fangcheng port development, must have scientific logistics demand forecasting to the development of the logistics industry in Fangcheng Port economic zone. It is difficult to quantify and statistics due to the value index accessibility, therefore this paper selects quantitative index ports in Fangcheng Port regional logistics, so as to provide reference for the strategic development of port.

Grey forecasting model method

The introduction of Grey theory

Due to imperfection and availability of statistical data on port cargo of Fangcheng Port, we use GM(1,1) model based on gray theory through comparison and study to forecast port cargo of Fangcheng Port. After analyzing the data on port cargo volume of recent years, we establish the port cargo volume forecasting model by use of grey theory. Now there are some factors to use this method: because the growth of port cargo is effected by economic development, industry policy,

natural condition and industrial structure ect. Among these factors, some are certain and others are not, so it can be seen as a grey system. The grey prediction method can better grasp system evolution law. At the same time ,it keeps out of fatal weakness about related insufficient data, also can avoid the subjective judgment caused by the personal experience, knowledge and preferences and macro policy factors.

GM(1,1) model is a first derivative differential equation model with a variable, it is suitable for the prediction to system development and change of behavior characteristic value. The essence is to give once accumulation Generating Operation sequence（1-AGO） through a series of original data, which generate data sequences present certain rules and structure prediction model.

GM(1,1) model

The basic process of modeling of Grey prediction model as follows:

If there are n primitive data, the original sequence is $x_k^{(0)}$,and that $x_k^{(0)} = (x_1^{(0)}, x_2^{(0)}, ..., x_n^{(0)})$.

Accordingly, the sequence of accumulation generation is $x_k^{(1)}$， and $x_k^{(1)} = (x_1^{(1)}, x_2^{(1)}, ..., x_n^{(1)})$

$$x_i^{(1)} = \sum_{n=1}^{i} x_n^{(0)} \qquad (1)$$

Accumulation generation sequence must be smooth .If sequence meet the following a few requirements, is considered a smooth. Smooth degree can be expressed by ρ_k

$$\rho_k = \frac{x_k^{(0)}}{x_{k-1}^{(1)}} \qquad (2)$$

$$\frac{\rho_{k+1}}{\rho_k} \prec 1, k = 2,3,...n \qquad (3)$$

$$\rho_k \in [0,0.5], k = 3,4,...n \qquad (4)$$

Tectonic background values sequence $z_k^{(1)} = (z_1^{(1)}, z_2^{(1)}, ..., z_n^{(1)})$ on the basis of $x_k^{(1)}$

$$z_k^{(1)} = \frac{x_k^{(1)} - x_{k-1}^{(1)}}{\ln x_k^{(1)} - \ln x_{k-1}^{(1)}}$$

Suppose $x_k^{(1)}$ have approximate law of permeability changing , the equation for:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \qquad (5)$$

Discretize (5), get Gray differential equation

$$x_k^{(0)} + az_k^{(0)} = b \qquad (6)$$

Using least-square method, a and b parameters can be given,

$$\alpha = [a,b]^T = (B^T B)^{-1} B^T Y \qquad (7)$$

Among (7)

$$B = \begin{bmatrix} -z_2^{(1)} & 1 \\ \vdots & \vdots \\ -z_n^{(1)} & 1 \end{bmatrix} \qquad (8)$$

$$Y = \begin{bmatrix} x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{bmatrix} \qquad (9)$$

The formulas of predicting model were established：

$$\hat{X}^{(1)}(k+1) = (X^{(0)}(1) - \frac{b}{a})e^{at}$$

$$\hat{x}_{k+1}^{(1)} = (x_n^{(1)} - \frac{b}{a})e^{-a(k-n+1)} + \frac{b}{a}, k = 1,2,\cdots,n \quad (10)$$

Then restore prediction value is

$$\hat{x}_{k+1}^{(0)} = \hat{x}_{k+1}^{(1)} - \hat{x}_k^{(0)}, k = 1,2,\cdots,n \quad (11)$$

The prediction precision of model is inspected by error analysis.

If $\hat{x}_k^{(0)} = (x_1, x_2, \cdots x_n)$ was analog sequence, then residual inspection for:

$$\varepsilon_k = x_k^{(0)} - \hat{x}_k^{(0)} \quad (12)$$

The relative error inspection for:

$$\delta = \frac{|\varepsilon_k|}{x_k^{(0)}} \quad (13)$$

From above the modeling process can be seen, the essence of Gray model is made index simulation after raw data accumulation generation. Fitting precision depends on the value of a and b, the background value and initial conditions. According to the literature, the background value and initial conditions have improved in this paper.

The prediction on ports of coasts in Fangcheng Port

Through searching, get follow data:

Table Ⅰ The original data of relative index

year	2000	2001	2002	2003	2004	2005	2006	2007
Ports of coasts	919	1003	1116	1320	1609	2006	2506	3032

$$x_k^{(0)} = (919,1003,1116,1320,1609,2006,2506,3032)$$

$$x_k^{(1)} = (919,1922,3038,4538,5967,7973,10479,13511)$$

The time response function can be obtained according to the grey forecasting model

$$\hat{x}_k^{(1)} = 4122.83e^{0.1986k} - 3203.826$$

Which can be obtained by simulation, value of simulation and the relative error of simulation test value in 2000 to 2008, see table Ⅱ below.

Table Ⅱ Simulation statistics of Guangxi Fangcheng Port port cargo throughput prediction

year	$x_k^{(0)}$	$\hat{x}_k^{(0)}$	ε_k	δ
2000	919	907.23	1.56	0.00082
2001	1003	906.15	-96.84	0.096
2002	1116	1105.32	-10.68	0.21
2003	1320	1348.26	28.25	0.22
2004	1609	1644.59	35.58	0.00002
2005	2006	2006.05	0.05	0.22
2006	2506	2446.96	-59.03	0.11
2007	3032	2948.78	-47.22	0.15

What Can be seen from table Ⅰ,the effect of simulation residual and relative error is better.

According to the grey model, can calculate the 2008-2015 port freight throughput of Guangxi Fangcheng port, see table Ⅲ.

Table Ⅲ The prediction value of Guangxi Fangcheng Port port cargo

years	2008	2009	2010	2011	2012	2013	2014	2015
The prediction value	3640	4441	5417	6607	8060	9831	11992	14628

Conclusion

(1) In this paper, gives the prediction model to predict the Guangxi Fangcheng Port cargo according the data of official statistics and compounding some scholars' research achievements get a good prediction results. The predicted results show that the throughput growth faster in the next few years, and will achieve Historical breakthrough in 2014.

(2) Using the grey system theory model, can overcome the limitations of related data and avoid the effect of factitious factors. The simulation results and the prediction model is close to practical value, the predicting error is lesser, model precision is better, this shows that the prediction method based on the grey theory is suitable for Guangxi Fangcheng Port cargo throughput forecast This method is feasible and effective and has important reference value on the development of the Fangcheng Port planning in Guangxi.

(3) Along with the development of things, some of the future disturbance factors will enter the system to influence the system, when using the model to give further extrapolation, will find grey intervals more and more big, that is, the effect of long-term forecast will get worse if using this model, therefore this model can not use to predict future all values. The model parameters should adjust continuously according to the new information, avoiding "once for all" in estimation of grey parameters in model. The practice shows that, to better fit to the change trend of system which is unknown and changing, there is no other choice but to increase new information, a new data constantly.

(4) The main purpose of this paper is to explore scientific and effective method to predict port throughput, and to provide a thought to solve the problems, not to provide a "universal" formula. Any prediction technology has its limitations, although grey model can forecast throughput, but still exist deficiencies, we will explore constantly more effective and more scientific prediction method in future.

Acknowledgment

This research was supported by the Technology development Projects of Guangxi Province (0992006-1)and the Foundation of Research Projects of Department of Education of Guangxi (Nos.200911LX165, 200911LX168, and 200911MS117) and Science Foundation funded project of Guangxi University of Technology (0977206) .

References

[1] Liu si-feng,guo tian-bang,dang yao-guo ect, "The theory of grey system and its application (second edition)"[M].*Beijing:Science Press*,1999

[2] Zhang rui,gao huanwen, "Based on the grey GM (1,1) model of agricultural mechanization levels predicted,"[J] *Transactions Of The Chinese Society Of Agricultural Machinery*, 2009, 40(2):91-95

[3] Tian yi-jun,zhu fengfeng. Clerk Maxwell, "Based on the GM (1,1) model and gray relational combined forecasting methods" [J]*Science Technology and Engineering* , 2009,9(4): 842-844,851

[4] China Statistics Press Editorial Board. *China Statistical Yearbook, 200*8, China Statistics Press

[5] China Stat ist ics Press Editorial Board. *China Stat ist ical Yearbook, 200*9, China Stat istics Press

[6] Deng ju-long. “Grey Prediction and Decision[M]”.wuhan:*huangzhong Science and Technology University Publishuing* House,1986.

[7] Liu wen-fang, “The application of GM (1,1) model in Prediction of guangxi sugar production”, *Journal of Guangxi University of Technology* 2009,20(2):27-29

doi:10.4028/www.scientific.net/AMM.50-51.615

Finite Element Simulation of Springback in Sheet Metal Forming

Lei Chen[1, a]

[1]Institute of Mechanical & Electrical Engineering Design, Jiangxi Blue Sky University, Nanchang, 330098, China

[a]ujs_cl@163.com

Keywords: Sheet metal forming, Springback, FEM

Abstract. Springback during unloading affects the dimensional accuracy of sheet metal parts. A finite element model is proposed to predict springback with contact evolution between the sheet and dies based on U-bending benchmark of Numisheet 1993. The underlying formulation is based on updated Lagrangian elastoplastic materials model. BHF is modeled as an effective force on the edge of the flange. Sheet unloading is simulated by the tools' reverse motion. Sensitivity analysis done using the finite element method (FEM) demonstrated that the magnitude of springback and the overall accuracy are highly influenced by the element size, integration points and yield criteria. The solutions validated with experimental data of Numisheet 1993 show more accurately than 90%.The optimal model derived from FEM analysis provides an accurate solution and small convergence. It also has merit in not requiring excessive computation time.

Introduction

Springback is a common phenomenon in sheet metal forming, caused by the elastic redistribution of the internal stresses during unloading. It has been recognized that springback is essential for the design of tools used in sheet metal forming operations because it affects the dimensional accuracy of sheet metal parts [1]. Traditionally, springback design based on trail and error method is a long cycle and high cost. In these years, a large amount of high strength steel and aluminium alloy sheets are used in car panels owing to the requirement of safety, economy and environment. These kinds of sheets have high Yang's modulus and high springback using in the car panels.

With the development of the computer science and FEM theory, numerical simulation of springback is performed. Since 1993, four benchmarks (U-bending, S-rail, AUDI door and unconstrained bending) are setup for springback prediction in Numisheet. And all the benchmarks of Numisheet 2005 have something to do with springback. There are mainly two kinds of finite element methods for springback simulation: static implicit method and dynamic explicit method [2, 3]. The advantage of static implicit method is the non-condition stability and no time step limit, whilst it is possible not converged due to the iterations. The dynamic explicit method does not need iterations, but the time step is restricted by the system. At present, the accuracy of springback prediction is not high commonly owing to mainly two reasons. One is the stress field acquired from simulation is not accurate. Another is the method using for simulation. The factors affecting stress field are materials model, constitute equation, yield criterion, element type, element size and contact algorithm, etc [4]. There are three problems in finite element solution for springback prediction. Firstly, since the flexibility of the whole panel is high, the boundary condition needs to be added to cancel the rigid displacement. Different boundary conditions, different simulation results. Secondly, local contact between the blank and tools exists during unloading, whilst it is often treated as the whole elastic unloading in many commercial softwares such as Ls-dyna. Thirdly, some material parameters are changed during springback, but often omitted.

In this paper, a finite element model of U-bending benchmark of Numisheet 1993 was proposed to simulate springback with contact evolution between tools and blank based on static implicit method. The simulation results were compared with the experiment. Then the effects of integration points, element size and yield criteria on springback results are investigated.

Finite element model

A commercial finite element analysis package MSC.MARC was used for our numerical simulation of the U-bending forming process (see Fig.1). Since the problem is symmetric, only one half of the length of the entire blank (175 mm×35 mm) was modeled. The binder, the die, and the punch are modeled as three separate rigid surfaces, and a modified Coulomb friction model is assumed.

$$F_t = -\mu F_n \cdot \frac{2}{\pi}\arctan\left(\frac{v_r}{C}\right)t \tag{1}$$

where F_t is tangential force, [N]. μ the friction coefficient, here is 0.162. F_n is normal reaction force, [N]. v_r is the relative sliding speed. C is a constant connected with v_r, here is 0.1. t is the tangential vector in the direction of the relative speed.

$$t = v_r / |v_r| \tag{2}$$

The blank is model as shell element (MARC type 139) with a mesh of 70×14. Boundary conditions are specified to create a symmetric condition. The material is modeled to be anisotropic, elastoplastic following the Barlat 1991 yield criterion and isotropic strain hardening. The plastic behavior of the sheet material is modeled using a power law relation ($\sigma = K(\varepsilon_0+\varepsilon)^n$), where σ is stress, K is material strength coefficient, ε_0 is initial strain, ε is strain, n is strain hardening coefficient. The material parameters are shown in Table1.

The blank holding process is modeled as a blank holder gap with an effective force exerted on the edge of the flange (see Fig.2). The blank holder gap is used to control the wrinkle of the flange, and the effective force controls the blank flow [7]. The effect force can be expressed as follows.

$$F_2 = \frac{2\mu F_1}{s} \tag{3}$$

where F_2 is the effective force, [MPa]. F_1 is actual blank holding force, [N]. s is the sectional area of the blank perpendicular to F_2, [mm^2].

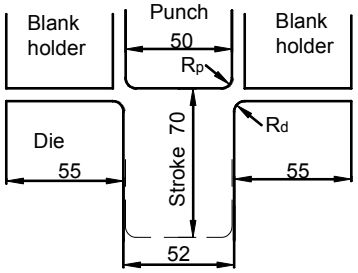


Fig.1 U-bending geometry

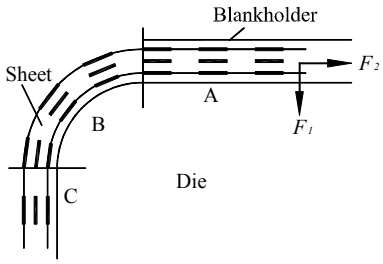


Fig.2 Force exerted on the flange

Table 1 Materials parameters

	E[GPa]	ν	Y [MPa]	r	Stress-strain relation [MPa]
0°			137.0	0.71	
45°	71.0	0.33	136.4	0.58	$\sigma = 579.79(0.01658+\varepsilon)^{0.3593}$
90°			136.4	0.7	

Two different strategies can be used to simulate springback during unloading. The first one can be understood as a simple continuation of the forming process, where the tools' motion is reversed and calculations are carried out until the loss of the contact [4]; in the second one, at the end of the punch stroke the tools' action is replaced by the corresponding forces and these are then consecutively decreased until vanished [5]. In this paper the first one is used because its physical meaning seems to be in better agreement with the real processes. The whole forming simulation is divided into three parts: firstly, the die and blankholder is stationary, the punch moves in a fixed speed. Then the punch

is stationary, the motions of the die and the blankholder are reversed. Finally, the punch's motion is reversed

Results and discussions

Simulation results and validation. Fig.3 shows the diagram for the springback measurement. Fig.4 shows the shape of the blank before and after springback. Table 2 shows the comparison of simulation and experiment results. Error Δ=|Simulation-Exp.|/Exp.. From the table, we can see that the simulation results of θ_1 and θ_2 agree well with the experimental results of Numisheet 1993 [1], whilst the ρ have a larger divergence and the CPU time is long about 8.3 hrs (29891 seconds). So it should be optimized.

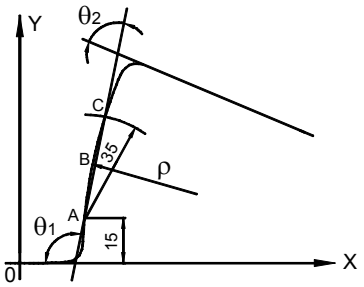


Fig.3 Diagram of springback measurement

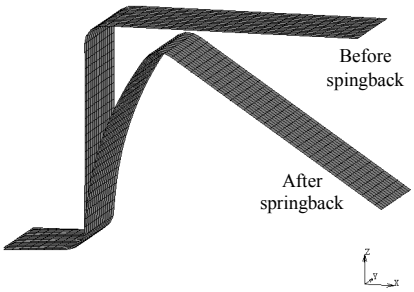


Fig.4 Shape of the blank before and after springback

Table 2 Comparison of simulation and experimental results

	θ_1 [°]	θ_2 [°]	ρ [mm]
Avg. (Exp.)	112.4	72.8	106.0
Simulation	114.9	69.7	92.0
Δ	2.2%	4.3%	13.2%

Element size. Due to the geometrical symmetry, only one quater (175 mm×17.5 mm)of the global specimen was simulated. From the analysis above, the blank under the blank holder is almost undeformed, so the element size can be larger and has little effect on springback. Whilst the element size of side wall is the focus owing to the large deformation. This portion of the specimen was discretized with a 4-node mesh onto five different plane mesh refinements.

Table 3 shows the simulation results with different mesh refinements. From the table, we can see that with the refinements of mesh, the CPU time increases rapidly. When the mesh refinements reach 5/3 mm, the simulation results agree well with the experiment more than 90%. So the c refinement of 5/3 mm is the best.

Table 3 Simulation results with different mesh refinements

Mesh	Size [mm]	Element	Node	θ_1 [°]	$\Delta\theta_1$[%]	θ_2 [°]	$\Delta\theta_2$[%]	ρ [mm]	$\Delta\rho$[%]	Time [s]
a	5	140	180	91.7	18.4	84.6	16.2	797.4	652.3	1306
b	5/2	232	295	108.3	3.7	71.9	1.2	92.5	12.7	3795
c	5/3	324	410	112.4	0	71.7	1.5	96.4	9.1	6036
d	5/4	416	525	112.2	0.2	71.6	1.6	99.4	6.2	12714
e	1	508	640	113.2	0.7	72.1	1.0	97.3	8.2	20630

Integral points. The number of through-thickness integration points (NIP) required for accurate springback analysis following sheet forming simulation is the subject of confusion and controversy. Li and Wagoner recommended based on a finite element analysis (FEA) of draw-bending springback, the use of 25 integration points (IP) [6]. They also reported that up to 51 IP could be required to obtain springback solutions with accuracies of 1% [3]. Several researchers [6] since then have reported that NIP between 5 and 11 are adequate, and even that 7 or 9 IP are optimal, with reduced accuracy for more IP. So we do research on this. Table 4 shows the simulation results with different NIP. From the

table, we can see that CPU time increases with the increase of the NIP, whilst the simulation results is stable when NIP is larger than 5. Taking CPU time and accuracy into account, NIP=7~9 are optimal. The different NIP results may be owing to the different amount of the plastic deformation. When the plastic deformation is small, the springback is larger and NIP is larger to get a accurate simulation result. On the contrary, only a small NIP is needed.

Table 4 Simulation results with different NIP

NIP	θ_1 [°]	$\Delta\theta_1$[%]	θ_2 [°]	$\Delta\theta_2$[%]	ρ [mm]	$\Delta\rho$[%]	Time [s]
3	106.3	5.4	75.1	3.2	125.1	18.0	4541
5	112.4	0	71.7	1.5	96.4	9.1	6036
7	111.5	0.8	72.1	1.0	100.3	5.4	7913
9	111.8	0.5	71.7	1.5	95.8	9.6	9626
11	110.9	1.3	71.8	1.4	101.8	4.0	11161

Yield criteria. There are several yield criteria for simulation, Von. Mises, Hill and Barlat. Table 5 shows the simulation results. From the table, we find that the angle of θ_1 and θ_2 both agree well with the experiment. However, the bending radius of ρ is in a large variety. Only using Barlat yield criteria agree well with the experiment. And when m=8, the simulation result is the best.

Table 5 Simulation results with different yield criteria

		θ_1 [°]	$\Delta\theta_1$[%]	θ_2 [°]	$\Delta\theta_2$[%]	ρ [mm]	$\Delta\rho$[%]
Von Mises		112.5	0.1	70.9	2.6	92.0	13.2
Hill		112.1	0.3	70.9	2.6	93.5	11.8
Barlat	m=6	111.4	0.9	71.3	2.1	96	9.4
	m=8	111.5	0.8	72.1	1.0	100.3	5.4

Summary

In this paper, a finite element model is set to predict springback with contact evolution between the sheet and dies based on U-bending benchmark of Numisheet 1993. Sensitivity analysis done using the finite element method (FEM) demonstrated that the magnitude of springback and the overall accuracy are highly influenced by the integration points, element size and yield criteria. A conventional optimization method combined with FEM was used to obtain optimal parameters that can predict springback more accurately than 90%.

Acknowledgement

The financial support for this work is from Youth Foundation of Jiangxi Educational Committee (No. GJJ09620) and is greatly appreciated.

References

[1] G. Liu, Z. Lin, and Y. Bao: Finite element in analysis and Design Vol. 39 (2000), p. 107
[2] N. Narkeeran, L. Michael: Finite Elements in Analysis and Design Vol. 33 (1999), p. 29
[3] K. Li, W. Carden, and R. H. Wagoner: International Journal of Mechanical Sciences Vol. 44 (2002), p. 103
[4] P. Y. Manach, S. Thuillier, and L. F. Menezes, et al: Proceedings of Numisheet 2002, p. 91
[5] M. Samuel: Journal of Materials Processing Technology Vol. 105 (2000), p. 382
[6] R. H. Wagoner, M. Li: AIP conference proceedings Vol. 778 (2005), p. 209
[7] L. Chen, J. C. Yang, and L. W. Zhang: AIP conference proceedings Vol. 778 (2005), p. 125
[8] D. E. Green: AIP conference proceedings Vol. 778 (2005), p. 894
[9] J. Cao, M. C. Boyce: SAE 930517, p. 694

©(2014) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.619

Visual Simulation of Straight-racket-hit Skill of Table Tennis Based on MD2 Model

Tianqi Jia[1,a], Gaopin Guo[2,b], Yong Yu[1,c], Xijuan Guo[2,d]

[1]Institute of Physical Education, Yanshan University, Qinhuangdao, Hebei, China

[2]Institute of Information and Engineering, Yanshan University, Qinhuangdao, Hebei, China

[b]guogaopin@126.com

Keywords: straight-racket-hit; mechanical arm; MD2 model; scene simulation

Abstract. To the analysis and further development of straight-racket-hit skill of table tennis, visual simulation is established to help finding its character through comparing with straight racket backhand skill. To overcome such flaws as unable to exert enough force and difficult to determine a strike point, an optimal angle to achieve the skill is got through analyzing force of different angle of wrist and forearm. Through analyzing the visual simulation of mechanical arm, a systematic theory is established to orient the training of straight-racket-hit skill.

1. Introduction.

Straight-racket-hit skill is to deliver the attack using the back of the racket used by athletes of pen-holding style. It is a new table tennis skill developed by Chinese athletes in late 80s. When hitting the ball with this skill, it is more easily to tilt the racket forward to fully exert the strength of forearm. Therefore, for a penholder, it delivers more strength than a backhand strike. Moreover, comparing with the forehand strike, more technical actions can be adopted when using this skill. It is a new skill developed on the basis of traditional pen-holding style skills[1]. Resent years domestic athletes has greatly improved this skill and is in the lead in the world's table tennis circle. To maintain the lead, new skills must be developed and then enhanced to stand up to the competition. Thus straight-racket-hit skill will be adopted by more and more athletes.

2. Visual simulation of the mechanical arm.

The mechanical arm is the most widely used automatic machinery in the field of robot technology. Its application domains include industrial manufacture, medicine treatment, entertainment, military, space exploration and so on[2].

According to its structure, mechanical arm can be classified as articulated mechanical arm, Cartesian mechanical arm, spherical mechanical arm, polar mechanical arm and cylindrical mechanical arm[2].

In this paper, the technology of OpenGL is used to simulate the mechanical arm. First get the force and torque of each joint (especially finger joints) to calculate the movement state of the mechanism. Therefore the two-buffer technology of OpenGL is used to simulate the straight-racket-hit skill of the mechanical arm. To enhance the quality of animation, OpenGL provides two color buffers. When one of them is flushed to the screen, the other is calculating and preparing for the next flush. After a frame is drawn, the two buffers are exchanged, namely the buffer which is used to flush the screen is now used to calculate, and vice versa. This technology only flush the screen when the frame is finished, thus ensures that no uncompleted frame would be shown on the screen. The code to use two-buffer is: void glutSwapBuffers(void) [3]. The structure of the simulation system is shown in Figure 1.

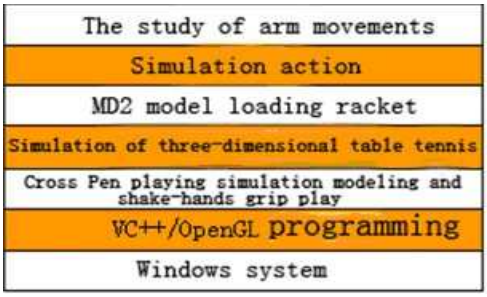


Fig. 1 Structure of visual simulation system of straight-racket-hit skill

3. Simulation of table tennis instrument

The instruments of table tennis include racket, ball, table and net[4]. Since it is inconvenient to establish complex models using OpenGL, the model of these instruments is established by MD2 modeling. It can be easily loaded by OpenGL, which is another reason to use MD2 model. MD2 is a file format used by 3D engine Quake2 to store models. It is simple and easy to realize, which makes it a widely used classic animation model file format. It is composed by two parts. One is the file head, which contains ID, version number and initial address of data of the model. The other is the main part of the file, which contains various data of the model, such as vertex data, texture data and vector data. MD2 is based on the key frame animation technology. The formula of the key frame interpolation is as follows[4]:

$$p(t) = p(0) + t\,(p1 - p0) \quad (1)$$

where t denotes the time, with 0 as the beginning and 1 as the end; p(t) denotes function of time; p0 denotes the start location; p1 denotes the end location.

Establishment of the model of instruments

There is a common way to establish a MD2 model, which can be described in the following two steps.

The origin model, which is not animation model but simply static model, can be easily established by 3D Max and then stored as 3DS file format.

Using MilkShape 3D1.7 to transform 3DS into MD2 file format[5]. MilkShape 3D is a modeling software developed by Mete Ciragan to establish models and edit actions of game models with less polygons. Models with different file format in various games such as Half Life and Quake can be easily edited by it. Furthermore, it can transform different 3D model files into other file format, including the transformation of 3DS into MD2.

4. Control of the table tennis instruments

The developing environment of the system is OpenGVS, a high level graphic developing system based on OpenGL. Function library of OpenGL is used to realize the loading function. The principle of MD2 modeling is to describe complex objects with triangles. Therefore, the vertex location and texture coordinate of each triangle is loaded from the file to display the graphs.

MD2 file is composed by five part: file head, set of texture coordinate, set of 3D object points, set of index of the constructing triangles and set of index of triangle texture coordinate[4]. The file head contains ID, version number, number of data of the model and relative address of the data to the initial address. The main part of the file contains data of the model, such as vertex data, texture data and normal vector data.

Vertex data in MD2 file can not be loaded directly, it must be decompressed firstly. To load a model, there are two steps: first, data structures must be defined to store the information coming from the MD2 files; second, after the structures are defined, the MD2 model is then loaded from the file. The instruments of table tennis is displayed and shown in Figure 2.

Fig. 2 The simulation ground of table tennis

5. Simulation and analysis of straight-racket-hit skill

5.1. Simulation of the holding gesture of straight-racket-hit skill

It is the holding gesture that makes straight-racket-hit skill special. Comparing to the commonly used gesture to pen-hold a racket, this skill demands a deeper and harder grip of the thumb to the handle, the forefinger moves towards the edge of the racket and loose the grip a little to make the racket tilt forward slightly to improve the adjustment of the racket. At the back of the racket lies the rest three fingers, they should stretch a little to improve the exertion of force and the stability of the racket[6]. The machinery arm is flexible and is capable to adjust each finger in small scale to get a correct holding gesture. The gesture is showed in Figure 3.

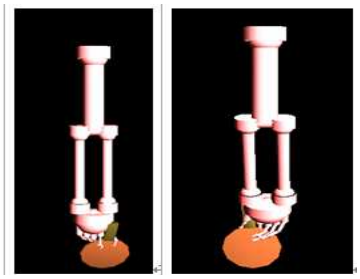


Fig. 3 Simulation of holding gesture of straight-racket-hit skill

5.2. Simulation of backhand half-volley of straight-racket-hit skill

Tilt the racket forward and hit the lower part of the ball when it is rising. Use the forearm to lead the wrist to exert the force forward and upward a little. The action is showed in Figure 4.

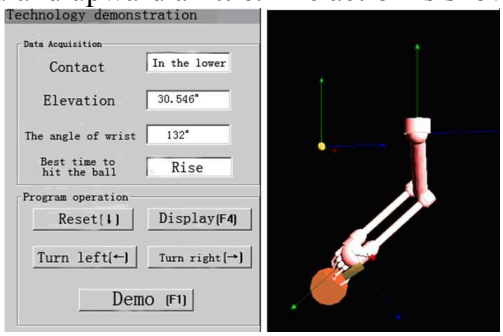


Fig. 4 Simulation of backhand half-volley of straight-racket-hit skill

5.3. Analysis of the simulation

To study the straight-racket-hit skill, a comparison with data of straight racket backhand skill is made. Based on character of table tennis skills, this paper makes the analysis through action simulation. The action to achieve loop ball is divided into five detail actions: preparing, wielding the racket, racket touching the ball, wielding racket with the ball on it, recovery. The analysis focuses on the action of wielding racket with the ball. Through analyzing the data get from the simulation, apart from velocity, the result shows that it has little different between actions of hitting different loop ball. Here data of straight-racket-hit skill and straight racket backhand skill is analyzed. The time, distance and velocity of the wielding are constant in the comparison.

5.3.1. Comparison of the time used to hit the ball between the two skills

Some definition will be given before further analysis.

The preparing action begins when racket is posed on the preparing location and end when the racket reaches the maximum location to the bottom right direction. Wielding racket action is to wield the racket from the maximum location to meet the ball. Wielding racket with the ball on it action begins with the touch of the ball and ends when the ball reaches the maximum location to the top left direction. Recovery returns the gesture to the preparing position[6]. Through analyzing the time between the two skills when the ball is in the middle of the rising period and when it begins to fall, we can see that less time is used by the straight-racket-hit skill when achieving the preparing action, the wielding racket action and wielding racket with the ball action. The comparison is list in Table 1.

Table 1. Comparison of time between the two skills [unit: second]

			Preparing	Wielding racket	Wielding racket with the ball	Recovery
	Holding gesture	Total time	Time	Time	Time	Time
Rise	Pen-holding	1.102	0.384	0.334	0.117	0.267
Medium-term	Straight-holding	1.000	0.334	0.150	0.084	0.434
Falling period	Pen-holding	1.102	0.401	0.317	0.150	0.234
Begins to fall	Straight-holding	1.069	0.284	0.234	0.117	0.434

5.3.2. Comparison between the distances of the wielding of the two skills

When playing table tennis, draw the arm back a little is helpful to improve the retractile power of the muscle, thus increases the speed of the ball when being hit. But it should be done appropriately, for researches shows that when muscle contracted too much, the retractile will decrease instead. The distance of the racket wielded from the preparing action to the wielding action influences the speed of the wielding. As the distance increases, the distance for the racket to accelerate increases as well. Time to hit the ball also influences the speed. The result is showed in Table 2.

Table 2. Comparison between distances of the two skills [unit: meter]

			Preparing	Wielding racket	Wielding racket with the ball	Recovery
	Holding gesture	Distance in total	Distance	Distance	Distance	Distance
Rise	Pen-holding	1.614	0.455	0.781	0.097	0.281
Medium-term	Straight-holding	1.585	0.332	0.628	0.153	0.471
Falling period	Pen-holding	1.530	0.494	0.712	0.120	0.203
Begins to fall	Straight-holding	2.199	0.401	0.968	0.200	0.630

From Table 2, we can see that the preparing distance of straight-racket-hit skill is longer than that of the straight racket backhand skill in both period of when the ball is in the middle of the rising period and when it begins to fall. When wielding the racket, it will generally accelerate to the maximum velocity and then the racket will hit the ball. The meeting of the ball and the racket is the key to guarantee a good hit. It will be better if they meet when the racket gets its highest speed, the hitting can get a bigger strength. As is showed in the formula $s = v0t + at/2$, when the initial velocity and time is the same, as the acceleration increases, velocity increases. And when the acceleration is the same, distance will increase with time. That is the way the racket gets its speed. Therefore a longer wielding distance means a faster racket speed. At the middle of the rising, the wielding distance of straight-racket-hit is greater than that of the straight racket backhand skill. When the ball begins to

fall, the wielding distance of straight-racket-hit is smaller than that of the straight racket backhand skill. At the middle of the rising, the wielding distance for straight-racket-hit is 0.153m greater.

5.3.3. Comparison between the arm angles of the racket of the two skills

Here the arm angle means the angle of the wrist and the forearm. Straight racket backhand skill hits the upper part of ball when it is rising, which makes the racket tilts forward a little. The forearm will meet the ball positively and exert the force forward and upward while the wrist moves forward to rub the ball[6]. It makes the straight racket backhand skill react more rapidly. Straight-racket-hit always uses the coming force to act, which demands the athlete to hit the ball quickly when it just bounced up. Namely act when the ball is rising, with the hand and elbow almost parallel to each other and the elbow moving near to the body. Based on the block action, the forearm exerts the force forward and rubs the ball to make it spin and increase its speed, as is showed in Table 3.

Table 3. Angles and best time to hit the ball of the two skills[unit:°]

Holding gesture	preparing	Wielding racket	Wielding racket with the ball	Best time to hit the ball
Pen-holding	Elbow parallel to wrist, wrist bend 50° and form an angle of 130° with the forearm, racket tilt forward	Racket tilt downward, wrist bend outward and form an angle of 142° with the forearm	Wrist rotate 10° and form an angle of 132° with the forearm, racket tilt upward	When the ball begins to rise
Straight-holding	Prepare the racket near the left and front of the belly, elbow form an angle of 35° with the wrist, and form an angle of 145° with the forearm, racket tilt downward	Rub the upper part of the ball, racket tilt forward, wrist and forearm form an angle of 150°	Forearm meet the coming ball positively, exert the force upward and forward, wrist rotate 17° forward, and form an angle of 163° with the forearm	When the ball rises

6. Conclusions.

Based on MD2 modeling, the simulation of straight-racket-hit is realized, and the analysis of the detail of each action is given. Then the result is compared to orient the training of the table tennis. Still there are many problems such as the simulation of the actions is not so fluent and the detail is not clear enough. The research must carry on to overcome such problems to orient the training, thus play an proper role in the development of table tennis.

REFERENCES

[1] X.D.Xu,X.M.Xu and S.G.Liu: *Table Tennis Players Hold Backhand of tactical training method*[J].Nanjing Institute of Physical Education.2004.

[2] G.M.Lu,L.N.Zhang and B.Zhang: *Five-DOF upper limb rehabilitation robot kinematics analysis*[J].Journal of Harbin Institute of Technology.2006.

[3] S.Dave,W.Mason,N.Jackie,D.Tom: OpenGL Programming Guide (Fourth Edition)[M]. Posts & Telecom Press.2005.

[4] Dove studio: OpenGL visualization of high-level programming and systems development • high-level programming papers (second edition)[M] . Water Resources and Hydropower Press. 2006.

[5] Dove studio: OpenGL advanced programming and visualization systems with open system development • Fat article (second edition)[M]. Water Resources and Hydropower Press.2006.

[6] X.G.Song,Y.G.Li: *Hold Backhand Table Tennis Technology Teaching*[J]. Chinese School Physical Education.2001.

doi:10.4028/www.scientific.net/AMM.50-51.624

Power Transformer Fault Diagnosis Based on Least Squares Support Vector Machine and Particle Swarm Optimization

Xin MA[1, a]

[1]School of Management and Economics, North China University of Water Conservancy and Electric Power, Zhengzhou, China, 450011

[a] maxin72@163.com l

Keywords: Least squares support vector machine (LS-SVM); Particle swarm optimization (PSO); Power transformer; Fault diagnosis; Dissolved gas analysis

Abstract. Dissolved gas analysis (DGA) is an important method to diagnose the fault of power t ransformer. Least squares support vector machine (LS-SVM) has excellent learning, classification ability and generalization ability, which use structural risk minimization instead of traditional empirical risk minimization based on large sample. LS-SVM is widely used in pattern recognition and function fitting. Kernel parameter selection is very important and decides the precision of power transformer fault diagnosis. In order to enhance fault diagnosis precision, a new fault diagnosis method is proposed by combining particle swarm optimization (PSO) and LS-SVM algorithm. It is presented to choose σ parameter of kernel function on dynamic, which enhances precision rate of fault diagnosis and efficiency. The experiments show that the algorithm can efficiently find the suitable kernel parameters which result in good classification purpose.

Introduction

Power transformer fault diagnosis is to determine the nature, location and severity of fault under the failure sign of power transformer. As the complexity of the power transformer fault, the diagnostic process can not only use a single method, but also through a combination approach. Therefore, the wide range of disciplines, methods and means must be conducive to fault diagnosis, which makes the fault diagnosis technology presents multidisciplinary characteristics.

Dissolved gas analysis (DGA) is considered as the most convenient and effective approach for power transformer fault diagnosis. However, there are many defects such as ‘short code’ and ‘encoding the border is too absolute’ in the three-ratio method or improved three-ratio method on the basis of data of DGA in the practice of transformer fault diagnosis. Recently, many artificial intelligence methods such as genetic programming [1], neural network [2], fuzzy theory [3] and expert systems [4] are applied to the diagnosis of transformer fault and achieved some results.

According to the mechanism of oil dissolved gas under the transformer failure, there is no clear function mapping between the gas content of oil and transformer fault types, the distribution of gas content is also very difficult to speculate, the accuracy and the number of actual field data collection are also very limited, therefore the traditional diagnostic classification based on experience risk minimization principle can not get enough knowledge to learn which resulting the lower classification accuracy and the higher diagnostic bias.

SVM has the outstanding classification ability in small samples and applied in fault diagnosis, but the kernel parameter selection is often not satisfactory. The parameter optimization method of least squares support vector machine is exhaustive search method in parameter space, that is within certain parameters, each parameter selected a value in accordance with certain intervals, each value is called a level, different levels of each parameter constitutes a combination of several sets of parameters, and a set of the minimum expected the upper bound of risk level is selected as optimal parameter values. In practical problems, it is difficult to know the range of optimal parameter in advance, so this method is not only difficult to accurately find the optimal parameters, but also has lower efficiency by way of the over election. PSO is a new evolutionary computation technique to find the optimal solution

through group collaboration and information sharing between individuals. This method not only fast convergence, but also easy to implement and there is no need to adjust many parameters. In this paper, the optimal value of kernel parameters are searched and seized based on PSO algorithm and combined with LS-SVM for power transformer fault diagnosis, in order to improve the performance of fault diagnosis system.

Parameter Selection of LS-SVM

Model Construction. Least squares support vector machine (LS-SVM) is an improved algorithm based on support vector machines, which uses quadratic loss function replace of not sensitive loss function in SVM, the loss function is constructed to turn the quadratic optimization into solving a linear equation, and this approach has the quickness solution. Objective optimization of LS-SVM can be described as follow.

$$\begin{cases} \min J(\omega,\xi)=\frac{1}{2}\omega^T\omega+\frac{1}{2}\theta\sum_{i=1}^{N}\xi_i^2 \\ s.t. \quad \phi_i=\omega^T\varphi(x_i)+b+\xi_i, i=1,\cdots,N \end{cases} \tag{1}$$

The Lagrange function is introduced in order to solve this optimal problem.

$$L(\omega,b,\xi,a)=\frac{1}{2}\omega^T\omega+\frac{1}{2}\theta\sum_{i=1}^{N}\xi_i^2-\sum_{i=1}^{N}a_i[\omega^T\varphi(x_i)+b+\xi_i-\phi_i] \tag{2}$$

$\varphi(x_i)$ is the mapping function of kernel space and maps the input data into the high demission space. ω is the weight vector. ξ is the error variable. θ is regularization parameter. b is deviation. a_i is Lagrange multiplier. The equation (2) is partial derivative for ω, ξ, θ, b and a_i respectively, and setted as 0, the following linear equations can be derived.

$$\begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & K(x_1,x_1)+\theta^{-1} & \cdots & K(x_1,x_N) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & K(x_N,x_1) & \cdots & K(x_N,x_N)+\theta^{-1} \end{bmatrix}\begin{bmatrix} b \\ a_1 \\ \vdots \\ a_N \end{bmatrix}=\begin{bmatrix} 0 \\ \phi_1 \\ \vdots \\ \phi_N \end{bmatrix} \tag{3}$$

Where,

$$K(x_i,x_j)=\varphi(x_i)^T\varphi(x_j),\ \text{i,j=1,…,N} \tag{4}$$

Kernel Function Selection. The LS-SVM model can be showed as:

$$\phi(x)=\sum_{i=1}^{N}a_iK(x,x_i)+b \tag{5}$$

The kernel function of LS-SVM includes radial basis function, polynomial function, sigmoid function, linear function and so on. The radial basis function is used as kernel function for its good generalization ability and its expression is:

$$K(x_i,x_j)=\exp(-\frac{\|x-x_i\|^2}{2\sigma^2}),\ \text{i=1,…,N} \tag{6}$$

The parameter selection of LS-SVM is the combination of kernel function parameter σ and regularization parameter θ.

Evaluation Indicators. For the parameters (σ, θ) which need to optimized, the sample standard deviation is selected as the evaluation indicator and also used as the objective function of PSO algorithm.

$$E_{SSD} = \left\{ \frac{1}{N} \sum_{j=1}^{1} [\phi(x_j) - \phi_j]^2 \right\}^{\frac{1}{2}}$$

$$= \left\{ \frac{1}{N} \sum_{j=1}^{N} [\sum_{i=1}^{N} a_i \exp(-\frac{\|x_j - x_i\|^2}{2\sigma^2}) + b - \phi_j]^2 \right\}^{\frac{1}{2}}, \text{i,j=1,…,N} \quad (7)$$

PSO Algorithm

Mathematical Description. There is a swarm consisted of m particles in a D dimension searching space, the velocity of particle is a vector: $v_i=[v_{i1}, v_{i2}, \ldots, v_{id}]$, i=1,2, …,m; d=1,2, …,D. The initial location of particle is a stochastic vector: $u_i=[u_{i1}, u_{i2}, \ldots, u_{id}]$, i=1,2, …,m; d=1,2, …,D. The best location p_{besti} searched by particle i is show as: $P_i=[p_{i1}, p_{i2}, \ldots, p_{id}]$, i=1,2, …,m; d=1,2, …,D. The best location g_{best} searched by particle i is show as: $P_g=[p_{g1}, p_{g2}, \ldots, p_{gd}]$, d=1,2, …,D.

The optimal value is searched by iteration according to the velocity renew equation (8) and the location renew equation (9).

$$v_{id}^{k+1} = wv_{id}^{k} + c_1 r_1 (p_{id} - u_{id}^{k}) + c_2 r_2 (p_{gd} - u_{id}^{k}) \quad (8)$$

$$u_{id}^{k+1} = u_{id}^{k} + v_{id} \quad (9)$$

v_{id}^{k}: the d dimension velocity of the i particle in the k iteration.
u_{id}^{k}: the d dimension current location of the i particle in the k iteration.
c_1,c_2: positive acceleration factor.
r_1,r_2: random number between [0,1].
w: inertia weight factor.

Inertia Weight Factor. The larger w the better globe optimal searching capacity and the local maxima will be avoided. The smaller w the better local optimal searching capacity and the higher convergence speed and accuracy.

$$w^k = w_L - \frac{(K_{\max} - k)(w_{ini} - w_L)}{K_{\max}} \quad (10)$$

K_{max}: the largest iteration time.
k: current iteration time.
w_{ini}: initial inertia weight factor.
w_L: inertia weight factor when reach the largest iteration time.

Parameter Optimization of LS-SVM/PSO Algorithm

The parameter optimization (kernel function parameter σ and regularization parameter θ) of LS-SVM used by PSO algorithm as following.

Step1. The initial parameters of PSO algorithm are set: the size of population m. positive acceleration factor c_1,c_2. Initial inertia weight factor w_{ini}. Inertia weight factor when reach the largest iteration time w_L. The initial velocity and location of particle.

Step 2. The fitness function of each particle is calculated and compared with the optimal objective value, if it is better than optimal objective value, then it is set as the new optimal objective value and iterated again.

Step 3. The inertia weight factor is calculated according to the equation (10), the velocity and location of particles are renewed according to the equation (8) and (9), the p_{best} and g_{best} are also be renewed.

Step 4. If the suspend conditions are satisfied, the iteration is end and the best classification accuracy is outputted. Otherwise go to step 2 and begin to the next iteration.

Experimental Methodology

Power Transformer Oil Dissolved Gas Generating Mechanism. Power transformer internally generated gases can be divided into the normal gases and fault gases. Normal gases are produced due to normal aging of transformer insulation system under normal operation, the concentration of hydrocarbon gases in the oil of normal transformers are show as table 1. Fault gases are produced by the thermal decomposition or discharge decomposition of insulation failure, the composition of gases under various power transformer fault are show as table 2, where the symbol '▲' means main components and '△' means minor components.

Table 1. Statistic limit concentration of hydrocarbon gases in the oil of normal transformers

hydrocarbon gases	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
Concentration [μL /L]	150	60	4	70	10

Table 2. The composition of gases under various power transformer fault

composition of gases	failure type					
	overheating		arc discharge		local discharge	
H_2	△	△	▲	▲	▲	▲
CH_4	▲	▲	△	△	△	▲
C_2H_6	▲	△				
C_2H_4	△	▲	△	△		
C_2H_2			▲	▲		
C_3H_8	△	△				
C_3H_6	△	▲				
CO		▲		▲		△
CO_2		▲		△		△

Simulation Framework and Result. The Characteristics gases usually used for power transformer fault diagnosis are five types: H_2, CH_4, C_2H_6, C_2H_4, C_2H_2, the data of gases are normalized in order to reduce the affections caused by value differences among different gases.

$$\hat{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{11}$$

$\hat{x}_i$: normalized data of gases.

x_i : the original concentration data of gases, i=1,2, ⋯,5.

x_{max}: the maximum value of original concentration data of the same gas.

x_{min} : the minimum value of original concentration data of the same gas.

The initial size of population m=30, positive acceleration factor $c_1=c_2=2$, Initial inertia weight factor $w_{ini}=0.9$, Inertia weight factor when reach the largest iteration time $w_L=0.38$. The faults of power transformer is divided into four types: low-energy discharge(D_1), high-energy discharge(D_2), high temperature overheating(T_3) and low temperature overheating(T_1), which high-energy discharge generally refers to arc discharge and more intense spark discharge, low-energy discharge generally refers to partial discharge and relatively weak spark. In order to verify the validity of this algorithm, the same input / output data are used to compare the proposed LS-SVM/PSO algorithm and

self-organizing polynomial neural networks (SOPN) [5], the diagnosis models are trained for the same number of training samples, diagnosis of the test samples as shown in table 3.

Table 3. Result of fault diagnosis

project		fault type			
		T_1	T_3	D_1	D_2
The size of test samples		60	28	23	41
SOPN	number of correct judgement	57	23	20	36
	recognition rate[%]	95	82.1	86.9	87.8
	total recognition rate[%]	89.50			
LS-SVM/ PSO	number of correct judgement	57	26	21	38
	recognition rate[%]	95	92.9	91.3	92.7
	total recognition rate[%]	92.98			

Summary

Considering the fact that parameters in support vector machine are usually decided by cross validation, LS-SVM/PSO algorithm was presented in which the parameters in SVM method are optimized by PSO algorithm. It was then applied to the fault diagnosis of power transformer based on dissolved gas analysis. The five characteristic gases dissolved in transformer oil are the inputs of support vector machine. The radial based kernel is selected to the model, and the superiority of SVM in processing finite samples is fully considered. The test shows the proposed method can find out the optimum accurately in a wide range and the value can be used to diagnose the transformer effectively.

References

[1]. Z. Zhang.: Application of Genetic Programming Based Discriminate Functions in the Insulation Fault Diagnosis of Power Transformers. Journal of Shanghai Jiaotong University, Vol. 40(2008), p.558-562.

[2] Z. Wang: An Artificial Neural Network Approach to Transformer Faults Diagnosis. High Voltage Research,Vol.13 (2009), p.1224-1229.

[3] Fu Y.: Power Transformer Fault Diagnosis Using Weighted Fuzzy Kernel Clustering. High Voltage Engineering, Vol. 36(2010), p.371-374.

[4] H. Xiong: Study on Kernel Based Possibility Clustering and Dissolved Gas Analysis for Fault Diagnosis of Power Transformer. Proceedings of the CSEE,Vol.25(2005), p.162-166.

[5] H.T. Yang: Intelligent Decision Support for Diagnosis of Incipient Transformers Faults using Self-organizing Polynomial Networks, IEEE Trans.on Power System, Vol. 13(1998), p.946-952.

Online: 2010-10-01 Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.629

Control of stochastic energetic resonance in a bistable system

Ying Meng, Min Lin

College of Metrology and Measurement Engineering, China Jiliang University
Hang zhou, China
E-mail: linm@cjlu.edu.cn

Keywords: stochastic energetic resonance, control, bistable system

Abstract. The motion of a Brownian particle in a bistable system is studied. We propose a method of controlling stochastic energetic resonance by adding a period signal to the system. The work done by the input signal is investigated as a quantifier of stochastic energetic resonance. By changing the control signal, the cooperative interplay between input signal, noise and nonlinearity of the system is adjusted effectively. Stochastic energetic resonance can be achieved or enhanced when the amplitude or frequency of the control signal is modified**.**

Introduction

The original work on stochastic resonance is carried out by Benzi et al. [1] while modeling the switching of the Earth's climate between ice ages and periods of relative warmth. SR defines the effect as amplification of a weak signal applied to the input of a non-linear system by tuning the noise intensity. In recent years, it has been studied extensively due to the potential application, such as for nonlinear memory storages, image reproduction and sensors to detect weak signal. SR, so defined, leaves a lot of liberty as to what is the observables that actually quantified the effect. These observables should be physically motivated, easily measurable, and be of technical relevance [1, 2-4]. The response amplitude (RA) was originally used to quantify SR by Benzi et al. The signal-to-noise ratio (SNR) is a measure of SR that can be observed in many systems. Besides, hysteresis loop area and area under the first peak in the residence time distribution have turned out to be good quantifiers characterizing SR as a bona fide resonance. Recently, Sekimoto [5] proposed a method of stochastic energetics (SE), which enable us to analyze the energetic relationships of SR described by Langevin equations. Subsequently, I.wai discovered the work done by the periodic force is a better qualifier of SR [6,7]. Later, Mamata Sahoo et al investigated heat fluctuations in a driven double-well.

The SR effect requires three basic ingredients：a weak input signal, a source of noise and a system with threshold. However, in realistic systems, the input signal and noise are fixed and the system is also determined. It is difficult to achieve SR when all the parameters are unchangeable. Gammaitoni et al [8] introduced a control scheme which allows us to either enhance or suppress the spectral response in the basic SR effect in 1999, then Locher et al [9] and other groups [10,11] proposed a variety method to control SR. The present work is based on the theory of SE and the theory of controlling SR. We add a periodic force to the SR model and propose the method of controlling stochastic energetic resonance. The paper is organized in the following way. In Section 2 and Section 3, the model of SR with a control signal and the work done by the external agent are explained. In section 4, we discuss the effect of the control signal to the system with different noise intensity. Finally, we present the summary of our results and discussions in Section 5.

Controlling stochastic energetic resonance with a period signal

The typical model of SR is given by

$$\frac{dx(t)}{dt} = -\frac{\partial U(x,t)}{\partial x} + \xi(t). \tag{1}$$

where $U(x) = -\frac{1}{2}\mu x^2 + \frac{1}{4}x^4 - xA\sin(\Omega t)$ is the reflection-symmetric quartic potential, μ is the potential parameter, $\xi(t)$ denotes a zero-mean Gaussian white noise with $\langle \xi(t)\xi(t') \rangle = 2D\delta(t-t')$, and intensity D. The model describes the overdamped motion of a Brownian particle in a bistable potential in the presence of noise and periodic force. The potential U with A=0 has bistable minima at $x = \pm\sqrt{\mu}$ and a maximum at $x = 0$.

We add a periodic signal $B\sin\Omega t$ to the system to control the stochastic energetic resonance artificially by changing the amplitude and frequency of the control signal, as shown in Fig.1. The corresponding dynamic equation is

$$\frac{dx(t)}{dt} = \mu x - x^3 + A\sin\omega t + B\sin\Omega t + \xi(t). \tag{2}$$

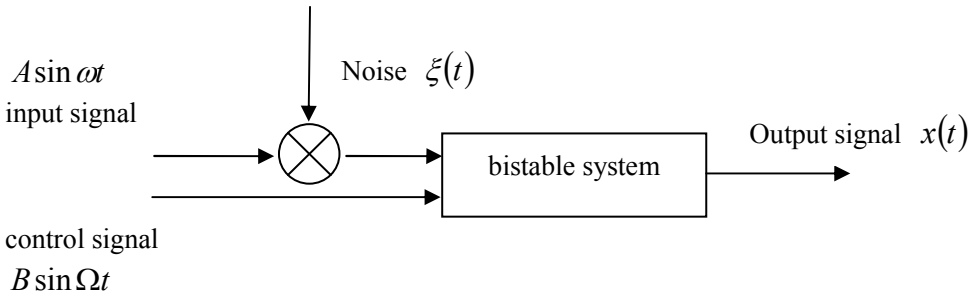


Figure.1 controlling stochastic energetic resonance with a periodic signal

Followed from the stochastic energetics formalism developed by Sekimoto, the work done by the external agent on the system is defined as

$$W = \int_{t_0}^{t_0+T_\omega} \left\langle \frac{\partial U(x,t)}{\partial t} \right\rangle dt = -\int_{t_0}^{t_0+T_\omega} A\omega \langle x(t)\cos\omega t \rangle + B\Omega \langle x(t)\cos\Omega t \rangle dt. \tag{3}$$

Where $T_\omega = \frac{2\pi}{\omega}$, t_0 is the initial transient time that is $2T$ in our simulation and angle brackets mean the ensemble average. For simplicity, we define W_1 as the work done by the input weak signal

$$W_1 = -A\omega \int_{t_0}^{t_0+T_\omega} \langle x(t)\cos\omega t \rangle dt. \tag{4}$$

By tuning the amplitude and frequency of the control signal, the work W_1 is analyzed. We set $\mu = 1$ throughout the paper. Numerical simulation of our model was carried out by using Heun's method. To calculate W_1 we first evolve the system and neglect initial transients. To get better statistics, we have calculated W_1 for 2000 different cycles .

Results and Discussions

In Fig.2 we plot the average work done by the input signal over a single period as a function of noise intensity D for $A = 0.3$, $\omega = 0.02\pi$ and $B = \Omega = 0$. The curve is resonance-like and the work W_1 takes a maximum value at $D = 0.33$ when the stochastic energetic resonance happens.

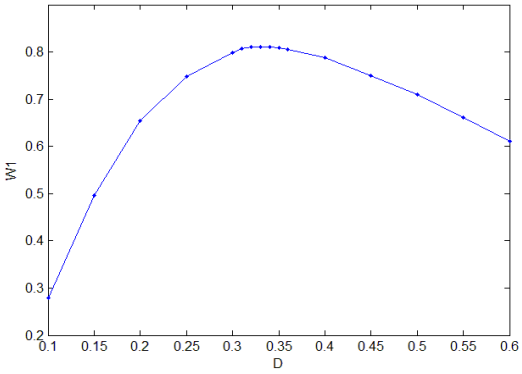


Figure.2 The average work W1 done per period with the noise intensity D

1) Small intensity of the noise

For the case where noise is small (D=0.25), the particle remains in one of the potential wells with an occasional random jump to another well and intrawell motion dominates. When introducing a control signal $B\sin\Omega t$ ($\Omega = 0.08\pi$) into the system, the work W_1 exhibits a peak as B varied. In Fig.3, the work W_1 increases with B increasing from 0 to 0.21 showing the enhancement of the stochastic energetic resonance. However, W_1 decreases for larger B and the stochastic energetic resonance weakens.

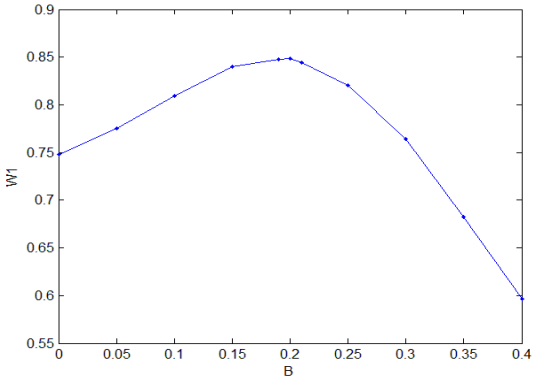


Figure.3 The work W1 versus the amplitude of the control signal B (D=0.25, $\Omega = 0.08\pi$)

If the frequency of the control signal is reduced to $\Omega = 0.04\pi$, the optimal amplitude of the control signal B is changed to 0.19. It is shown that different B is needed to achieve stochastic energetic resonance when the control frequency Ω is different.

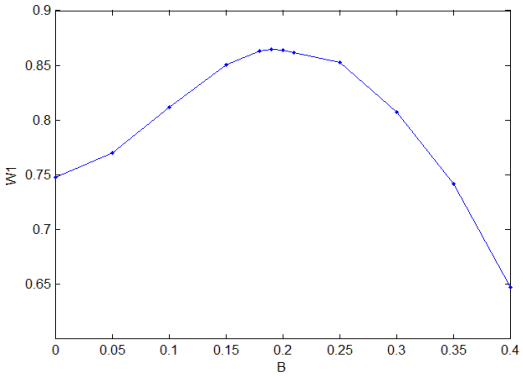


Figure.4 The work W1 versus the amplitude of the control signal B (D=0.25, $\Omega = 0.04\pi$)

2）Suitable intensity of the noise

In figure 2, the work W_1 takes a maximum value at D=0.33. When applying a control signal $B\sin\Omega t$ ($\Omega = 0.08\pi$) to the system, the effect of the stochastic energetic resonance can be enhanced significantly (see Fig.5). It is noteworthy that the work W_1 at B=0.11 is larger than that at B=0.

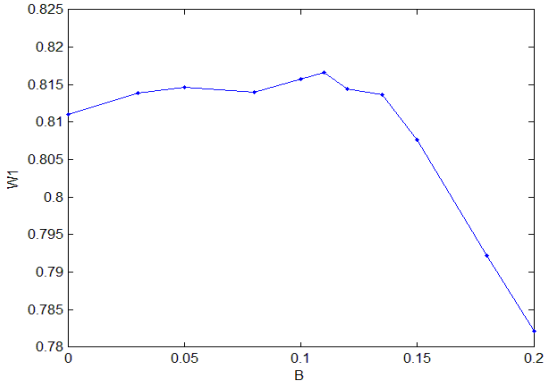


Figure.5 The work W1 versus the amplitude of the control signal B (D=0.33)

3）Large intensity of the noise

For the case when $D = 0.36$, a large number of interwell random switches occur for each period of the input signal and the system response is again random. If we add a control signal to the system with fixed frequency $\Omega = 0.08\pi$, the work W_1 when the intensity of the control signal B varied is shown in Fig.6. It is noted that the work W_1 at B=0.41 is larger than that at B=0. The result shows that the control signal leads the interwell transition more in order, thus enhancing the stochastic energetic resonance.

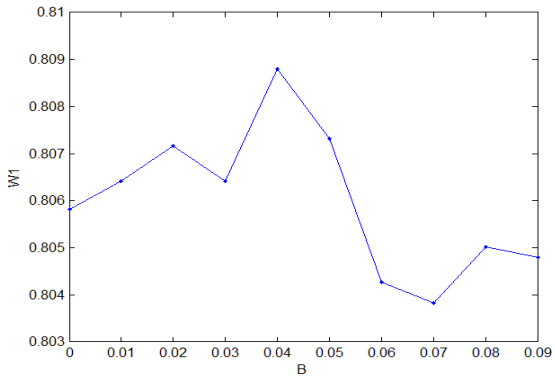


Figure.6 The work W_1 versus the amplitude of the control signal B (D=0.36)

In figure 7, we have plotted average work W_1 as a function of noise intensity D for different values of control signal intensity B for $A = 0.3$, $\omega = 0.02\pi$ and $\Omega = 0.08\pi$. As we increase B, the

maximum of the work W_1 increases and the peak position shifts towards higher values of D. It is clearly seen from the figure that the control signal enhances the stochastic energetic resonance effectively.

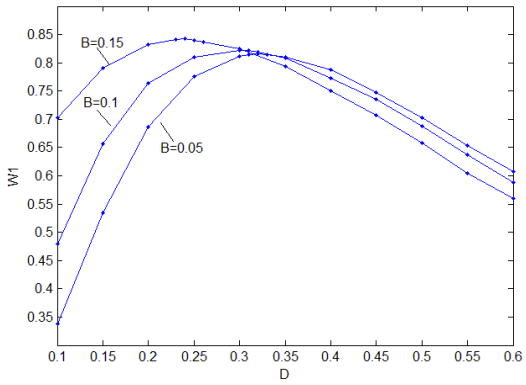


Figure 7. The work W_1 as a function of D for different values of the intensity of control signal (B).

Summary

In this paper , with the aid of stochastic energetics, we propose a method of controlling stochastic energetic resonance by introducing a periodic signal. Adding the system with a periodic signal can lead the fluctuations in the single well and thus affects the transition between the two wells. An effective control of stochastic energetic resonance can be achieved while changing the amplitude or frequency of the control signal.

References

[1] R. Benzi, A. Sutera, and A, Vulpiani. Phys. A, Mathematical and general, vol. 14(11) (1981), p. 453-457

[2] J.S.Lim, M.Y.Choi and B.J.Kim, Phys. Rev. B, vol. 68 (2003)

[3] S. Fauve and F. Heslot, Phys. Lett. vol. 97 (1983) p. 5-7

[4] L.Gammaitoni, F. Marehesoni, and S. Santucc, Phys.Rev.Lett, vol 74 (1995), p. 1052-1055

[5] K. Sekimoto, Prog. Theor. Phys. Supp, vol. 130 (1998) p. 17-23U.Seifert, Eur. Phys. J. B, vol. 64 (2008), p. 423-431

[6] T. Iwai, J. Phys. Soc. Jpn. vol 70 (2000), p. 353-358

[7] T. Iwai, Phys. A. vol 300 (2000), p. 350-358

[8] Gammaitoni L, Locher M, Bulsara, Hanggi P, Neff J, Wiesenfeld K, Ditto W L and Inchiosa M E, Phys. Rev. Lett.vol (82) (1999), p. 4574

[9] Locher M, Inchiosa M E, Neff J, Bulsara A R, Wiesenfeld K, Gammaitoni L, Hanggi P and Ditto W L. Phys. Rev. E. vol (62) (2000), p. 317

[10] M.Lin, Q.M.Mao, Y.J.Zheng and D.S.Li. Acta Phys. Sin. vol (56) (2007), p. 5021

[11] M.Lin, L.M.Fang and R.G.Zhu. Acta Phys. Sin. vol (57) (2008), p. 2638

Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.634

Robust Intrusion Detection Algorithm Based on K-means and BP

Yangjun Zhong[a], Shuiping Zhang[b]

Faculty of Information Engineering,Jiangxi University of Science and Technology, Ganzhou,341000,China

[a]zhongyangjun@mail.jxust.cn, [b]zsp@mail.jxust.cn

Keywords: Intrusion detection;k-means;BP;Abnormal behavior

Abstract: Nowadays the traditional intrusion detection models have two disadvantages:low work efficiency and high false positive.Considering these disadvantages,this paper proposes a new intrusion detection method combining k-means clustering algorithm and BP neural network.The experimental results show that the improved intrusion detection model in this paper can save the calculation time significantly with same detection capabilities for the abnormal behavior.

1.Introduction

Along with the rapid development of computer and internet, people have realized the importance of network safety problem[1].Intrusion detection system has become an essential part in the field of network safety and it is different from the traditional network safety protection technologies such as firewall and digital encryption. In 1987,Denning proposed the first intrusion detection model. From then on,intrusion detection technologies have developed from the pure theoretical research to commercial products. Beacause of the fast development,the intrusion detection technologies cann't been neglected in the future research of security protection.

K-means is the most widely used clustering analysis method which is very simple and rapid. This algorithm can customize the number of clustering and ensure high similarity in the same clustering and low similarity between different clustering[2].

In 1986,the group leaded by Rumelhart and McCelland proposed BP(Back Propagation) neural network.As the most popular neural network currently,it is a network of single transmission and multilayer feedforward[3].

This paper puts forward an intrusion detection model combining k-means with BP neural network.The model can be divided into two parts: learning units and work units.Learning units is also referred to data analysis units.In the phase of data analysis,k-means clustering algorithm is used for training data sets to produce clustering centers and then clustering centers are sent to BP neural network.The training data sets are reduced and so this way not only can increase the efficiency of training,but also can avoid the various errors in massive samples training.In the simulation experiments,we use KDDCUP99 data sets and gain good testing results.

The remainder of this paper is organized as follows.We first introduce the data pretreatment in Section 2.The BP neural network for intrusion detection is presented in Section 3.Then experimental results are given in Section 4.Section 5 concludes the paper.

2. Data Pretreatment

The preliminary data collection and processing are the foundation of a complete invasion examination system[4].This article uses the most widely used network packets as the original data

source.We can get the network packets in many ways.Sniffer and Ethereal are two excellent softwares to obtain the network packets.At present,the majority of theory studies use the Winpcap interface which is based on Windows platform.Although we can obtain the network packets by the interface,the original datas have different types.In order to extract comparable characteristics,we need to set a special data preprocessing unit for changing characters into numbers and normalizing.

Select initial centers of the k clusters $c_1,...,c_k$

When the cluster membership is changing,
..Generate a new partition
..For i=1,…,n
….Assign each data point p_i to its closest cluster centers c_j .
..end
..Compute new cluster centers.
..For j=1,…,k
…. $c_j = \frac{1}{n_j}\sum_{p_i \in c_j} p_i$; n_j =number of points in c_j .
..end
end

Fig.1 k-means algorithm

K-means clustering is a well known partitioning method which aims to partition datas into k clusters in which each sample belongs to the cluster with the nearest mean. Given a set of samples,where each sample is a d-dimensional vector, k-means clustering aims to partition the n observations into k sets (k < n). Fig.1 shows the detailed process of k-means algorithm.

The training samples obtained after preprocessing will be sent into the k-means clustering algorithm.The purpose why we use k-means clustering algorithm is to reduce the training samples of BP neural network. After the clustering,we can much simplify the training samples[5]. Note two questions here:

1.Data pretreatment should not blindly set all data to normalization which will cover distinguishing features of the original data.We should be selective on the normalized in data pretreatment.The object which is suitable for normalizing should be the data that have obvious difference numerical.

2.The samples which have few number can be used as training samples directly and do not need to cluster.Mandatory clustering plays an adverse effect.

Table 1 k-means clustering results

Type of behavior	Number of cluster training	Number of cluster centers
nmap imap	408	15
pod	264	5
guess_passwd	353	13
Smurf	803	23
teardrop	1000	7
Satan	2015	14
Back	3254	20
Ipsweep	1305	6

3. The BP Neural Network For Intrusion Detection

BP neural network performs a highly adaptive nonlinear decision function from training samples and is a supervised learning mode[6].In this paper,the network is trained with 41 dimensional samples

as input and the value of detection results as output.We construct three layers BP with 41,10 and 1 neurons in the input,hidden and output layer respectively.

The purpose of BP network used here is to improve the detection capability on unknown intrusion behavior in the detection model[7,8].Because the BP network can judge known behavior and unknown behavior,then we put the unknown behavior into quarantine area.Through regular filtration,unknown behavior determine and perfecting BP network training,we can ensure their own safety and productivity of the entire intrusion detection model.

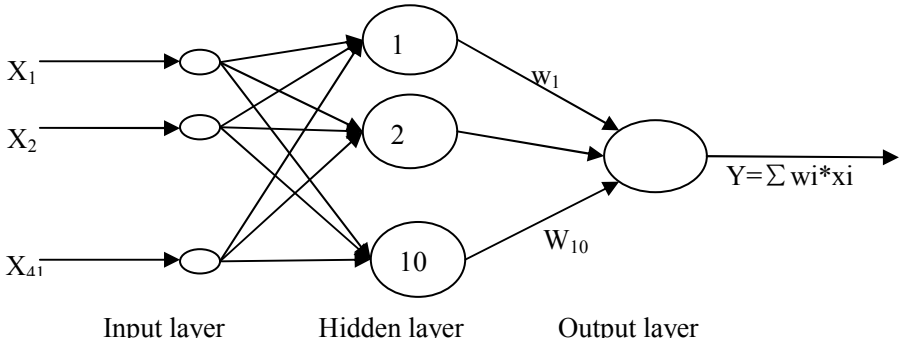


Fig.2The Back Propagation Neural Network model

Table 2 BP training statistics

Type of behavior	the number of BP samples	Learning results
nmap imap	15	00000001
pod	5	00000010
guess_passwd	13	00000100
Smurf	23	00001000
teardrop	7	00010000
Satan	14	00100000
Back	20	01000000
Ipsweep	6	10000000

The whole algorithm of this paper is outlined in Fig.3.

4.Experiment

In our experiments,we select KDDCUP99 data set which includes 494021 datas.Thereinto,there are 97278 normal datas and 396743 abnormal datas.The abnormal datas can be divided into four categories:dos,probe,u2l and r2l. In order to improve the accuracy of intrusion detection model, we classify each categorie further according to the specific types of invasion.

To test the detection capability of our method for unknown intrusion behavior,we choose some type of intrusion behavior randomly and we get two categories of data set needed by simulation experiments:the training data sets and the test data sets.

After preprocessing the datas uniformly, we design three schemes for simulation experiments:

1.We use k-means clustering algorithm for training data sets to produce clustering centers.The test data is compared with clustering centers and the distance of which less than prescriptive threshold is considered as abnormal behavior.

2.We use training data sets to train BP neural network directly.The test data will be sent to BP network and we can identify the test data through the output of the BP.

3.We use k-means clustering algorithm for training data sets to produce clustering centers and then send clustering centers to BP neural network. The BP network trained by the clustering centers is used to identify the test data.

Table 3 Experimental comparison table

Items / Experiment	Time of Training	Correct classification rate of test samples
Experiment 1	18 seconds	57%
Experiment 2	372 seconds	88%
Experiment 3	28 seconds	82%

In accordance with the experimental results, we come to the following conclusion:

1.The first model can't identify the unknown intrusion behavior and false positive is high.Moreover,the model doesn't have learning ability.

2.The recognition rates of the second model for unknown and known samples are both high, but the training samples are excessive and the time spent on training is too long.

3.Because the third model uses k-means clustering algorithm for training data sets firstly and only clustering centers are sent into the BP neural network for training,the coding and training spend very short time and the detection results on known and unknown samples between the third model and the second model are nearly the same.

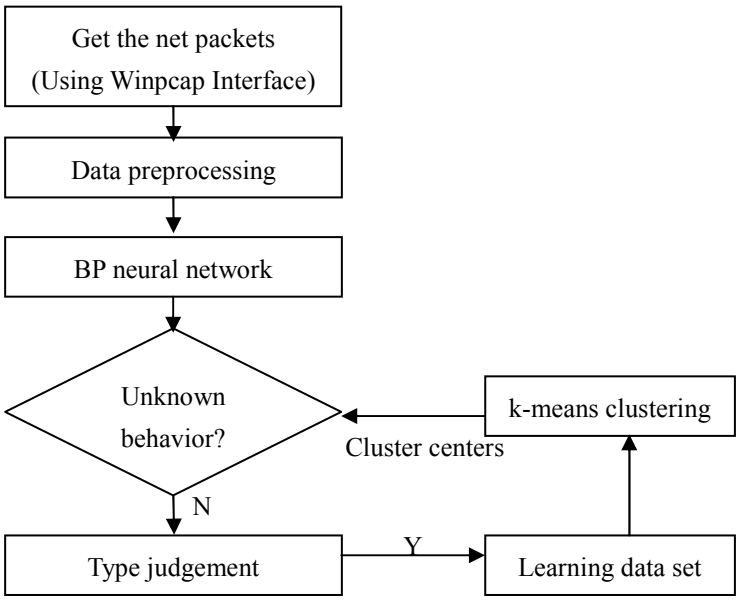


Fig.3 The flow chart of our algorithm

5.Conclusion

At present,high-speed internet communication has put forward rigorous requirements to the intrusion detection model. The intrusion detection model based on k-means clustering algorithm has great advantages for analysis of large-scale network packets.However, it is inacceptable for the detection capability to unknown intrusion behavior.As long as the prophase training of BP is sufficient,the detection speed can meet the requirement of the high-speed networks. Nevertheless,it also exists problems for the training of BP network to large-scale network packets.In this paper,we use k- means clustering algorithm for learning data sets and then the clustering centers produced by k-means will be provided for the training of BP neural network.The simulation experiments prove

that this method can effectively combine the advantages of both.What's more, the efficiency of intrusion detection model based on our method is greatly improved.

References

[1] Chih-Fong Tsai,Yu-Feng Hsu,Chia-Ying Lin and Wei-Yang Lin.Expert Systems with Applications.Vol.36(2009),p.11994.

[2] Safaa O.Al-Mamory,Hongli Zhang.Computer Communications.Vol.32 (2009),p.419.

[3] Rachid Beghdad.Computers & Security.Vol.27(2008),p.168.

[4] Neil C. Rowe,Sandra Schiavo.Computers & Education.Vol.31(1998),p.395.

[5] Kwok-Yan Lam,Lucas Hu and Siu-Leung Chung.Journal of Systems and Software.Vol.33 (1996),p.101.

[6] Uwe Aickelin,Julie Greensmith.Information Security Technical Report.Vol.12 (2007),p.218.

[7] Wun-Hwa Chen,Sheng-Hsun Hsu and Hwang-Pin Shen.Computers&Operations Research. Vol.32(2005),p.2617.

[8] Chia-Mei Chen, Ya-Lin Chen and Hsiao-Chung Lin.Computer Communications.Vol.33 (2010),p.477.

doi:10.4028/www.scientific.net/AMM.50-51.639

Divergence color histogram for Content-based Image Retrieval

Shi-xiong Xia[a], Hong-bing Zhou[b] and Yong Zhou[c]

School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, Jiangsu 221116

[a]xiasx@cumt.edu.cn, [b]zhou@cumt.edu.cn, [c]yzhou@cumt.edu.cn

Keywords: Divergence color histogram, Annular color histogram, Angular color histogram, Hybrid color histogram, Content-based Image Retrieval.

Abstract. Combined With the color and spatial partitions of image, this paper puts forward divergence to describe the dispersion information of pixels in each annular subset, and all those divergences of annular subsets constitute the divergence color histogram. It is invariant with translation and rotation, and suitable as a basic feature of image. Finally, simulation results show the improvement of the divergence color histogram over the spatial color histogram are: 16% (annular), 27% (angular)and 26% (hybrid), so the divergence color histogram has a significant better retrieval performance.

Introduction

In content-based image retrieval, color histogram is a very common statistical feature, it is efficiency, and not sensitive with translation and rotation. It can quickly retrieve many similar images from image database based on those advantages, but unfortunately in these images, the proportion of images with similar content is very small, that means the precision rate is low. The reason leading to these results is that the color histogram ignored spatial distribution of information treats different models of the images as a same model.

Combined with spatial information, Aibing Rao etc made a total of three kinds of *spatial color histogram*[1], that is the annular color histogram, the angular color histogram and the hybrid color histogram, these three histograms are invariance with rotation and translation, When the they are applied for image retrieval, the retrieval performance proven to be superior to the traditional color histogram.

In a small region, the dispersion of pixels is also very important information, so the paper will introduce another type of spatial color histogram named *divergence color histogram* based another hybrid partition.

The spatial color histogram[1]

Let $I=[I_{xy}]_{m\times n}$ be an image of size $m\times n$ where $I_{xy}, 0\le I_{xy}\le c$ is the color of pixel (x,y). Suppose $B_1, B_2, ..., B_M$ are disjoint sets is a partition of color Interval $[0,c]$ .Let $S_q=\{(x,y)\,|\,1\le x\le n, 1\le y\le n, c_{xy}\in B_q\}$

be a set of coordinates in which all colors of pixels are in color set B_q ,next consider the set S_q as geometric set on 2-D plane, compute the centroid denoted as $C^q=(x^q,y^q)$ of S_q and define x^q and y^q as

$$x^q=\frac{1}{|S_q|}\sum_{(x,y)\in S_q}x;\quad y^q=\frac{1}{|S_q|}\sum_{(x,y)\in S_q}y.$$

Next Compute the Maximum distance between the centroid and coordinate in set S_q ,Given an annular partition number H, S_q is divided into H annular subsets $R_{q1}, R_{q2}, \cdots, R_{qH}$ by H circles with

C^q as the center and with $\frac{hr^q}{N}$ as the radius for each $1 \le h \le H$ .This is illustrated in Figure 1, the annular partition is invariant with translation and rotation.

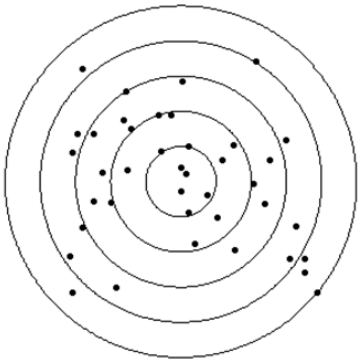


Figure 1. The annular partition of set S_q.

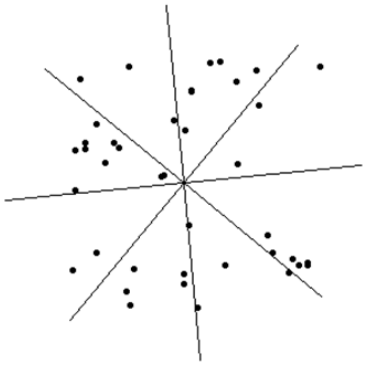


Figure 2. The angular partition of set S_q.

Let $A_{qh} = |R_{qh}|$, the matrix $A = [A_{qh}]_{M \times H}$ is called the annular color histogram of the image.

Similar to the annular partition, to get angular partition, we should compute the set S_q and the centroid C^q firstly, and then compute the S_q 's principal direction that is denoted as $\Theta(S_q) = \frac{1}{|S_q|} \sum_{(x,y) \in S_q} \theta(x,y)$ where $\theta(x,y)$ is defined as the direct angle from coordinate (x,y) to the centroid C^q, finally Given an angular partition number K, S_q is divided into K angular subsets $R_{q1}, R_{q2}, \cdots, R_{qK}$ by K radials with C^q as the origin and with $\Theta(S_q) + \frac{k}{N} 2\pi$ as the angle for each $1 \le k \le K$. This is illustrated in Figure 2, the angular partition is invariant with translation and rotation

Set $A_{qk} = |R_{qk}|$,the matrix $A = [A_{qk}]_{M \times K}$ is called the Improved annular color histogram of the image. By combining above two approaches, Given an annular partition number H and an angular partition number K, S_q is divided into $H \times \mathrm{K}$ hybrid subsets $R_{q11}, R_{q12}, \cdots, R_{qhk}, \cdots, R_{qHK}$ by H circles with C^q as the center and with $\frac{hr^q}{N}$ as the radius for each $1 \le h \le H$ and K radials with C^q as the origin and with $\Theta(S_q) + \frac{k}{N} 2\pi$ as the angle for each $1 \le k \le K$. This is illustrated in Figure 3, the hybrid partition is invariant with translation and rotation.

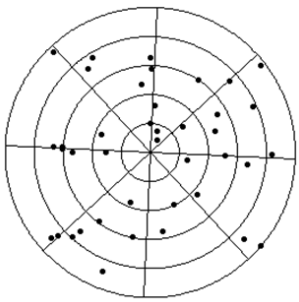


Figure 3. The hybrid partition of set S_q.

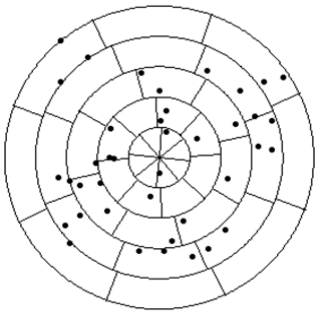


Figure 4.The another hybrid partition of set S_q.

Set $A_{qhk} = |R_{qhk}|$,the matrix $A = (A_{q11}, A_{q12}, \cdots, A_{qhk}, \cdots, A_{qHK})$ is called the hybrid histogram of the image.

The divergence color histogram

The paper will introduce another type of spatial color histogram named divergence color histogram based another hybrid partition.

Given an annular partition number H and an angular partition number K, firstly compute the set S_q and the centroid C^q for each $1 \le q \le M$, divide S_q into N subsets $R_{q1}, R_{q2}, \cdots, R_{qN}$ with the above annular partition method. Secondly, calculated R_{qh} 's principal direction that is denoted as $\Theta(R_{qh}) = \frac{1}{|R_{qh}|} \sum_{(x,y) \in R_{qh}} \theta(x,y)$ for each $1 \le h \le H$ where $\theta(x,y)$ is defined as the direct angle from coordinate (x,y) to the centroid C^q, R_{qh} is divided into K angular subsets $R_{qh1}, R_{qh2}, \cdots, R_{qhk}, \cdots, R_{qhK}$ by K radials with C^q as the origin and with $\Theta(R_{qh}) + \frac{k}{K} 2\pi$ as the angle for each $1 \le k \le K$. $R_{qh1}, R_{qh2}, \cdots, R_{qhk}, \cdots, R_{qhK}$ is the another hybrid partition and illustrated in Figure 4,the another hybrid partition is invariant with translation and rotation.

Set $A_{qh} = \frac{\sum_{k=1}^{K} \max_{(x,y) \in R_{qhk}} I_{xy}}{K \cdot \max_{(x,y) \in R_{qh}} I_{xy}}$,the matrix $A = [A_{qk}]_{M \times H}$ is called the divergence color histogram of the image and A_{qh} is called the divergence of set R_{qh} .If we don't consider the color information in B_q , A_{qh} is proportion of the nonempty set in $R_{qh1}, R_{qh2}, \cdots, R_{qhk}, \cdots, R_{qhK}$,reflects the divergence information of coordinates of set R_{qh} in the h's annular region. when A_{qh} is more bigger, the divergence is more bigger. In most cases, the variance is general used to reflect the dispersion of points, but because our set is distributed in the ring, the variance is not a good tool to describ their dispersion. The dimension of divergence color histogram is $M \times H$,and of the hybrid color histogram is $M \times H \times K$,so the dimension of divergence color histogram is far less than the hybrid color histogram. It makes the retrieval of divergence color histogram more efficiency than the hybrid color histogram with the same M, H and K.

Experiments and Comparisons

In the experiment, the cpu is Genuine Intel Cpu 2140@1.6GHz, the size of physical memory is 1.49GB, the softwares are Microsoft Windows Server 2003 and Matlab 6.5. This paper collected 500 images from the network as an image database, 100 images as the retrieval database. The Image database has 10 categories of images and contains pandas, mobile phones, snakes, mountain, people, guns, cars, horses, flowers, airplanes, and each category has 50 images. The retrieval database also contains the 10 categories of images, each including 10 images. In order to improve computational efficiency, the average for each image size after compression is about 100kb. The color space is RGB, and the detailed parameters are shown in Table 1.

Table 1. The color and spatial granularity for 4 histograms

M	40	40	40	
	R	G	B	Dimension
Ann(H)	25	25	25	3000
Ang(K)	25	25	25	3000
Hyb(HxK)	5x5	5x5	5x5	3000
Dive(HxK)	25x25	25x25	25x25	3000

In table 1, the color and spatial granularity for 4 histograms are listed where M is the division number of each dimension of the RGB color space, H is an annular partition number, K is an angular partition number.

Performance Comparison Figure 5 lists the average precision-recalls taken over 100 queries.

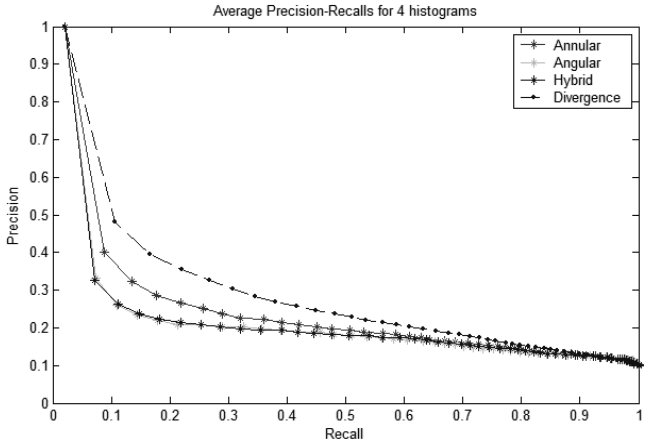


Figure 5. Average Precision-Recalls for 4 histograms

In figure 5, The Performance of the divergence color histogram is better than the others histograms where angular color histogram is very close to hybrid color histogram. In order to quantitatively determine differences in retrieval performance, the areas surrounded by the four histograms with the two axes are calculated and listed in Table 2.

Table 2. The areas surrounded by the four histograms with the two axes

	Area	Ratio
Ann.	0.2281	0.8354
Ang.	0.1994	0.7303
Hyb.	0.2001	0.7325
Dive.	0.2731	1

In table 2, The improvement of the divergence color histogram over the spatial color histogram are: 16% (annular), 27% (angular) and 26% (hybrid).

Time complexity of these four histograms can be divided by two parts, one is to partition the color and spatial region of the image, and the second is to calculate the feature value of all the partitions. so the annular color histogram is $O(m \cdot n + M \cdot H)$, the angular color histogram is $O(m \cdot n + M \cdot K)$,both the divergence and hybrid color histogram is $O(m \cdot n + M \cdot H \cdot K)$.In general, there is $M \cdot H \cdot K < m \cdot n$,It can be thought that the four histograms are the same time complexity $O(m \cdot n)$, so the divergence color histogram is little much inferior in efficiency.

Summary

In this paper, the partition of the divergence color histogram is invariant with translation and rotation, the divergence color histogram reflects the divergence information of coordinates in each ring, such information has a very good retrieval performance, and the divergence color histogram has the same time complexity with the three spatial color histograms.

Simulation experimental results show the improvement of the divergence color histogram over the spatial color histogram are: 16% (annular),27%(angular)and 26% (hybrid), so the divergence color histogram has a significant better retrieval performance.

References

[1] A.B. Rao, R. K, Srihari: *11th IEEE International Conference on Tools with AI* (1999), p.183

[2] S.A. Lim, G.J. Lu: *Information Technology: Computers and Communications* (2003), p.28

[3] G. Pass and R. Zabih: *IEEE Workshop on Applications of Computer Vision* (1996), p.96

[4] X.J. Zhang, J.G. Sun and Z.Y. Zhang: submitted to *Journal of Ningbo University* (2007)

[5] M. J. Swain and D. H. Ballard: submitted to *International Journal of Computer Vision* (1991)

doi:10.4028/www.scientific.net/AMM.50-51.644

A kind of Web Database Classification Based on Machine Learning

Xiaoqing ZHOU [1, a], Xiaoping TANG [1, b]

[1]Center of Computing, China West Normal University, Nanchong, Sichuan, China）

[a]zhousun123@163.com [b]369656820@qq.com

Key words: deep web; web database; feature extraction; feature valuation; Naive Bayes classifier

Abstract: The traditional search engine is unable to correct search for the magnanimous information in Deep Web hides. The Web database's classification is the key step which integrates with the Web database classification and retrieves. This article has proposed one kind of classification based on machine learning's web database. The experiment has indicated that after this taxonomic approach undergoes few sample training, it can achieve the very good classified effect, and along with training sample's increase, this classifier's performance maintains stable and the rate of accuracy and the recalling rate fluctuate in the very small scope.

Introduction

In order to use the information including in Deep Web database more effectively, it is need to categorize by its fields. At present the classification of Web database is primary and needs further improvement [1]. The paper advances the classification of web database based on machine learning and uses a new method to calculate weight. The classification method can work very effectively after trained with some samples tested by experiment; it will work better with more samples. And the stable performance can be available, accuracy and recall will vary within small range accordingly.

Classification of Web database

Feature extraction of the query interface

Feature extraction is the preface to categorize automatically. HTML form contains complicated information and also magnanimous useful information. How to choose proper feature to describe HTML form is very urgent. The pattern of HTML form is <form action="URL"method="get|post">, and the content of HTML form is </form>. The argument of Action means the processing program to submit HTML data to server. The argument of Method means the method to transfer the content of HTML form to sever [2].

HTML includes so much useful information that it can show its features by using all the control. But some property of outlook can be emitted, such as: size, length, and so on. Additionally the features can be made by the show times of control; HTML can be in one textbox or several textboxes.

According to the above-mentioned, the following features of HTML will be withdrawn:

1) The property of name in HTML <form> tag;
2) These words which withdrawn from the property of action in <form> tag;
3) The type of control which included in the HTML form;
4) The value of name property and value property in input control;
5) The property of name in select control and textarea control;
6) These words between control tags

Many features of HTML have been gotten after the above process. Then we will standardize the features to make it more accurate. Examples as the following figure1 and Figure2. Figure1 is for the code of HTML and figure2 is the features withdraw automatically.

<Form name=”searchadvformnew” method=”get”
Action=”http://search.book.dangdang.com/search.aspx”>
<input type=”text” name=”title”>
<input type=”text” name=”author”>
<input type=”text” name=”publisher”>
<input type=”image” name=”search_btn”>
</form>

Fig. 1 The HTML code for web form

InputType=text
InputType=image
Textname=title
Textname=author
Textname=publisher
Imagename=search_btn
Formname=searchadvformnew
Action=http
Action=search.book.dangdang.com
Action=search.aspx

Fig. 2 Features extracted from the web form automatically

Naïve Bayes Classifier

Classifier of Naïve Bayes is the most common of Probability-based classifier instructors [3]. The classifier parameters are formed by the priori class probabilities $P(c_j)$ and class-based entry of conditional probability $P(w_t \mid c_j)$, determined entirely by the training set of documents have been marked。The formula of each class c_j of priori class probability is such as the equation (1):

$$P(c_j) = \frac{1 + \sum_{d_i \in D} P(c_j \mid d_i)}{|C| + |D|} \qquad (1)$$

In equation (1), $|C|$ is the number for the class; $|D|$ is the text number of training in the collection. $P(w_t \mid c_j)$ is estimated by the equation (2):

$$P(w_t \mid c_j) = \frac{1 + \sum_{d_{i \in D}} N(w_t, d_i) P(c_j \mid d_i)}{|V| + \sum_{k=1}^{|V|} \sum_{d_{i \in D}} N(w_k, d_i) P(c_j \mid d_i)} \qquad (2)$$

In equation (2), $N(w_j \mid d_i)$ are frequencies of features appear in the text d_i; $|V|$ are on behalf of numbers of all the different features of term in text collection. For the text d_i, when d_i is defined as belonging to category c_j, $P(c_j \mid d_i)$=1, otherwise, $P(c_j \mid d_i)$=0.

For the text without annotations in test text set, we can be obtained the posterior probability $P(c_j \mid d_i)$ by using the classifier has been trained. Characteristics of the first k-term in the text d_i are expressed with $w_{d,k}$.Formula as the equation (3):

$$P(c_j \mid d) \propto P(c_j)\prod_{k=1}^{|d|} P(w_{d,k} \mid c_j) \quad (3)$$

The weight estimates of inquiring interface feature

It is need to create a model after getting the features of query interface. And it is also need evaluate the weight of the features by the method of statistics and mathematics in order to improve the classification accuracy. The famous weight function is the TF-IDF formula by Salton in 1988[4].

Recently, some scholars represent with the rationality of the TF-IDF to weight the features [5]. They think that the useful word sheet is only small ratio and most word sheet is not relative to the type which needs to confirm, so they are noise word sheet. The noise may cover useful information and induce the classification accuracy. So we use a new method to adjust weight, which is to evaluate every feature with evaluation function. The new weight function is called TF-TEF function, TEF is for the evaluation function of the features. The TF-TEF function is as follows:

$$w_{ik} = TF - TEF(t_{ik}) = TF(t_{ik}) \times TEF(t_{ik}) \quad (4)$$

Among them, $TF(t_{ik})$ are the frequencies of first k-term feature in the text d_i, $TEF(t_{ik})$ is the evaluation function to rate various characteristics of term and reflect the degree of correlation between various types of entry. This article uses information gain as evaluation function (equation (5)). $P(w)$ is the probability of Feature words w; $P(\overline{w})$ is probability of feature words w not appear. $P(c_t \mid w)$ means probability of feature words w to belong to the class c_t.

$$TEF_{\inf oGain}(w) = P(w)\sum_t P(c_t \mid w)\log\frac{P(c_t \mid w)}{P(c_t)} + P(\overline{w})\sum_t P(c_t \mid \overline{w})\log\frac{P(c_t \mid \overline{w})}{P(c_t)} \quad (5)$$

Weight adjustment techniques are on the basis of feature adjustment to evaluation function, and not Simple feature selection. During weight-adjusted, the classifier role is modified the characteristics. The Calculation of $P(c_j \mid d)$ changes as the equation (6):

$$P(c_j \mid d) \propto P(c_j)\prod_{k=1}^{|d|} P(t_{d,k} \mid c_j)^{TE-TEF(t_{j,k})} \quad (6)$$

In equation (6), $TF-TEF(t_{d,k})$ is the new weighting function of feature entry $t_{d,k}$. The higher weight means greater role in Naïve Bayes model, and the lower weight means smaller role .As $TF-TEF(t_{d,k})=0$, $P(w_{d,k}\mid c_j)$ ceases actually.

Experimental Design

This article has selected 195 database interfaces from three areas under UIUC Web in order to verify the feasibility of classification algorithm of Web database. We conduct manual classification of these data sources by selecting randomly 10,20,30,40,50,60,70 deep Web query interfaces as training samples, the rest are test samples.

Index of Web Database Categories can adopt precision, recall and F value. We suppose Web database of dw_1, $dw_2 \ldots dw_n$,and predetermined categories of l_1, $l_2 \ldots l_n$. TP assigned the dw_i to the correct category l_i; FN means not to assign the dw_i to the correct category l_i (Leakage points). FP Incorrectly assigned the dw_i to the l_i (False alarm), Then:

$$P=\frac{TP}{TP+FP},\ R=\frac{TP}{TP+FN},\ F_1=\frac{2PR}{P+R} \qquad (7)$$

The results shown in Figure 3.Firstly,the application of classification achieve accuracy of 89.4% and recall of 88.2% after training random sample of 30(Random 10 for each area).Secondly, the classifier performance is stable fundamentally, and the accuracy, recall and F_1 value wave in the small range with the increase of training sample.

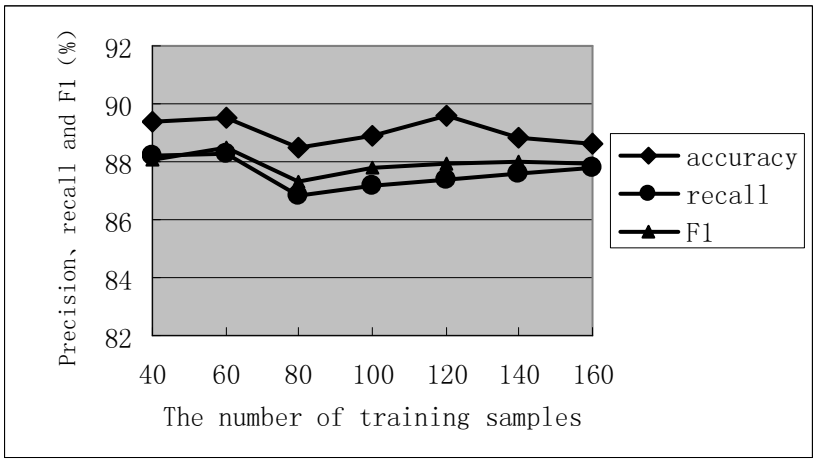


Fig. 3 Accuracy, recall and F_1

Conclusions

Web data includes query interface and result. The classification of web data is the key to integrate classification and search. The paper discusses feature extraction from HTML and classification with naïve Bayes. The result shows that the method used in the paper can work well

with small samples. But the paper does not consider more content of the words, so it is our further study.

References

[1] GAO Ling, ZHAO Peng-peng, CUI Zhi-ming. Automatic Judgment of Deep Web Query Interfaces. *Computer Technology and Development* Vol.17-5(2007), p.150-151

[2] YU Wei, LI Shi-jun, WEN Li-juan, TIAN Jian-wei.Ranking of Deep Web Sources Dased on Data Quality. *Journal of Chinese Computer Systems* Vol.31-4(2010), p.641-646

[3] JIN Ling-zhi, WANG Xiao-ling, ZHU Shou-zhong.Automatic classification of Deep Web sources. *Microcomputer Information* Vol.25-4(2009), p.227-230

[4] Tang Huanling, Sun Jiantao, Lu Yuch. A Weight Adjustment Technique with Feature Weight Function Named TEF-WA in Text Categorization. *Journal of Computer Research and Development* Vol.42-1(2005), p.47-51

[5] Wei Huang, Zhongliang Jing.Multi-focus image fusion using pulse coupled neural network.*Pattern Recognition Letters* Vol.28-9(2007), p.1123-1132.

doi:10.4028/www.scientific.net/AMM.50-51.649

Water Exit Dynamic Analysis of Underwater Vehicle

Tianqing You[a], Jiazhong Zhang[b], Cong Wang[c]

School of Astronautics, Harbin Institute of Technology, Harbin, 150001, China

[a] ytianqing@gmail.com, [b] zhangjz@hit.edu.cn, [c] alanwang@hit.edu.cn

Keywords: water exit, Fluid Structure Interaction, structural dynamic, Timoshenko beam, cavity

Abstract. The cavity, formed at the nose part of underwater launched vehicle due to the high launching velocity, will collapse while crossing water free surface and has a great impact on the strength of vehicle structure. In order to study the effect of the cavity collapse to vehicle structure, water exit dynamic analysis has been conducted. Before that, the computational model has been established. The external load, caused by cavity collapse, is assumed to be concentrated pulse forces applied on the vehicle at different time. And the flight vehicle structure is simplified into a free-free Timoshenko beam. The Fluid Structure Interaction (FSI) effect is treated as added mass and hydrodynamic drag force varying with wet surface area, which decreases with the water exit of vehicle. The dynamic response, excited by the external load, is calculated. Result presents the necessary of taking account of FSI and the influence of time space between the two concentrated forces.

1. Introduction

Because of the high velocity of underwater launched vehicle such as submarine-launched missile, the drop of the fluid pressure to its vapor pressure will give rise to cavity at the nose part [1]. While the vehicle crossing water free surface, severe and complexly formed load will apply on the vehicle structure suddenly. The high pressure, caused by collapse of the cavity [2-4], is one of major load that leads to strong vibration.

Besides the serious external load, the effect of adjacent fluid on the vehicle vibration can not been neglected [5-7]. At low frequencies, the fluid appears to the vehicle structure like an added mass[8]. Moreover, the area of wet surface decreases as the vehicle moves out of water. The effect of surrounding fluid on the vehicle vibration changes accordingly. For the vehicle, the dynamic property is time-dependent. The vehicle during water-exit process can be a time-varying Fluid Structure Interaction (FSI) dynamic system.

This paper deals with dynamic analysis of a slender underwater launched vehicle crossing water free surface. The external load caused mainly by cavity collapse is assumed to be two concentrated forces, applied on the vehicle at different time. And the flight vehicle is simplified into a free-free Timoshenko beam [9]. The FSI effect is treated as added mass varying with wet surface area. Dynamic response of the vehicle subjected to external load has been calculated. The necessary of taking account of FSI effect during water exit has been proved. The transient calculation of different time space between the two concentrated forces has been conducted. The impact of cavity collapse on the strength of vehicle structure has been analyzed.

2. Computational model

When underwater launched vehicle moving in the water at high velocity, the fluid pressure dropping to vaporization pressure at the nose part will give rise to cavity. Besides water vapor, the gas originally in launching tube can be sucked in the cavity. When the vehicle crosses water free surface, severe and complexly formed load will apply on the vehicle structure suddenly. The Computational Fluid Dynamics (CFD) has been conducted to analyze hydrodynamic load formed during water exit. As Fig. 1 shows, the free surface and cavity deformation has been obtained. Moreover, the pressure

contour has been present. As the result indicate, when the vehicle crosses the free surface, the low pressure cavity bubble will collapse. The pressure gradient makes the water near the low pressure bubble flow into the original cavity at high speed. The water flow impacts the hull of vehicle severely, causing sudden pressure rise up at relatively small location. Therefore, the force trigged by cavity collapse is assumed to be mainly concentrated force F_1 and F_2.

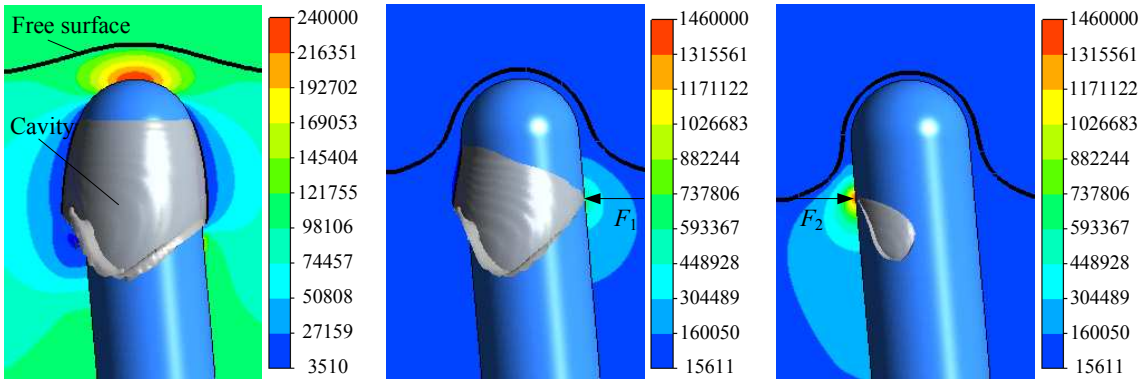


Fig. 1 free surface, cavity profile, and pressure distribution / Pa

Based on the CFD analysis, the water exit dynamic model has been established. As Fig. 2 shows, the cylinder shaped vehicle has the length of L=12.0m and diameter of D=1.60m. Compartment structure is widely adopted in underwater lunched vehicle. At the joint part between two chambers, there is 31% and 17% decrement of mass density and stiffness at the location of xD^{-1}= 1.24 and xD^{-1}= 3.79. And mass and stiffness distribution at the other position are assume to be uniform along the vehicle. The concentrated pulse forces F_1 and F_2 , formed by cavity collapse, apply on vehicle vertical to the axis at different time. And there is time space Δt between the two pulse forces.

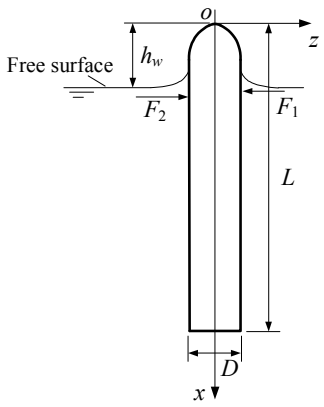


Fig. 2 Configuration of computational model.

When the vehicle moves underwater or crossing water free surface, the flow field is very complex. The fluid motion trigged by structure vibration is just a small part of flow field, and assumed to be linear small disturbance. Therefore, the water is treated as incompressible inviscid ideal fluid. Discretize the governing equations for water, and substitute into the structure dynamic equation. The structure dynamic equation considering FSI can be simplified into the following form:

$$\left(\mathbf{M}+\mathbf{M}_a\right)\ddot{\mathbf{u}}+\mathbf{C}\dot{\mathbf{u}}+\mathbf{K}\mathbf{u}=\mathbf{F} \tag{1}$$

Compared with the vibration equation in vacuum, the structure dynamic equation considering FSI has an additional matrix $\mathbf{M}_a$, which is added mass matrix.

The underwater launched vehicle is cylinder shaped, thus the three-dimensional added-mass can be approximated by a strip theory synthesis, in which the flow at each section is assumed to be locally two dimensional. For the circular section, the added mass of a two-dimensional circular cylinder is precisely equal to the displaced mass of fluid [10]. Convert the added mass into lumped mass matrix

form, the two nodes element below the water free surface, which needs to take the FSI into account, has the added mass matrix expressed as follows:

$$m_{ij} = \pi\rho_f R^2 L_e \delta_{ij} / 2 \tag{2}$$

Where, m_{ij} is the element of added mass matrix, R is the radius of local circular section, L_e is the length of beam element, the subscript i and j is the degree of freedom on the orientation of deflection at each node that need to consider the effect of added mass.

In order to verify the numerical method mentioned above, the frequencies of free-free uniform beam has been calculated. The uniform beam has the following properties of elastic modulus E=2.1 $\times 10^{10}$Pa, mass pre unit length m=3.24$\times 10^{4}$kg·m^{-1}, moment of inertia of cross-section I_z=5.97$\times$ 10^{-2}m^4 and length L=12.0m. As table 1 show, the present numerical result and the result calculated by ANSYS agree well.

Table 1. The comparison between the present result and ANSYS result.

Mode	In vacuum (Hz)		Fully submerged (Hz)	
	The present result	ANSYS result	The present result	ANSYS result
1	10.69	10.45	9.32	9.10
2	27.62	27.28	24.06	23.78
3	49.59	50.24	43.21	43.79

As the vehicle move out off the water free surface, the added mass matrix and hydrodynamic darg force is updated at each time step. As a result, the property such as mode shape and natural frequencies are change accordingly. For transient calculation, the method of mode superposition, based on fourth order Runge-Kutta algorithm, is adopted to calculate the dynamic response.

3. Transient analysis

The sudden concentrated pulse forces F_1 and F_2, with the magnitude of 50.0kN, is applied both at the location of xD^{-1}=1.05. The initial nondimensional water exit height is $h_{w0}L^{-1}$=1.00. And, the vehicle moves out of water at the velocity of v=30ms^{-1}.

The transient caculation has been carried out. The internal moment and shearing force response has been obtained. To present the necessary of taking account of FSI, the comparison of moment distribution along the vehicle is shown in Fig. 3. As the result shows, the moment response presents distinct difference between the result with FSI and without FSI.

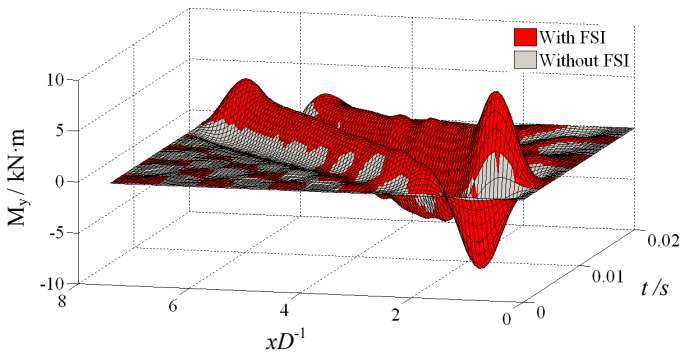


Fig. 3 Configuration of moment distribution.

Generally, the internal force responce level should be relieved, if FSI is taken into account. Bacuse the wet surface is constantly changing and the external load is applied at the cavity zone, the kind of FSI presents the unique phenomena. The numerical result indicates that the water exit can modify the structure property and has the tendency of deteriorating mechanical environment.

Although the application of the sudden impulse load, caused by cavity collapse, lasts very short time, and there is no need to consider resonance, the time space between the two concentrated impulse forces can have influence on the response level.

In order to analyze the effect of time space Δt, the structure dynamic simulation of the vehicle with different concentrated force locations $x_{F1}D^{-1}$ and $x_{F2}D^{-1}$ have been conducted. The concentrated force F_1 is applied at the location of $x_{F1}D^{-1}$=2.10, The concentrated force F_2 is applied differently at the location of $x_{F2}D^{-1}$=2.10, 2.25 and 2.40. At each loaction, the transient caculation of different time space Δt has been made. The curve of max moment along the vehicle *vs* time space Δt is shown in Fig. 4. As the increase of time space Δt, the max moment magnitude climb up. At the point of Δt=0.0025s, the magnitude reaches its peak. After that, the max moment drops and trends to be stable at relatively small value.

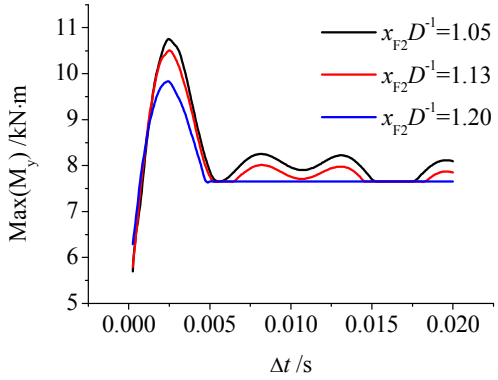


Fig. 4 Curves of max moment along the vehicle *vs* time space

The system is assumed to be linear, and superposition is satisfied. This phenomenon indicates that, at certain time space Δt, the response of the two forces can reinforce each other and cause the internal moment grow larger than usual. The same phenomenon happens to the shearing force response, and the max shearing force response along the vehicle *vs* time space Δt has been drawn, as Fig. 5 shows.

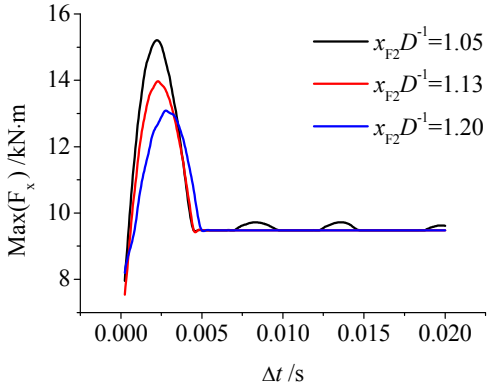


Fig. 5 Curves of max shearing force along the vehicle *vs* time space

The load casued by cavity collpase is assumed to be two concentrated forces applied on different time. The simulation result indicates that the time space between the two force has a significant influence on the internal force responce level. And the time space is determined by complex factors

during cavity collpase, such as head shape, piching angle, velocity and etc. The randomicity also makes this time space hard to control. Thus, in order to increase the vehicle structure reliability, it is critical to determine the max internal load.

4. Conclusion

Based on a simplified free-free Timoshenko beam model, the structure dynamic analysis of underwater launched vehicle crossing water free surface, including FSI effect, is conducted. The FSI effect is treated as added mass and hydrodynamic darg force varying with wet surface area, which decreases as the vehicle crossing free surface, has been introduced. The load casued by cavity collpase is assumed to be two concentrated impulse forces applied on different time.

The dynamic response subjected to external load, which mainly is caused by the collapse of cavity, is calculated. There is an apparent difference between the FSI and non-FSI result, which presents the necessary of taking account of FSI.

Because of the influence on dynamic response, the time space between the two force is found to be an important condition. Certain time space can cause the vehicle response larger than usual. Due to the uncontrollability of this time space value, it is nessary to to determine the max internal load to increase the structure reliability.

References

[1] R.T. Knapp, J.W. Daily and F.G. Hammitt: *Cavitation* (Mcgraw-Hill book company, 1970).

[2] Christopher Earls Brennen: *Cavitation and Bubble Dynamics* (Oxford University Press, 1995).

[3] Luo Jing and Li Jian: Two-dimensional simulation of the collapse of vapor bubbles near a wall. Journal of Fluid Engineering. Vol. 130 (2008), 091301

[4] Quan Xiaobo, Li Yan, Wei Haipeng and Lv Haibo: Cavitation collapse characteristic research in the out-of-water progress of underwater vehicles. Journal of Ship Mechanics. Vol. 12 No. 4 (2008), p. 545-549

[5] C. Conca, A. Osses and J. Planchard: Added mass and damping in fluid-structure interaction. Comput. Methods Appl. Mech. Engrg. Vol. 146 (1997), p. 387-405

[6] Ray P.S. Han and Hanzhong Xu: A simple and accurate added mass model for hydrodynamic fluid-structure interaction analysis. J. Franklin Inst. Vol.333B, No. 6 (1996), p. 929-945

[7] Tomomi Uchiyama: Numerical prediction of added mass and damping for a cylinder oscillation in confined incompressible gas-liquid two-phase mixture. Nuclear Engineering and Design. Vol. 222 (2003), p. 68-78

[8] G.C. Everstine: Prediction of low frequency vibrational frequencies of submerged structures. Journal of Vibration and Acoustics. Vol. 113 (1991), p. 187-191

[9] Tsukasa Ohshima and Yoshihiko Sugiyama: Effect of aerodynamic loads on dynamic stability of slender launch vehicle subjected to an end rocket thrust. ASME 2002 International Mechanical Engineering Congress and Exposition (IMECE2002). New Orleans, Louisiana, USA. (2002).

[10] J.N. Newman. *Marine Hydrodynamics* (The MIT press, 1977).

doi:10.4028/www.scientific.net/AMM.50-51.654

Dynamic Performance of Vehicle-Turnout-Bridge Coupling System in High-Speed Railway

CHEN Rong[1, a], WANG Ping[1, b], QUAN Shun-xi[1, c]

[1]MOE Key Laboratory of High-speed Railway Engineering, Southwest Jiaotong University, Chengdu 610031, China; PH: (86)028-87601210

[a]chenrong@home.swjtu.edu.cn, [b]wping@home.swjtu.edu.cn, [c]jdqsx@126.com

Keywords: vehicle, bridge, turnout, coupling vibration, turnout-bridge relative position, wheel/rail contact relation

Abstract. In order to study dynamic behavior of vehicle-turnout-bridge coupling system, a vehicle-turnout-bridge dynamic analysis model is established by employing the dynamic finite element method (FEM). When No.18 crossover turnouts(with a speed of 350km/h) are laid symmetrically on the 6×32m continuous beam, influences of turnout/bridge relative position and wheel/rail contact relation in turnout zone on the system dynamic responses are analyzed. The result shows that: wheel/rail contact of turnout zone (especially the frog) has great effect on dynamic responses of turnout on bridge, thus the nose rail height of frog should be optimized to mitigate the wheel load transition and its longitudinal gradient. In terms of the 32m-span continuous beam, the best relative position is frog part of turnout arranged in the range of 1/8 and 1/4 of span.

Introduction

With the rapid development of high-speed railway, a lot of researches have been made in train/track/bridge dynamics both at home and abroad [1~3]. Turnout is one of important railway infrastructures in high-speed passenger dedicated railway lines (PDLs). The dynamic optimization of turnout design remains a key problem that need to be solved because a turnout has complex structure and variable wheel/rail contact geometry. When it is laid on the bridge, two critical technologies need to be considered. One is longitudinal static interaction between bridge and turnout; the other is vertical and lateral dynamic interactions within train/turnout/bridge coupling system. There are many factors that have influences on the dynamic interaction of this system, such as vehicle's type and speed, turnout's type, track irregularity of turnout, foundation under rail, track stiffness, bridge structure type and stiffness, relative position between turnout and bridge, and so on.

In this study, the dynamic performance of vehicle-turnout-bridge coupling system and optimization of its parameters for structure design are studied by applying a numerical simulation method through investigating such aspects as train's running quality, rail stress state of turnout, vibration of turnout and bridge, deformation, etc.

Vibration Model of Vehicle-Turnout-Bridge System

Vehicle Model. This paper mainly studies the dynamic characteristics of vehicle-turnout-bridge system under high-speed running condition. Focusing on four-axle vehicle, a dynamics model is established. The carbody, bogies and wheels are all viewed as rigid body that has 5 degrees of

freedom (DOFs), i.e. lateral vibration, vertical vibration, side rolling, shaking, and nodding [4]. The whole vehicle subsystem has 35 DOFs, see Table 1.

Table 1 DOFs of four-axle vehicle model

DOFs	Lateral vibration	Vertical vibration	Side rolling	Shaking	Nodding
Carbody	Y_C	Z_C	ϕ_C	ψ_C	β_C
Bogies(i=1，2)	Y_{ti}	Z_{ti}	ϕ_{ti}	ψ_{ti}	β_{ti}
Wheels(i=1～4)	Y_{wi}	Z_{wi}	ϕ_{wi}	ψ_{wi}	β_{wi}

Model of Turnout on Bridge. Fig.1~Fig.3 is the vibration model of No.18 single turnout with movable nose rail on bridge and it has following features:

(1)Vibration of each rail in turnout is considered. Sleeper supporting point is viewed as a node, and the rail structure is dispersed. Rail is regarded as an Euler-beam which is bendable vertically and laterally. Switch rail, movable nose rail and wing rail are non-uniform beams, and the rest are uniform beams. Each rail node has 4 DOFs, namely, vertical displacement and declination, lateral displacement and declination [5].

(2)For ballasted track, eccentric load and bending deformation of sleeper are considered, while for ballastless track, vibration of the whole track bed is considered. Rail's buckle node is viewed as a node, and the sleeper is dispersed. Sleeper is viewed as a one-way bendable Euler-beam in vertical plane, and a rigid mass-block in lateral plane. Each sleeper node has 3 DOFs, i.e. vertical displacement and declination, lateral displacement.

(3)The connection between rail and sleeper is viewed as a spring/damping system, whose stiffness and damping vary with the different supporting situation. The connection between track bed and sleeper is also viewed as spring/damping system, and the supporting stiffness and damping of the track bed are uniform distribution in the longitudinal of sleeper.

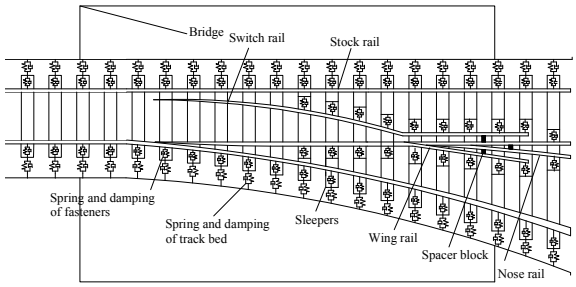


Fig. 1 Plan of vibration model for turnout on bridge

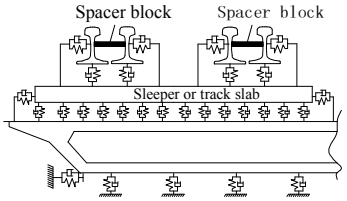


Fig. 2 Cross section of the switch part model

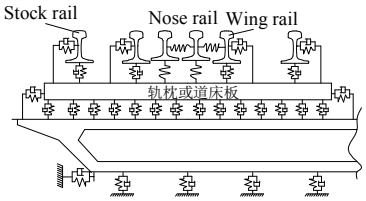


Fig. 3 Cross section of the frog part model

(4)The girder is viewed as an Euler-beam with variable section. As to ballastless track, the rail and beam are directly connected, thus there are interaction between the two. As to ballasted track, the interaction between track and beam is disperse as a series of point-point interaction, and the action point is connected by linear spring and damping [6].

Vibration Equation of the System. Vibration equation of this vehicle-turnout-bridge coupling system is established by adopting the FEM. Vehicle, turnout and bridge are dispersed as an integrated group of finite elements respectively. A proper displacement mode is chosen for each element. The kinetic energy and potential energy of each element and the whole system is calculated, and then the vibration equation is deduced via Hamilton theory. Therefore, vibration differential equation group of

the vehicle, turnout and bridge are obtained, as shown in Eq. 1[7].

$$M(\ddot{u}) + C(\dot{u}) + K(t,u) = P(t,u,\dot{u}) \quad (1)$$

In Eq. 1, M is the total mass matrix of system; C is the total damping matrix of system, K is the total stiffness matrix of system; P is load vector of the system; T is time.

When the vehicle passes each element in sequence at time T, the mass matrix, damping matrix, stiffness matrix and load vector of each carbody are at first obtained. Then according to the relation of vehicle/turnout and bridge/turnout, the mass matrix, damping matrix, stiffness matrix and load vector of the whole system are assembled. Park method is adopted to solve this vibration equation.

Influences of Relative Position between Turnout and Bridge

Calculation Parameters. No.18 crossover turnouts (with a speed of 350km/h) are laid symmetrically on the 6×32m continuous beam, and the foundation under rail is longitudinal continuous slab track. CRH3 EMUs passes the turnout through the main line at 385km/h; frog part of turnout is set on the third span of the continuous beam, considering its position change between the left and right pier. The influence of turnout/bridge relative position on turnout on bridge and vehicle's dynamic responses are analyzed.

Dynamic Responses and Its Analysis. Table 3 shows the change of dynamic responses varies with relative turnout-bridge position:

(1)When turnout/bridge relative position changes, it causes the change of beam's vertical deflection and rotation. When the frog moves from the left pier to the right pier, the vertical deformation of the beam first increases and then gradually decreases to 0. And the maximum deformation appears during frog's moving from midspan to the right pier. Accordingly, the vertical rotation value changes from negative to positive and then becomes negative again. Fig.4 and Fig.5 show that the vertical deflection reaches its minimum near 1/8-span, and it becomes bigger near 3/8-span and during frog's moving from midspan to the right pier.

Table 2 Dynamic responses change with relative position of turnout/bridge

Position of frog	Left pier	1/8 span	1/4 span	3/8 span	1/2 span	5/8 span	3/4 span	7/8 span	Right pier
Dynamic wheel-load(kN)	188.1	185.0	178.4	179.4	178.4	186.0	182.9	187.0	187.0
Flange force (kN)	58.7	59.7	60.0	58.7	58.1	58.7	58.7	58.7	58.7
Wheel unloading rate	0.78	0.71	0.70	0.77	0.72	0.81	0.72	0.82	0.75
Derailment coefficient	0.70	0.61	0.61	0.66	0.65	0.59	0.59	0.60	0.76
Dynamic stress of rail (MPa)	159.4	166.6	171.2	177.4	183.3	187.4	193.2	197.6	201.0
Excursion of switch rail (mm)	0.41	0.41	0.41	0.41	0.41	0.42	0.42	0.42	0.43
Excursion of nose rail (mm)	0.46	0.29	0.39	0.44	0.37	0.26	0.46	0.30	0.39
Lateral vibration acceleration of carbody (g)	0.036	0.036	0.036	0.037	0.037	0.038	0.037	0.037	0.037
Vertical vibration acceleration of girder (g)	0.025	0.066	0.097	0.187	0.207	0.166	0.078	0.036	0.029
Lateral vibration acceleration of girder (g)	0.022	0.025	0.029	0.039	0.037	0.033	0.031	0.025	0.024
Girder's deflection at max wheel unloading rate(mm)	-0.017	-0.034	0.008	0.094	0.291	0.318	0.263	0.092	0.003
Vertical deflection change rate at max wheel unloading rate(0.001rad)	-0.020	0.013	0.024	0.032	0.015	-0.035	-0.022	-0.039	-0.030

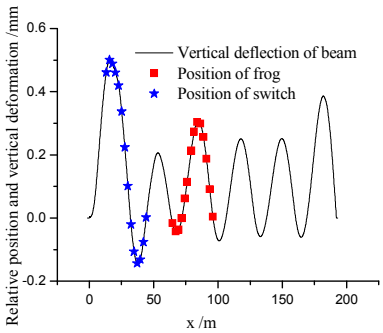


Fig.4 Vertical deflection distributing of beam

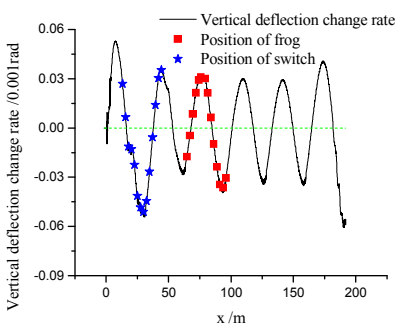


Fig.5 Vertical deflection change rate of beam

(2)The turnout/bridge relative position has some but not significant influence on dynamic wheel load, flange force, vehicle running stability, excursion of switch rail and nose rail. Dynamic wheel load at the pier is slightly bigger than that at midspan, where the carbody's lateral vibration acceleration is a bit smaller than that at the pier.

(3)The maximum wheel unloading rate at the frog part of turnout increases with the increasing of vertical deflection change rate of beam. The change rate of vertical deflection becomes 0 near 1/8 of span, and the maximal wheel unloading rate is 0.69; near 3/8 of span, the vertical deflection change rate of beam reaches its maximal positive value, and the maximal wheel unloading rate is 0.77; near 5/8 and 7/8 of span, vertical deflection change rate of the beam reaches its maximal negative value, and the dynamic wheel load is increasing, thus the maximal wheel unloading rates both exceed the allowable value, reaching 0.81、0.82 respectively.

(4)Wheel/rail lateral force doesn't change a lot with the change of turnout/bridge relative position, but the fluctuation of dynamic wheel load is great at the pier's (where supporting stiffness is big) and in the midspan (where beam's vibration) is violent. As a result, derailment coefficient of the vehicle increase significantly when the frog is at the left and right pier or near the midspan. Besides, as more vehicles run onto the bridge, the further the frog is from the left pier, the bigger the vertical and lateral deformation of beam become, and the rail's dynamic stress increases linearly.

(5)Continuous beam is usually a viable section structure. Near the pier the beam section is higher, and mass, vertical and lateral flexure stiffness of the beam are all bigger. Meanwhile, the pier restrains the vertical and lateral deformation of beam so that the beam's vertical and lateral acceleration increase as frog moves to near the midspan.

By comparison, when the frog part of turnout is arranged in the range of 1/8 and 1/4 of span, the dynamic responses of the whole system is small, so it is the best turnout/bridge relative position.

Influences of Wheel/Rail Contact Relation in Turnout Zone

Table 3 shows the design of wheel/rail relation at frog [8]. The former is listed at the end of this paper.

Table 3 Design of wheel/rail relation at frog

Designs of	Reduction value of nose rail's height (mm)					Horizont	Rail base slope
	Twnr 5mm	Twnr 20mm	Twnr 22.5mm	Twnr50mm	Twnr 65mm		
Scheme 1	14	5.4	4	0	0	9	1:40
Scheme 2	8	4.6	4	0	0	9	1:40
Scheme 3	14	3	2.75	0	0	9	1:40
Scheme 4	14	5.4	4	1	0	9	1:40
Scheme 5	14	5.4	4	0	0	0	1:40
Scheme 6	14	5.4	4	0	0	5	1:40
Scheme 7	14	5.4	4	0	0	9	1:30
Scheme 8	14	5.4	4	0	0	9	1:20

Twnr =Top width of nose rail (mm)

The results show that dynamic responses of the coupling dynamic system vary a lot when the wheel/rail relation changes, such as reduction value of nose rail's height, rail base slope and how much nose rail is hidden horizontally by wing rail, etc.

(1) At frog, when the top surface width of nose rail is within range of 20～50mm, wheel load is transited partially between nose rail and wing rail. If the top width is <20mm, when the longitudinal slope of nose rail top surface is reduced, the influence on the system is small. If the top width is =20mm, when the longitudinal slope (within the range of wheel load transition) of nose rail top surface is reduced by adjusting nose rail's height, such responses as the dynamic wheel load, wheel unloading rate, derailment coefficient and vertical vibration acceleration of girder are greatly reduced. When nose rail is restricted by the mechanic and electric integration structure at the first

traction point, longitudinal slope of its top surface can only be reduced by decreasing the height reduction value of nose rail which has a top width of 50mm.

(2) The less the tip of nose rail is hidden horizontally, the greater the lateral and vertical irregularity of track become, and accordingly the dynamic responses strengthen.

(3) LMA wheel tread was used for EMUs. As the rail base slope increases from 1:40 to 1:20, the vertical irregularity caused by the change of wheel/rail relation decreases. So do such responses as the dynamic wheel load, wheel unloading rate, derailment coefficient and vertical vibration acceleration of girder. But the vehicle's lateral vibration acceleration increases due to the expansion of wheel load transition caused by the increase of rail base slope.

Summary

According to the study, the following conclusions can be made:

(1) Dynamic interactions strengthen when a turnout is laid on a bridge where it is further excited by bridge deformation and vibration. Therefore, it should be judged from such aspects as train's safety and comfort, structure intensity of turnout and dynamic evaluation indicators of turnout and bridge structure to estimate whether the bridge structure meets the laying requirement of turnout or not.

(2)With the change of frog's position on bridge, dynamic responses of vehicle-turnout-bridge coupling system vary a lot. The derailment coefficient of the EMUs increase significantly when the frog is at the left and right pier or near the midspan; maximal wheel unloading rate increases significantly as the change rate of vertical deflection increases; the beam's vertical and lateral acceleration increase as frog moves to near the midspan. As for 32m-span continuous beam, the best turnout/bridge relative position is the frog part of turnout arranged in the range of 1/8 and 1/4 of span.

(3) Responses of coupling dynamic system vary a lot when the wheel/rail relation changes in turnout zone. The reduction value of nose rail's height, rail base slope and how much nose rail is hidden horizontally by wing rail should be more optimized.

Acknowledgment

The authors wish to acknowledge the support and motivation provided by Natural Science Foundation of China (N0.51008256), Fundamental Research Funds for the Central Universities (No.SWJTU09BR017).

References

[1] C.B. Cai: *Theory and Application of Train-Track-Bridge Coupling Vibration in High-Speed Railways.* Ph.D. Thesis, Southwest Jiaotong University (2004).

[2] K.H. Chu, V.K. Garg, A. Wiriyachai: Vehicle System Dynamics, Vol. 9(1980), p207.

[3] K. Liu, G.De Roeck and G.Lombaerta: Journal of Sound and Vibration, Vol. 325(2009), p 240.

[4] W.M. Zhai: *Vehicle-Track Coupling Dynamics*. (Science Press, China 2007).

[5] P. Wang: *Research on Wheel/Rail System Dynamics on Turnout.* Ph.D. Thesis, Southwest Jiaotong University (1997).

[6] R. Chen: *Theory and Application of Vehicle-Turnout-Bridge Coupling Vibration in High Speed Railway*. Ph.D. Thesis, Southwest Jiaotong University (2009).

[7] R. Chen, P. Wang, X.P. Chen: Journal of Southwest Jiaotong University, Vol. 43 (2008), p. 361.

[8] P. Wang, X.Y. Liu: *Computational Theory and Design Method of Jointless Turnout*. (Southwest Jiaotong University Press, China 2007).

doi:10.4028/www.scientific.net/AMM.50-51.659

Automatic texture mapping algorithm for plant leaf model

Yuandi Zhao[1,2,a] Zhixun Su[2,b] Chunjiang Zhao[1,c,*] Xinyu Guo[1,d] Yanqi Liu[2,e]

[1]China National Engineering Research Center for Information Technology in Agriculture, Beijing, China, 100097

[2]School of Mathematical Sciences, Dalian University of Technology, Dalian, China, 116024
[a]doppzyd@gmail.com, [b]zxsu@dlut.edu.cn, [c]zhaocj@nercita.org.cn, [d]guoxy@nercita.org.cn, [e]liuyanqi1986@163.com
*Corresponding author

Keywords: Texture mapping, Plant leaf model, Parameterization, Boundary mapping

Abstract. Texture mapping of plant leaf is a challenging task because of the shape diversity. Based on the barycentric mapping parameterization method, a novel algorithm is proposed. A suitable boundary mapping is computed by automatically extracting and matching the feature points on boundaries of the mesh and image. The matching process consists of two steps, which utilize the symmetry and curvature information respectively. Through iteratively minimizing a weighted energy, the inner constraints given by users in order to make the details clearer can be satisfied by the presented algorithm. Experiments show that this robust algorithm can produce plant leaves with fine texture details and increase the realism of the models.

Introduction

Plant visualization is a key component in virtual agriculture. The triangle mesh of plant with its related texture image would be more realistic and easily edited on computer, such as animation, morphing, and shape analysis. Texture mapping of a given triangle mesh with its texture image is implemented by the parameterization technique, which computes a bijective mapping between the mesh and image.

Parameterization technique has been studied extensively over the past decades. There are many parameterization methods based on Tutte's barycentric mapping theorem [1, 2, 3], where boundary mapping should be computed in advance. Generally speaking, the boundary of the parameter domain should be a regular one, such as a square or a circle, which would not be appropriate for the complicated plant model. As a crucial step in the parameterization process, boundary mapping will greatly affect the validity and conformality of the parameterization results.

Leaf is a major organ of the plant. Compared with other models, the boundary mapping is hard to construct due to its various shapes. Noting that every kind of leaves more or less has some feature points on its boundary, the structure will be much simplified if we make use of these feature points to separate the boundary into several parts. However, most existing methods specify these points by hand, which reduce the efficiency a lot.

Presented is a novel approach which automatically produces a suitable boundary mapping for the parameterization of leaf mesh based on feature points. With this boundary mapping, the implementation of texture mapping will be simple and effective. In addition, our algorithm can also effectively solve the common problem that the users want to constrain some inner points to the given positions in order to keep texture details better.

Method

There are two processes in our algorithm, which are called boundary mapping process and mesh parameterization process respectively.

Boundary mapping: The boundary curves of mesh and image could be obtained by classical segmentation method [4]. And the feature points on their boundaries could be extracted by Douglas-Peucker algorithm [5] for its simplicity.

The most important problem is how to match the feature points between mesh and image boundaries. To solve this problem effectively, we adopt two matching procedures as follows.

Firstly, the coarse matching of feature points is achieved easily through the symmetry information, which is supplied by applying Principal Component Analysis (PCA) algorithm to all points in mesh and image. Fig 1 is an example which illustrates feature points and symmetry information in mesh and image. The symmetry axes separate the boundaries into two half-boundaries.

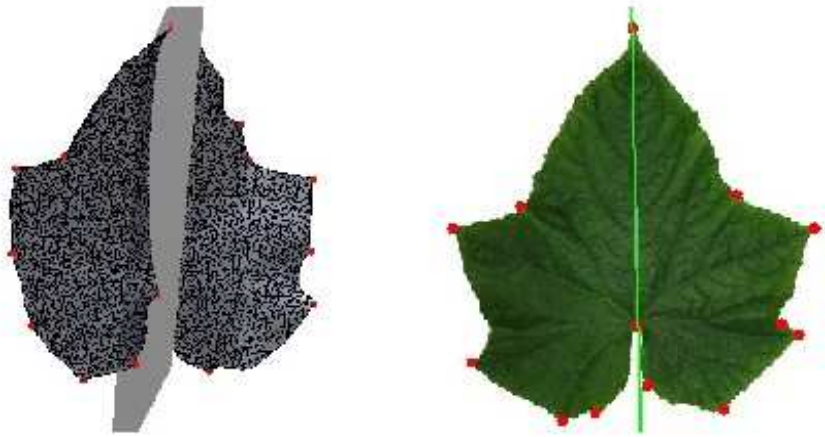


Fig. 1 Feature points and symmetry information in mesh and image

Secondly, the exact matching procedure will be executed. In this key step, two properties of every feature point are computed. One is the chord length relative to the half-boundary which it belongs to. The other is the signed curvature where positive means convex at that point. Considering that the boundary in image can be approximated by the polygon which is outlined with the feature points, we apply discrete differential geometry theory [6] to compute the curvature at a boundary point v both in mesh and image as follows:

$$K(v) = \alpha \tag{1}$$

where α is the signed exterior angle at point v, as shown in Fig 2.

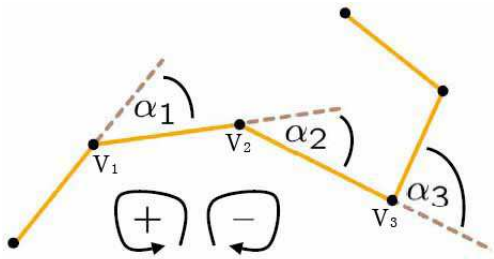


Fig. 2 Curvature of boundary point in mesh and image

After assigning the two properties described above, we execute the following steps in order to match feature points exactly (only take the left-half-boundary for example).

1. Select n feature points in mesh and image respectively according to the modulus of their curvature in descending order.
2. Classify the feature points into positive and negative categories according to their curvature values, and descendingly sort them in mesh and image respectively. After this step, we have four sets of

feature points, i.e. the convex and concave point sets for mesh and image: CVmesh = {cvm1, cvm2…}, CCmesh = {ccm1, ccm2…}, CVimage = {cvi1, cvi2…}, CCimage = {cci1, cci2…}.

3. As for feature point cvm1, search the most corresponding point in set CVimage by their difference of the chord length. If the least difference satisfies a given threshold, these two points are matched. Otherwise, there is no corresponding point in image for cvm1. The remaining points in CVmesh and CCmesh can be processed in the same manner.

When the two matching procedures are finished, some of the feature points on both boundaries are matched. These points separate the boundary into several parts that look like straight lines, so we can parameterize the mesh boundary piecewisely. The feature points on mesh boundary would be parameterized onto their related positions on image boundary.

Mesh parameterization: In this part, the whole mesh is parameterized through minimizing the following function according to the boundary condition as computed above.

$$E(p) = \sum_{(i,j)\in E} K_{ij} \left\| p_i - p_j \right\|^2 + \alpha_k \sum_{k\in C} \left\| p_k - p_k^0 \right\|^2 \qquad (2)$$
$$s.t. p_b = p_b^0, \quad b \in B$$

where the famous cotangent weight K_{ij} is computed as in [2]; α_k is the weight of the kth constrained point; E, C, B stand for the edge set of the whole mesh, the constrained point set given by users and the boundary point set respectively; p_i is the parameter coordinates of the ith mesh vertex; p_k^0 and p_b^0 denote the parameter coordinates of given constrained positions and boundary conditions.

Note that if users do not give additional constraints for inner points, which means all weights α_k are zero, the function degenerates to the classical one in [2]. If users constrain some inner points' positions, we set all weights $\alpha_k = 1$ for initialization. The function (2) should be minimized iteratively by updating α_k as follows until convergence.

$$\alpha_k = d_k \Big/ \min\{d_k\}, \; k \in C \qquad (3)$$

where d_k is the distance between current and specified positions of vertex p_k in previous iteration. By this rule, the conformality of the parameterization can be well preserved while the inner constrained conditions are satisfied.

Results

The developed algorithm is used to map the leaf image onto its original mesh. Two different kinds of leaves which can represent most leaf shapes are used in our experiments in order to show the efficiency and robustness of our algorithm. One is tobacco leaf which is long and thin. This kind of leaves looks like an ellipse, so the number of feature points is less. The other is cucumber leaf whose boundary has more feature points. The shape is so complicated that the desired boundary mapping is hard to obtain by traditional methods.

In the exact matching procedure, we set $n = 6$. Fig 3 shows the experimental results. Fig 3 (a) and (b) are the original meshes of tobacco leaf and cucumber leaf, and Fig 3 (c) and (d) are their corresponding texture images. Fig 3 (e) and (f) are the parameterization results, while Fig 3 (g) and (h) are the texture mapping results. It is displayed clearly that the leaf's features, not only the feature points of the boundary but also the details such as the veins, are mapped successfully with its texture image. The whole leaf meshes with texture images are more realistic.

Note that in Fig 3 (e) and (g), users constrain two extra inner points (red) in order to guarantee the long vein is mapped correctly.

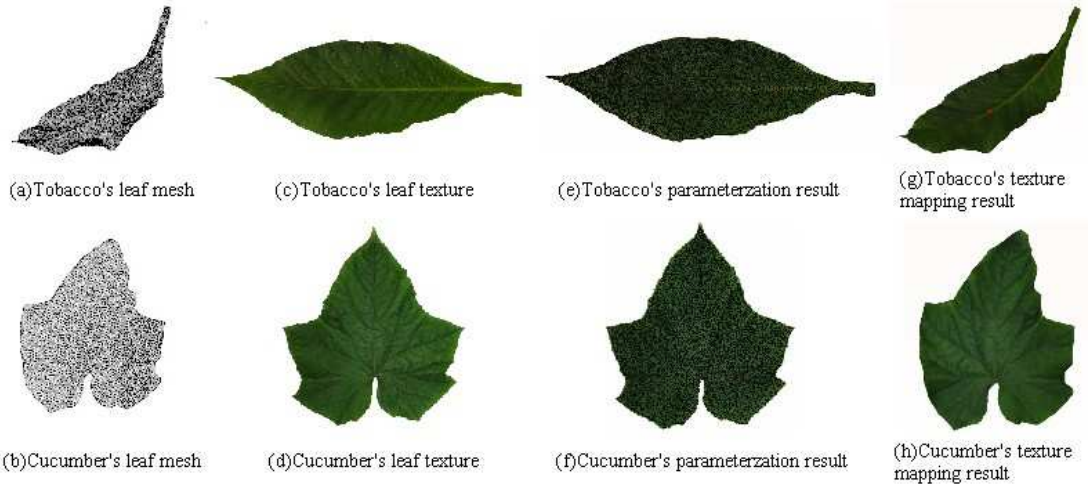


Fig. 3 Texture mapping experiments

Conclusions

Presented is an automatic algorithm for the texture mapping of plant leaf model. After feature points of boundaries are extracted, it takes two steps called coarse matching and exact matching to complete the matching process which is the most important mission. Then the boundary mapping is computed through piecewise chord length parameterization so that the parameterization of the whole leaf mesh is constructed. Besides, users can also specify some inner constraints for the purpose of obtaining a more realistic model. Experimental results show that the algorithm can produce a good texture mapping of the leaf model. Two typical kinds of leaves are tested to show our algorithm is robust to various leaf shapes.

Acknowledgement

This work is supported by Key Technologies Research and Development Program of China (Grant No. 2009BAK43B18), National High Tech. R&D Program of China (Grant No.2008AA10Z219), National Natural Science Foundation of China-Guangdong Joint Fund (No. U0935004), Fundamental Research Funds for the Central Universities, New Century Excellent Talents in University of China (No. NCET-05-0275).

References

[1] W. T. Tutte: How to draw a graph. *London Mathematical Society*, Vol. 13 (1963), 743-768.

[2] M. Eck, T. Derose, T. Duchamp, H. Hoppe, M. Lounsbery and W. Stuetzle: Multiresolution analysis of arbitrary meshes. *ACM SIGGRAPH 95*, (1995), 173-182.

[3] M. Floater: Mean value coordinates. *Computer Aided Geometric Design*, Vol. 20 (2003), 19-27.

[4] D. Comaniciu and P. Meer: Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.24 (2002), 603-619.

[5] D. H. Douglas and T. K. Peucker: Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *The Canadian Cartographer,* Vol.10 (1973), 112-122.

[6] M. Desbrun, K. Polthier, P. Schroeder and A. Stern: Discrete Differential Geometry: An Applied Introduction, course notes. *ACM SIGGRAPH* 06, (2006).

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.663

Intelligent Control Mining Management System In an Open Pit Based on GIS/GPS/GPRS/RFID

Qinghua GU[1, a], Caiwu LU [1,b] , Shigun JING [2,c] ,Xin MA[3,d]

[1] School of Management, Xi'an University of Architecture &Technology, Xi'an, China

[2] Luoyang Luanchuan Molybdenum Industry Group.Inc.,Luoyang,Luoyang, China

[3] Xi'an Institute of Measurement Technology, Xi'an, China

[a]qinghuagu@126.com, [b]lucaiwu@126.com, [c]lyjingshg@sina.com.cn,[d]maxin@126.com

Keywords: GIS; GPS; GPRS; RFID; linear programming; mining; open pit

Abstract: Using GIS(Geographic Information System), GPS(Global Positioning System), GPRS(General Packet Radio Service) and RFID(Radio Frequency Identification), an intelligent control mining management system in an open pit has been designed and developed. A linear programming model is set up in a practical application. By the model, the system can automatically draw up production plan of ore blending well every day. The system can monitor and dispatch open-pit trucks and shovels well, and can play back their historical paths. It can monitor and control the process of mining production in real time. By RFID, The system can also count the number of trucks' delivery and shovels' loading automatically. Experiments on real scenes show that the performance of this system is stable and can satisfy production standards of ore blending in open pits.

Introduction

The intelligent management system of mining production in developed countries has become mature. In some mines, special measures adaptable to local conditions have been taken to control ore mining. Especially, computer technologies are applied to control and manage all processes of mining, transportation, processing and storage for controlling in ore quality. For example, DISPATCH system of American Module Company is highly mature and universal. It has developed into the decision-making platform with omni-directional mining production management and control system, namely Intellimine[1]. In Minntac Mine, USA, the approach that loading by shovel-unloading orderly by railway vehicles is adopted to control the balance of ore quality. The ore quality control system dispatches trucks and railway vehicles by computer. The computer networks of ore mining have also been applied in other developed mines such as Newman iron ore and Paraburdoo mine in Australia. Another example is the dispatch system of Hibbing Taconite Company in America. The computer control center is connected with the vehicles dispatched by computer in this system. With the communication of dispatched vehicles' terminals and other information terminals, the workers can command each process in ore blending well according to the changes of ore quality[2]. The fundamental cause of their success in ore mining management is strict and meticulous control system and scientific, advanced technical means.

In most domestic mines, many management systems of mining production only just use computer-related technologies to manage a part of mining production. For instance, to set up ore blending model, such as linear programming, 0-1 integer programming and so on. The model can make short-term quality production plan in mines. However, the systems can't control and manage the entire ore production process in real-time. The control in entire ore production process is one of the most aspects to keep the stability of ore quality and develop digital mine. So, this paper puts forward that using GIS/GPS/GPRS/RFID technologies and linear programming model, the system can implement the real-time control in ore quality, automatically draw up scientific and rational short-term ore blending plan, count the number of trucks' delivery and shovels' loading automatically and dynamically track, dispatch and manage the production equipments to implement ore blending in an open pit.

GIS/GPS/GPRS/RFID

The Geographic Information System (GIS) is one kind of computer system gathering, storing, managing, analyzing, demonstrating and applying geographic information[3]. The task of an intelligent control mining management system in open pits is to track, monitor and manage production equipment, which depends largely on geographic spatial information. Therefore GIS plays an important role in visual supervisory systems of ore blending, real-time dynamic management and assistance in decision analysis.

The Global Positioning System (GPS) is a satellite-based navigation system made up of a network of 24 satellites placed in orbit. With GPS, the intelligent control mining management system provides accurate position (latitude and longitude), speed, bearing, time, track and more basic information of trucks and shovels[4]. Besides, the accurate position of much-pile boundary and blasting holes can be calculated by GPS units. The ore quality of much-pile boundary and blasting holes can be monitored in electronic map.

The General Packet Radio Service (GPRS) is one of GSMPhase2+ standard realization contents and can provide fast data transmission speed6. The intelligent control mining management system in open pits largely uses the advantages of GPRS such as faster speed and instant connections as the need arises and charging by amount of data. It can provide real-time wireless transmission and is very quick without a dial-up modem connection to GPS equipment for position information. That is very important because GPS position information has room for only a small amount of data and needs to be transmitted frequently. So this system can make good use of GPRS to transmit GPS position information and other information.

Radio Frequency Identification (RFID) is one kind of non-contact automatic recognition technologies. The intelligent control mining management system mainly use RFID long-distance recognition technology to improve the poor working conditions, in which the drivers have to record the workload by the manual way before. Then by RFID technology, the system can trace the ore real-timely to provide the basis for ore blending plan.

Principle of intelligent control mining management system

Structure of intelligent control mining management system

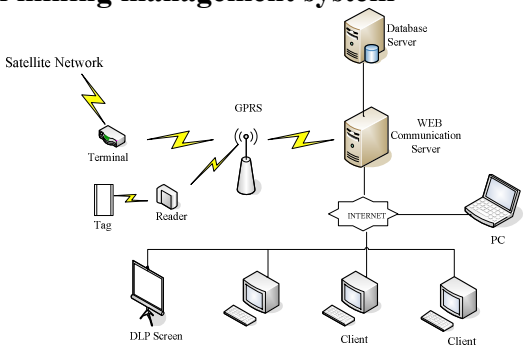


Fig. 1 Intelligent control mining management system structure

The intelligent control mining management system is composed of mobile terminals and RFID tags carried by vehicles, RFID readers, a communication network and a dispatching center, as shown in Fig. 1. In this system the mobile terminals receive GPS signals and then calculate the latitude, longitude, angle, elevation and speed of the vehicles. The expansion interfaces of mobile terminals can also meet many examination and control lines to obtain the information from the vehicles. When trucks arrive at the crushing station, RFID tag attached on trucks enters the scope of RFID read-write antenna, the tag can receive the radio-frequency signal sent by the reader. The tag is powered by the RF field generated by the reader. Upon being 'powered-up', the tag will continuously transmit its data, by damping the incoming RF power field. Then the reader can transmit the data to the GPRS wireless data transmission terminal through the RS232 connection. Each kind of information is transmitted to the monitoring center through GPRS and internet. GPRS, as a telecommunications network between

the mobile terminals and dispatching center, mainly transmits the information on position and condition of the vehicle and information in case of an alarm, to the dispatching center, which in turn transmits dispatch and control commands to the vehicles. In the dispatching center, the communication server, the database server and the console are connected by a 100M local area network. The dispatching center under control of the system software system receives and processes all kinds of information coming from the controlled vehicles. The position, ore grade and other information of the shovels are displayed at multimedia monitors and electronic maps in the dispatching center and from these the vehicles can be monitored and dispatched. Besides, the system may establish other centers having monitoring and dispatch subsystem in the different regions. This sub-center having the independent or the relatively independent monitoring dispatch functions can be controlled in the center [6].

The model and algorithm

1) Linear programming

For the actual mine production, the model of multi-shovels and multi-crushing stations is set up. Assume that the amount of shovels is m, the amount of crushing stations is n, the amount of ore from shovel(i) to crushing station(j) is x_{ij}, the ore grade supplied by shovel(i) is $g_{fi,}$ the ore grade needed by crushing station is $g_{sj.}$ The constraint conditions are as follows[5]:

a) The quantity of supplied ore and production capacity of shovels: The quantity of ore supplied by shovels should be within their range of production capacity and meet the requirement of production task. That is:

$$Q_i \le \sum_{j=1}^{n} x_{ij} \le A_i \quad (i = 1,2,...,m) \tag{1}$$

x_{ij} is the ore amount supplied by shovel(i) for crushing station(j); A_{i} is the most ore supplied by shovel(i); Q_i is the production task of shovel(i);

b) The least task of crushing station: In the ore blending plan, each crushing station will be assigned with least task. Assume that the least production task of crushing station(j) is Qjr, that is:

$$\sum_{i=1}^{m} x_{ij} \ge Q_{jr} \tag{2}$$

c) Non-negative constrain conditions: The ore amount supplied by shovel for crushing station can't be negative, that is:

$$x_{ij} \ge 0 \quad i = 1,2,\cdots,m \quad j = 1,2,\cdots,n \tag{3}$$

d) Objective function:The target ore grade of mine is g and The error range should be less than 5%. According to the actual production in mine, the deviation from actual ore grade and target ore grade in every crushing station should be smallest in the objective function, that is:

$$\min s = \sum_{j=1}^{n} \left| (Q_1 \cdots Q_{j-1} Q_{j+1} \cdots Q_n) \sum_{i=1}^{m} x_{ij} g_{fi} \right| \tag{4}$$

Of which, g_{fi} is the ore grade supplied by shovel(i); Q_j is the task of crushing station(j). According to actual mine requirement, this model is optimized by adding new constrain, that is:

$$\left| \sum_{i=1}^{m} x_{ij} g_{fi} \right| \ge Q_j \bullet g_{sj} . \tag{5}$$

The model is solved by the two-stage method. At the first stage, add many new non-negative variables (x_{n+1}, x_{n+2}, ..., x_{n+h}) to the model. The purpose is that m-moment unit sub matrix is included in the coefficient matrix $\boldsymbol{A}(\boldsymbol{A}=(b_{ij})_{m\times(n+h)}$ ($i=1, 2, ..., m; j=1, 2, ..., n, n+1, ..., n+h$)) of new composed initial simplex table.

At the first stage, use the minimum sum of all artificial variables added as its objective function, that is:

$$\min \quad Z_1 = \sum_{i=n+1}^{n+h} x_i \tag{6}$$

If the optimal solution of Z_1=0 can be obtained, all the added artificial variables are non-basic variables, while m variables before added are basic variables. There is an m-moment unit sub matrix left in the coefficient matrix when the corresponding column of artificial variables are deleted. Assume that it is the initial feasible base B_0, then turn to the second stage of the two-phase method to solve the problem. Otherwise, there is no feasible solution for this problem.

2) Computation of ore quantity

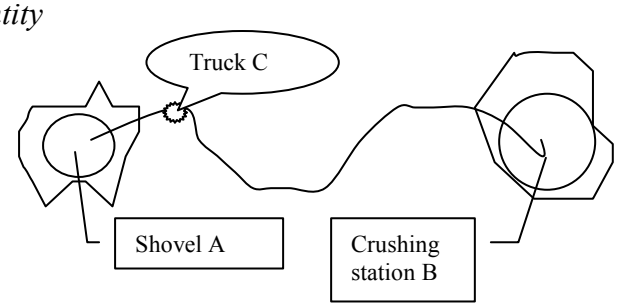


Fig. 2 Statistics algorithm schematic drawing

In system, a RFID reader is carried by each shovel in the loading mine position A and the other reader is fixed in the crushing station B. Then each tag is attached to each truck C, as shown in Fig. 2. It is assumed that *Distance*(*C, A*) is the distance between truck C and each shovel A, and *Distance*(*C, B*) is the distance between truck C and the crushing station. R_1 is the scanning distance of the reader carried by shovel. R_2 is the scanning distance of the reader fixed in crushing station. The value of R_1 and R_2 depends on the actual situation in open pit. For example, R_1= R_2=10m. The algorithm is to judge *Distance*(*C, A*) and R_1 , *Distance*(*C, B*) and R_2 to compute the number of transportation. And by this, ore quantity can be computed approximately.

Intelligent control mining management system in Sandaozhuang open pit

The Sandaozhuang open pit is part of the Luoyang Luanchuan Molybdenum Industry Group Inc. Rotary drilling machines, shovels and trucks are used in this mining process. The intelligent control mining management system in this open pit consists mainly of terminals, tags, readers and management software of mining production. And the software system is composed of modular subsystems. Software constitution and its leading function is as follows:

1) GIS monitoring dispatch subsystem

This part can be installed in the monitoring center mainly to manage, dispatch and monitor trucks and shovels. Its leading function is as follows:

Vehicles display: display some truck or shovel by different ways, different colors and different marks; don't display some truck or shovel according to command.

Historical paths' playback: play back the travel path of some truck in some time on the electronic map.

Trucks localization inquiry: inquire current position, speed, condition and driver so on of some truck at any moment.

Instruction dispatch: the monitoring center can send out the text dispatch instruction; the terminal carried by trucks will clue on remarkably by the red indicating lamp and display the dispatch instruction on the terminal. The topic application is shown in Fig.3.

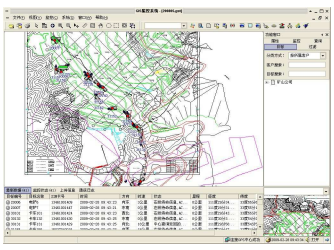


Fig.3 GIS monitoring dispatch interface

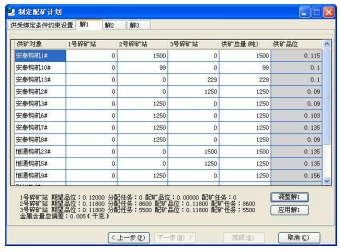


Fig.4 Making ore blending plan

2) Intelligent ore blending management subsystem

The function of this subsystem is mainly to implement the production management of ore blending for shovels in dispatching center, including the automatic drawing up of ore blending plan and dynamic monitoring and controlling the mining operation process of shovels in real time. The main functions are follows:

Muck-piles' management: import the coordinate data of the holes and the boundary of muck-piles, and ore grade of holes; then the holes and the boundary of muck-piles can be shown in the electronic map; according to the ore grade of holes and other properties, random shape ore block in the map can be selected by circularity or polygon and its average grade of ore and its quantity can be calculated automatically.

Drawing the isoline of ore grade: on the basis of muck-piles' data, different grade values can be randomly set and according to this, different isoline of ore grade can be drawn in different colors in the map; the quantity of ore in different area which is divided by isoline also can be calculate.

Making ore blending plan: by the linear programming model, the plan of ore blending every day can be worked out according to the current ore grade, ore loose coefficient and ore lithology of work zone, the shovels' production capacity and the task, capacity and target ore grade of crushing station.The topic application is shown in Fig.4.

3) The ore quantity statistics subsystem

Its function is to record the quantity delivered by trucks and the quantity loaded by shovels. The concrete function is as follows:

Take count of trucks' delivery: count accurately each truck's delivery quantity from shovels to crushing stations during some time interval by RFID and weighing machine; the interval can be set by users; the statistical results can be transformed to report forms of class, month and season according to users' need.

Take count of shovels' loading: count accurately each shovel's loading quantity during some time interval; the other functions are the same as the above.

Summary

1) The intelligent control mining management system in an open pit based on GIS/GPS/GPRS/RFID, uses multi-technologies to control the mining production in the open pit. The practical application indicates that it improves the mineral labor production ability and the production scheduling efficiency greatly and saves the investment and transportation costs.

2) Through the linear programming model and ore quantity computation algorithm, the system changes the irrationality that making the ore blending plan according to subjective experience and mining ore quantity is inaccuracy. It reduces the effects of human factors on production process and provides the scientific basis for ore blending in the open pit.

Acknowledgement

This work is supported by Supported by the key discipline of Shaanxi Province; Shannxi Science and Technology Project (the Key Industries R&D Programme) (2009K08-25);Innovation Scientists and Technicians Troop Construction Projects of Henan Province.

References

[1] L. M.Cheng, X. D. Jiang, J .W. Huang. Application of computer-based dispatch system Dexing copper ore mine. *China Mining Magazine*, 9(2000), p.54-56.

[2] H.Shu Research status and development trend of ore blending at home and abroad. Metal mine design and construction, 03,(1994), p.36-42.

[3] J. P. Liu, L. J. Wei, H .S .Wang. Programming information system of mineral resources based on GIS. Journal of China University of Mining and Technology, vol.33(2004) , p.580-591.

[4] Y.L. Zhang, H. Deng, J.Du. The technical analysis and design of the GPS vehicle position system software. *Computing Technology and Automation*, vol.22 (2003) , p.84-86.
[5] Y.Hou.The Principle and method for the ore blending in open-pit. Mining Engineering, 4(2006), p.22-24.
[6] Q. H. Gu, C. W. Lu, F .B. LI, C. Y .Wan. Monitoring dispatch information system of trucks and shovels in an open pit based on GIS/GPS/GPRS[J]. *Journal of China University of Mining & Technology*, vol.18(2008),p.288-292.

doi:10.4028/www.scientific.net/AMM.50-51.669

Aggregate homotopy method for min-max-min programming satisfying a weak-normal cone condition

H.J. Xiong[1, a] and B. Yu[2,b]

[1]College of science, Huazhong Agricultural University, Wuhan, China

[2]School of mathematical science, Dalian University of Technology, Dalian, China

[a]hj_xiong@mail.hzau.edu.cn, [b]yubo@dlut.edu.cn

Keywords: min-max-min programming; aggregate function; homotopy method; weak-normal cone condition; multiple-instance classification.

Abstract. Min-max-min programming is an important but difficult nonsmooth programming. An aggregate homotopy method was given for solving min-max-min programming by Bo Yu el al. However, the method requires a difficult to verify weak-normal cone condition. Moreover, the method is only theoretically with no algorithmic implementation. In this paper, the weak normal cone condition is discussed first. A class of min-max-min programming satisfying the condition is introduced. A detailed algorithm to implement the method is presented. Models arising from some applications such as support vector machine for multiple-instance classification in data mining, can be included in the problem. In the end, the aggregate homotopy method is given to solve the multiple-instance support vector machine model.

Introduction

Min-max-min programming arises in many applications such as engineering design, circuit design, heat exchange, data mining and etc (see [1,2,3] and etc. for reference). However, because of its non-smoothness and non-convexity, existing methods to solve the problem are few.

In [2], the problem of minimizing a real-valued function f(x) subject to a constraint of the form $\max_{1\le i\le r}\min_{1\le j\le l_i}\{f_{ij}(x)\}\le 0,$ is considered. A transcription technique was introduced such that existing software for smooth nonlinear programming can be suitable to solve the problem. However, the method needs some extra auxiliary variables, whose dimensions depend on the number of "max" and "min" constraints. If the number increased, the dimension of transformed problem will become very large.

Aggregate function approximation is an attractive smoothing technique for max-type nonsmooth optimization problem. In [3], an adaptive smoothing method was given for unconstrained min-max-min problems based on the aggregate function smoothing the inner min-function.

Based on the aggregate smoothing technique and a global convergent homotopy method, an aggregate homotopy method was presented to solve the following general min-max-min programming ([1])

$$\begin{aligned} &\min\ f(x)=\max_{1\le i\le m}\min_{1\le j\le l_i}\{f_{ij}(x)\}\\ &s.t.\quad \mathrm{g(x)}=\max_{1\le i\le k}\min_{1\le j\le p_i}\{g_{ij}(x)\}\le 0. \end{aligned} \tag{1}$$

Compared to some existing methods for solving min-max-min programming, the new method doesn't require any extra variables or solve any extra sub-problems, and it's globally convergent. However, the convergence of the method is only for problems whose constraint set satisfying a so-called weak-normal cone condition. For many problems, it's difficult to verify whether the condition is satisfied. Moreover the work in [1] is only theoretically and doesn't present implementable algorithm.

The paper aims at given some detailed discussions on the weak-normal cone condition with an implementable algorithm. Moreover, the application of the method to multiple-instance support vector machine will be introduced.

Discussion of weak-normal cone condition and implementable algorithm.

Aggregate homotopy method is an efficient globally convergent method to solve min-max-min programming. However, the convergence requires an extra weak-normal cone assumption ([1]). For purpose of applications, it's necessary to give some detailed analysis on the assumption.

Proposition 1. Denote $\Omega_1 = \{x \in R^n : g_{ij}(x) \le 0, 1 \le i \le k,\ 1 \le j \le p_i\}$, if Ω_1 has nonempty interior, and component functions $g_{ij}(x)$ ($i = 1,\ldots k,\ j = 1,\ldots, p_i$) are convex functions, then min-max-min programming defined as (1) is regular and the feasible set satisfies the weak-normal cone condition.

Proof From all $g_{ij}(x)$ are convex functions and Ω_1 has nonempty interior, there exists a $x^1 \in \Omega_1$, such that for arbitrary x satisfying $g_{ij}(x) = 0$, it has

$$(x^1 - x)^T \nabla g_{ij}(x) < 0. \tag{2}$$

If the weak normal cone condition ([1]) doesn't hold, then there exists a closed subset $N \subset \Omega_1$ and a $x \in N$, such that

$$x^1 = x + \sum_{i \in II(x)} \rho_i \sum_{j \in JJ_i(x)} \eta_{ij} \nabla g_{ij}(x),\ \sum_{i \in II(x)} \rho_i > 0,\ \sum_{j \in JJ_i(x)} \eta_{ij} > 0, \tag{3}$$

where $II(x) = \{i \in \{1, \ldots, k\}: g(x) = \min_{1 \le j \le p_i} g_{ij}(x)\}$, $JJ_i(x) = \{j \in \{1, \ldots, p_i\}: g_{ij}(x) = \min_{1 \le j \le p_i} g_{ij}(x)\}$. From (3), it has

$$\sum_{i \in II(x)} \rho_i \sum_{j \in JJ_i(x)} \eta_{ij} (x^1 - x)^T \nabla g_{ij}(x) > 0,$$

which contradicts to (2). As a result, the weak normal cone condition holds only if all $g_{ij}(x)$ are convex functions and Ω_1 has nonempty interior.

The aggregate homotopy method establishes an aggregate homotopy equation ([1]), the equation determines a smooth path approaching to a solution of the problem. In the following, a framework of an Euler-Newton predictor-corrector algorithm is given to path-following the smooth path.

Euler-Newton Predictor-Corrector Algorithm

Parameters. $\theta > 0$, $\varepsilon_p > 0$, $\varepsilon_N > 0$, $h > 0$;

Step 0 (Initialization). Take an initial point $\varpi^{(0)} = (x^{(0)}, \alpha_0) \in N \times R_{++}$, $t_0 = 1$, $d_0 = [0_{n+1}; -1]$, $k = 0$;

Step 1 (Euler Prediction Step). Solve the linear equation $\begin{bmatrix} DH(\varpi^{(k)}, t_k) \\ d_0^T \end{bmatrix} d_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ to obtain a unit tangent vector $d_1 = \dfrac{d_1}{\|d_1\|}$; predictor point $(\varpi 1^{(k)}, t1_k) = (\varpi^{(k)}, t_k) + h d_1$;

Step 2 (Newton Corrector Step). Solve equation $\begin{pmatrix} H(\varpi, t) \\ d_1^T [\varpi - \varpi 1^{(k)}; t - t1_k] \end{pmatrix} = 0$ by Newton iterations with stopping criteria ε_N to obtain a new point $(\varpi^{(k+1)}, t_{k+1})$;

Step 3 (Termination rule). If $t_{k+1} < \varepsilon_p$, the algorithm is terminated and $\varpi^{(k+1)}$ is the solution; else, $k = k+1$, go to Step 1.

Application to multiple-instance support vector machine models.

Multiple-instance classification is an important problem in data mining. It was first given in [4] to determine whether a drug molecule is a musk molecule. The problem can be formulated as:

Suppose every data in the training set is a bag, which consists of several instance, there is a label for every bag. If a bag A_i is labeled as a negative bag, then all instances in the bag are negative points, and if a bag B_i is labeled as positive, then at least one instance in the bag is a positive point. Based on the labeled training bags, the goal of multiple-instance classification is to generate a classifier (decision function), such that the classifier can make a prediction to unknown bags.

Since the multiple-instance classification was given, it has been used extensively in many applications such as image retrieval, stock prediction and etc. Up to now, it's seen as the fourth machine learning framework in machine learning fields (more introductions can be seen from the reference [5]). In [6], the mature Support Vector Machine (SVM) method was generalized first in machine learning to solve multiple-instance classification, and the following non-smooth problem was obtained,

$$\begin{cases} \min\limits_{\varpi,b,\xi} \ \dfrac{1}{2}\varpi^T\varpi + C\sum\limits_{i=1}^{m}\xi_i \\ s.t. \quad \max\limits_{1\le i\le n_+} \min\limits_{j\in P_i}\{1-\xi_i - B_j^i\varpi - b\} \le 0, \\ \qquad \max\limits_{1\le i\le n_-,\, j\in N_i}\{1-\xi_{n_++i} + A_j^i\varpi + b\} \le 0, \end{cases} \tag{4}$$

where P_i is the index set of the instance in the ith positive bag, B_i^j is the jth instance in the ith positive bag, N_i is the index set of the instance in the ith negative bag, A_i^j is the jth instance in the ith negative bag, n_+ denotes positive bags number and n_- denotes negative bags number, $m = n_+ + n_-$ is the total bags number.

All component functions in (4) are linear functions, and a strictly interior point is easy to find. From Proposition 1, model (4) is solvable by aggregate homotopy method.

By taking the initial point as $(\varpi_1^{(0)},\ldots,\varpi_n^{(0)},b_0,\xi_1^{(0)},\ldots,\xi_m^{(0)}) = (0,\ldots 0,0,2,\ldots,2)$, two test data, MUSK-1 and MUSK-2 ([6]), are chosen for comparisons of the aggregate homotopy method (AH) with methods MI-SVM([6]) and mi-SVM ([6]). The results on prediction accuracy are listed in Table 1. From the results, it can be seen that aggregate homotopy method is efficient for solving multiple-instance classification.

Table 1. Results of three methods

Data Set	AH	MI-SVM	mi-SVM
MUSK-1	85.7%	83.1%	87.4%
MUSK-2	90.7%	84.3%	84.9%

Acknowledgement

The research is supported by the Fundamental Research Funds for Huazhong Agricultural University (52204-09146).

References

[1] B. Yu, G.X. Liu and G.C. Feng: *The aggregate homotopy method for constrained sequential max-min problems*, Northeast Math. Vol. 19(2003), p. 287-290.

[2] C. Kirjner-N and E. Polak: *On the conversion of optimization problems with max-min constraints to standard optimization problems*, SIAM J. Optim., Vol. 8(1998), p. 887-915.

[3] E. Polak and J.O. Royset: Algorithms for finite and semi-infinite min-max-min problems using adaptive smoothing technique, JOTA, Vol. 119(2003), p. 421-457.

[4] T.G. Dietterich, R.H. Lathrop and P.T. Lozano: *Solving the multiple-instance problem with axis parallel rectangle*, Artificial Intelligence, Vol. 89(1997), p. 31-71.

[5] Information on http://cs.nju.edu.cn/zhouzh/zhouzh.files/publication/techrep04.pdf.

[6] S. Andrews, I. Tsochantaridis and T. Hofmann: Support vector machines for multiple-instance learning, in: Adv. Neural Inform. Proc. Sys., 561-568(2003).

doi:10.4028/www.scientific.net/AMM.50-51.673

Foreground extraction based on dual-camera system

Li Jue[1, a]

[1]School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou,China

[a]juelee@gmail.com

Keywords: Dual-camera system, stereo image pair, Residual image, Foreground extraction.

Abstract. This paper presents an automatic foreground segmentation algorithm for stereo image pair captured by a dual-camera system. Being different from the monocular image, binocular images contain the disparity map between the stereo image pair. For the disparity map is computationally expensive, our approach adopts the residual image with spatial displacement $(\Delta x, \Delta y)$ to segment the initial trimap automatically. From the residual image, rough region of foreground is clustered as the initial trimap of GrabCut algorithm. Compared with a rectangular region, the calculated trimap is more accurate. After running GrabCut algorithm, the images are segmented into foreground and background layers that comprises of the front objects and back environment. Experimental segmentation results with the original images captured by the dual-camera system indicate that our approach is efficient and promising.

Introduction

Foreground/background segmentation of photographs is a popular computer vision problem and has many important applications. It has long been an active area of research. The challenge of this problem is to segment the foreground/background layer both accurately and efficiently. For most existing methods concentrate on monocular image, which are either computationally expensive or require manual initialization, for instance, magic wand[7], intelligent scissors[4], graph cut[5,6], GrabCut[1,2] and so on. These efficient and accurate methods need user interaction to get an initial user-define trimap.

In this paper, we study the problem of extracting a foreground layer from stereo image pair in real time. Binocular stereo images contain the disparity information of objects in the view. It is easy to generate a 3D anaglyph using the pair of captured images. The key element of our problem is to generate trimaps without user's definition or interaction, and the trimaps ensure the accuracy of extracted results. For the calculation of disparity map between images pair is computationally expensive, our approach avoid the disparity computation and adopt the residual image to generate the rough trimap.

Fig.1 shows the work flow of our approach. The paper is organized as the work flow. In section 2 we describe the simple dual-camera system firstly. For the displacement exists, it is necessary to align the pair of captured images. In section 3 we present the residual image with spatial displacement and the generation of rough fore-mask from the residual image. In section 4 we use the Grabcut algorithm to split the more accurate foreground layer. Experimental results and conclusion are given in section 5.

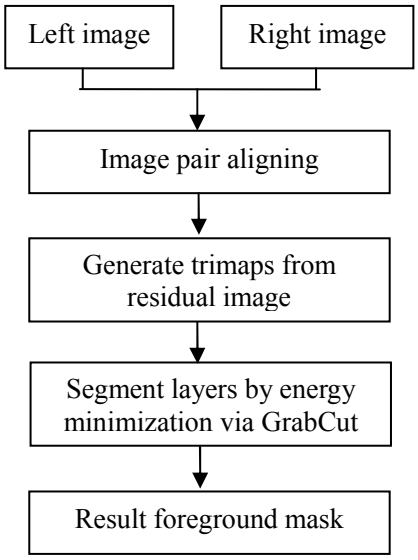


Figure 1. The work flow of our approach

Dual-camera system and stereo image pair aligning

In recent years, 3D technology develops rapidly. It is simple and cheap way to set up a dual-camera system: Two cameras with same brand and model are aligned along the same horizontal plane and positioned parallel at some distance apart from each other. Experimental results show that the best distance is about 62mm which is similar with human interpupillary distance. Using this system, it is easy to record or capture a stereo image pair. From stereo image pair, we can regenerate 3D scene by many modes, for instance, red-blue mode, page flip mode and so on. Note that due to the poor synchronization, this simple dual-camera system does not fit to capture or record the fast moving objects.

(a)The original captured Image pair

(b) Before aligning (c) After aligning

Figure 2.Stereo image pair aligning. (a) The original image pair captured from dual-camera system. (b) and (c) are non-aligned and aligned anaglyphs respectively.

Since two cameras are positioned manually, the displacements (horizontal and vertical direction) between left and right lens are inevitable. Fig.2 first row (a) shows an original image pair captured from the dual-camera system. In order to compare the difference between left and right images clearly, we use anaglyphs (Red-Blue 3D Pictures) to show the disparity between image pair. Fig.2 (b) and (c) show the non-aligned anaglyph and the aligned anaglyph respectively. Obviously, after image pair aligning, most pixels in background region match together. So the disparities of these pixels run to 0.

The basic idea of image pair aligning is to compute the x-dimensional displacement Δx and the y-dimensional displacement Δy between the two cameras. Experimental results indicate that it is enough to compute Δx and Δy once when these two cameras are fixed. Here we use the Optical Flow [3] to calculate Δx and Δy: The edge/corner pixels of image are extracted as feature points. We extract N pairs of feature points $\{P_i^l = (x_i^l, y_i^l) \mid i = 1 \cdots N\}$ from left image and $\{P_i^r = (x_i^r, y_i^r) \mid i = 1 \cdots N\}$ from right image, which are one-to-one mapping. Then

$$\Delta x = \frac{1}{N}\sum_{i=1}^{N}(x_i^r - x_i^l),\quad \Delta y = \frac{1}{N}\sum_{i=1}^{N}(y_i^r - y_i^l) \tag{1}$$

By such coordinate transformation, all left captured images are shifted with $(\Delta x, \Delta y)$ from the origin of itself coordinate system to align the corresponding right images. The motivation of image pair aligning is to remove background pixels in residual image as much as possible.

Remark: In Eq.1, good effect will be taken if these extracted feature points are inside background region, so a better way is to remove front objects when computing Δx and Δy. For our dual-camera system, we keep nothing within 2 meter distance when image pair aligning.

Generate trimap from residual image

It is obvious that the 3D-stereopsis image pair is very similar and most of the content of one image is repeated in the other, which is often described as "highly redundant". After images aligning, the residual image between the pair is given by

$$I(x,y) = I_r(x,y) - I_l(x+\Delta x, y+\Delta y) \tag{2}$$

Where Δx and Δy are the displacements which are computed above.

In order to avoid the influences of noise and illumination between the image pair, the residuals are preserved only where value exceed a predetermined threshold I_{th} .So Eq.2 describes the residual image could be written as

$$I(x,y) = \begin{cases} 0 & \text{if } I(\mathrm{x,y}) \le I_{th} \\ I(x,y) & \text{other} \end{cases} \tag{3}$$

From the residual image, a rough initial trimap is generated as the following steps:

Step 1. Binary processing of the residual image.

Step 2. Morphology operation(erode, dilate, clean) on the binary image.

Step 3. Region labeling and removing small regions (area limits), then removing thin strip.

Step 4. Flood filling of the connected region.

Fig 3 right four images show the result mask image of every step. The last result mask is regarded as the initial trimap for GrabCut input

Figure 3. Residual image and the procedure of trimap generation. The left is residual image. The right 4 images show the procedure of trimap generation (sequence from left to right, top to bottom)

Energy minimization via GrabCut

GrabCut is an innovative 2D image segmentation technique developed by Rother et al.[2]. It is a very user-friendly image segmentation algorithm. For images where the foreground and background are cleanly separated, GrabCut can robustly segment the image given only a rectangular region as

input. In our approach, we use the calculated trimap instead of a user-defined rectangular region. It is obvious that the calculated trimap is more accurate than a rectangular region.

Due to space limitations some details of GrabCut algorithm have been omitted, but can be found in [2]. Here, we simply describe the procedure of Grabcut:

1. Use the calculated trimap as input. Pixels inside the trimap are marked as unknown. Pixels outside of trimap are marked as known background.
2. computer generate an initial image segmentations, where all unknown pixels are tentative placed in the foreground class and all known background pixels are placed in the background class
3. Gaussian Mixture Models(GMMs) are built for initial foreground and background classes using the Orchard-Bouman clustering algorithm.
4. Each pixel in the foreground class is assigned to the most likely Gaussian component in the foreground GMM. Similarly, each pixel in the background is assigned to the most likely background Gaussian component.
5. The GMMs are thrown away and new GMMs are learned from the pixel sets created in the previous set.
6. A graph is built and Graph Cut is run to find a new tentative foreground and background classification of pixels.
7. Steps 4~6 are repeated until the classification converges.

Results and Conclusion

(a) Stereo image pair after aligning

(b) The trimap

(c) Foreground layer

Figure 4: Results using our approach. (a)The original aligned images captured from the dual-camera system. (b)The trimaps which are computed using the residual images. (c) The resut foreground layer extracted by our approach.

In our experiment, we use the simply dual-camera system to capture VGA (640*480) resolution

images. In order to balance the effect and performance, we resample the original image (640*480) into small image (160*120) to accelerate segmentation speed. After computing segmentation, foreground layer resample into VGA resolution. The results are shown in Fig. 4. The operating time is about 0.1sec for VGA images on Intel Core 2 Duo 2.83GHz CPU with 1GB RAM.

Two examples are given in Fig 4. The left is a common background, and the right we try to segment foreground from the complex background (the TV Set is playing behind the tester). The left result image is good quality, and the right result images show that GrabCut segmentation result is not enough accurate when the background is complex. The segmentation result includes part of the frame of TV Set. Because the frame is black and similar with the hair texture, low-contrast is the main factor of GrabCut failure.

In conclusion, a new approach for foreground extraction from binocular images has been proposed and demonstrated. Experimental results indicate that the calculated trimap is reliable, and the quality of segmentation results depends on the segmentation approach (here we use GrabCut). It might improve the foreground segmentation accuracy by adopting more accurate segmentation approach.

References

[1] Blake, A., Rother, C., Brown, M., Perez, P., And Torr, P. Interactive image segmentation using an adaptive GMMRF model. In *Proc. European Conf. Computer Vision*. (2004).

[2] Rother, C., Kolmogorov, V., And Blake, A. GrabCut—interactive foreground extraction using iterated graph cuts. In *Proc. ACM Siggraph* (2004).

[3] Lucas, B., and Kanade, T. An iterative image registration technique with an application to stereo vision. In *Proc. of IJCAI'81*, pp. 674–679. (1981).

[4] Mortensen, E., And Barrett, W. Intelligent scissors for image composition. In *Proc. ACM Siggraph*, pp.191-198. (1995).

[5] Kwatra, V., Schödl, A., Essa, I., Turk, G., And Bobick, A. Graphcut Textures: Image and Video Synthesis Using Graph Cuts. In *Proc. ACM Siggraph*, pp.277-286. (2003)

[6] Boykov, Y., And Jolly, M. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In *Proc. IEEE int. Conf. on Computer Vision, CD-ROM*. (2001)

[7] Adobe Systems Incorp. 2002. Adobe Photoshop User Guide.

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.678

Adaptive Polynomial Approximation to Circular Arcs

Lizheng Lu[1, a]

[1]Department of Mathematics, Zhejiang Gongshang University, Hangzhou 310018, PR China

[a]lulz99@yahoo.com.cn

Keywords: Circular arc; Polynomial; Hermite approximation; s-power series

Abstract. We present a new adaptive method for approximating circular arcs in polynomial form by using the s-power series. Circular arcs can be expressed in infinite series form, we obtain the order-*k* Hermite interpolant by truncating at the *k*th term. An upper bound on the error of the interpolant is available, so we can obtain the lowest degree polynomial curve that approximates a circular arc within any user-prescribed tolerance. And this degree can be further reduced through subdivision, which generates a spline approximation with C^k continuity at the joints.

Introduction

Circles and circular arcs are of great interest for geometric modelling and for problems of shape preservation because of their constant curvature. A polynomial representation for circular arcs is not only simpler for the rendering and curve-curve intersection problems, it is of special interest to the aircraft and the automobile industries where surfaces are frequently constructed from curves involving circular arcs using lofting or skinning techniques.

A circular arc can be explicitly represented by rational curves [1], or can be approximated by polynomial curves [2−4]. Floater [5,6] presented a framework for high-order approximation of conic sections by rational curves. Piegl and Tiller suggested using integral B-splines [7]. Yang and Ye constructed the approximation of circular arcs by C^2 cubic polynomial B-splines [8].

The *s-power series*, firstly introduced by Sánchez-Reyes [9,10] for geometry processing, can be regarded as the two-point analogue of Taylor expansions. The s-power basis combines many desirable properties of the Bernstein, Taylor, and Hermite polynomial bases, without their disadvantages. For more details about the comparisons among the above bases and the properties of the s-power series, we refer the reader to [9,10]. The s-power series has been used in several geometrical operations, including the inversion of analytic functions [11] and the representation of analytic functions [12].

We assume without loss of generality that the circular arc is defined as follows:

$$\mathbf{r}(\theta) = \left(\cos\theta, \sin\theta\right), \quad \theta \in [-\alpha, \alpha], \tag{1}$$

where α is a positive number. One can always achieve such a situation by rotation and translation.

The main idea of this paper is to represent Eq. 1 by the s-power series. We then obtain the polynomial approximation by a truncation at the *k*th term. A circular arc can be approximated within any given tolerance by simply adding the newly-introduced high-degree terms to the approximating polynomial. Since the approximation is expressed in polynomial form, it is very convenient to take differentiation and integration operations.

Preliminaries

A *s-power series* over the unit domain $u \in [0,1]$ is defined by (see [9,10])

$$a(u) = \sum_{k=0}^{\infty} a_k(u) s^k, \quad s = (1-u)u, \tag{2}$$

where

$$a_k(u) = (1-u)a_k^0 + u a_k^1. \tag{3}$$

Note that Eq. 2 is simply a power series of symmetric parameter s, whose coefficients $a_k(u)$ are linear functions of u rather than constants. The function $s^k(u)$ is nothing but the central (kth) scaled Bernstein polynomial $B_k^{2k}(u)$ of degree $2k$. For the sake of simplicity, a natural representation of $a_k(u)$ would be the pair of Bézier ordinates:

$$a_k(u) \to a_k = \begin{Bmatrix} a_k^0 \\ a_k^1 \end{Bmatrix}. \tag{4}$$

Throughout this paper, we adopt the simplified notation in Eq. 4 to represent linear functions. We will also use for a_k the term *coefficient*, although always bearing in mind that we are dealing with linear functions, not with constants. For more details on the s-power series, see [9,10].

Given the s-power series of a function $a(u)$, a truncation at the kth term clearly yields its order-k Hermite interpolant

$$H_k(a;u) = \sum_{i=0}^{k} a_i(u) s^i.$$

This is the unique degree-($2k$+1) polynomial that reproduces the derivatives $\left\{a^{(i)}(u)\right\}_{i=0}^{k}$ at the endpoints u = 0,1. Therefore, if we subdivide the domain into several segments and compute a kth truncation for each one, we get a polynomial spline with C^k continuity.

Polynomial Approximation

Let $x(\theta) = \cos\theta$ and $y(\theta) = \sin\theta$, $\theta \in [-\alpha, \alpha]$. Since the s-power series is always defined over the domain [0,1], we introduce the parameter transformation

$$\theta = \theta(u) = (2u-1)\alpha, \quad u \in [0,1]. \tag{5}$$

Firstly, we express $x(u)$ in the s-power form, i.e.,

$$x(u) = x\big(\theta(u)\big) = \cos\big((2u-1)\alpha\big) = \sum_{k=0}^{\infty} x_k(u) s^k. \tag{6}$$

The initial coefficient x_0 is the linear function interpolating $x(u)$ at the endpoints u = 0,1. Then we get

$$x_0 = \begin{Bmatrix} x_0^0 \\ x_0^1 \end{Bmatrix} = \begin{Bmatrix} x(0) \\ x(1) \end{Bmatrix} = \begin{Bmatrix} \cos\alpha \\ \cos\alpha \end{Bmatrix} = \cos\alpha. \tag{7}$$

The next coefficient x_1 is straightforwardly calculated from the derivatives at the endpoints:

$$x_1 = \begin{Bmatrix} x_1^0 \\ x_1^1 \end{Bmatrix} = \begin{Bmatrix} x'(0) \\ -x'(1) \end{Bmatrix} + (x_0^1 - x_0^0) \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} 2\alpha\sin\alpha \\ 2\alpha\sin\alpha \end{Bmatrix} = 2\alpha\sin\alpha. \tag{8}$$

The remaining coefficients x_{k+2} (k = 0,1,…) are given by the recursive formula (see [11]):

$$x_{k+2} = \frac{2}{k+2}\left(2(k+1)x_{k+1} - \begin{Bmatrix} x_{k+1}^1 \\ x_{k+1}^0 \end{Bmatrix} - \frac{2\alpha^2}{k+1} x_k\right). \tag{9}$$

Noting that the cosine function and $s = (1-u)u$ in Eq. 6 are all symmetric functions, the coefficients x_k must satisfy

$$x_k^0 = x_k^1, \quad k = 0,1,\ldots.$$

Therefore, every coefficient x_k degenerates to a constant rather than a linear function, and Eq. 9 can be further simplified to

$$x_{k+2} = \frac{2}{k+2}\left((2k+1)x_{k+1} - \frac{2\alpha^2}{k+1} x_k\right), \quad k = 0,1,\ldots. \tag{10}$$

Secondly, we express $y(u)$ in the s-power form, i.e.,

$$y(u) = y\big(\theta(u)\big) = \sin\big((2u-1)\alpha\big) = \sum_{k=0}^{\infty} y_k(u) s^k. \tag{11}$$

To obtain the coefficients y_k, we can also use the recursive formula Eq. 9 with the different starting values y_0 and y_1. However, using the relationship between the sine and cosine functions, we propose a simple alternative method that does not require the recursive formula as follows.

Obviously,

$$y(u) = -\frac{1}{2\alpha}x'(u) = -\frac{1}{2\alpha}\frac{\mathrm{d}}{\mathrm{d}u}\sum_{k=0}^{\infty}x_k(u)s^k.$$

Noting that x_k are all constants, we can easily obtain y_k by comparing the coefficients:

$$y_k = -\frac{k+1}{2\alpha}(1-2u)x_{k+1} = -\frac{k+1}{2\alpha}\begin{Bmatrix} x^0_{k+1} \\ -x^1_{k+1}\end{Bmatrix}, \tag{12}$$

which implies that $y^0_k = -y^1_k$ for $k = 0,1,\ldots$.

Applying the parameter transformation in Eq. 5, we can express the circular arc in Eq. 1 as

$$\mathbf{r}(u) = \sum_{k=0}^{\infty}\mathbf{r}_k(u)s^k, \tag{13}$$

where the vector $\mathbf{r}_k(u) = (x_k(u), y_k(u))$ denotes the kth coefficient of the circular arc in the s-power form. And, the components x_k and y_k are obtained by Eq. 10 and Eq. 12, respectively.

We obtain the order-k Hermite approximation by a truncation at the kth term of Eq. 13:

$$\mathbf{H}_k(\mathbf{r};u) = \sum_{i=0}^{k}\mathbf{r}_i(u)s^i. \tag{14}$$

From the *permanence* property of the s-power series [10], all the coefficients of $\mathbf{H}_{k+1}(\mathbf{r}; u)$ agree with those of $\mathbf{H}_k(\mathbf{r}; u)$, except for the newly-introduced one $\mathbf{r}_{k+1}(u)$. As shown by Eq. 10, x_{k+1} depends only on the two previous coefficients x_k and x_{k-1}. Note also that y_{k+1} is determined by x_{k+2}. So, it is very convenient and efficient to deal with polynomials of different degrees.

Error Analysis

The difference vector $\mathbf{R}_k(\mathbf{r}; u) = \mathbf{r}(u) - \mathbf{H}_k(\mathbf{r}; u)$ is called the remainder and measures the quality with which the circular arc $\mathbf{r}(u)$ is approximated by the polynomial $\mathbf{H}_k(\mathbf{r}; u)$. We propose an upper bound on the error of the approximation in the following theorem.

Theorem 1. $\mathbf{R}_k(\mathbf{r}; u)$ converges uniformly to zero over [0, 1] and has an upper error bound

$$\varepsilon_k = \frac{\sqrt{2}}{(2k+2)!}\alpha^{2k+2}. \tag{15}$$

Proof. By Theorem 3.5.1 in ([13], p. 67), the remainder is characterized by

$$\mathbf{R}_k(\mathbf{r};u) = \frac{(-s)^{k+1}}{(n+1)!}\left(x^{(n+1)}(\xi_1), y^{(n+1)}(\xi_1)\right), \quad n = 2k+1, \quad \xi_1, \xi_2 \in (0,1).$$

Note that $x(u) = \cos((2u-1)\alpha)$ and $y(u) = \sin((2u-1)\alpha)$. Applying the chain rule to the trigonometric functions leads to

$$x^{(n+1)}(u) = (2\alpha)^{n+1}\cos\left((2u-1)\alpha + \tfrac{n+1}{2}\pi\right), \quad y^{(n+1)}(u) = (2\alpha)^{n+1}\sin\left((2u-1)\alpha + \tfrac{n+1}{2}\pi\right).$$

Then, we have

$$d_k = \left\|\mathbf{R}_k(\mathbf{r};u)\right\| = \frac{(2\alpha)^{n+1}s^{k+1}}{(n+1)!}\sqrt{\cos^2\theta_1 + \cos^2\theta_2}, \quad \theta_i = (2\xi_i - 1)\alpha + \tfrac{n+1}{2}\pi, \ i = 1,2.$$

Since $s = (1-u)u \in [0, \frac{1}{4}]$, the conclusion can be obtained easily.

For any user-prescribed tolerance δ, it is trivial to find the lowest k such that $\varepsilon_k \le \delta$. So one can obtain the polynomial curve of a lowest degree n ($n = 2k+1$) that approximates the circular arc within the tolerance. It can be considered as an adaptive method, i.e., the user can choose a smaller δ for

higher accuracy and also can choose a larger δ for lower accuracy. Therefore, it brings a convenient interactive editing tool to geometric modelling design.

If the degree of the approximating polynomial is too high for a considerable small tolerance δ, we can subdivide the circular arc into several segments and then reduce the degree for each segment. One can see that the error bound in Eq. 15 will decrease exponentially by a factor $1/4^{k+1}$ if α is replaced by $\alpha/2$. As a result, we can achieve a good balance between accuracy and efficiency. And the resulting piecewise curves are C^k continuous at the joints.

Example

Let us consider the unit circle defined by $\mathbf{r}(\theta) = (\cos\theta, \sin\theta)$, $\theta \in [-\pi, \pi]$. As displayed by Fig. 1, the successive order-k Hermite polynomial curves gradually approximate the whole circle. When $k = 5$, we obtain a good enough approximation, i.e., one cannot distinguish the approximating curve from the circle. However, the degree of the polynomial is 11 in such case. If we subdivide the circle into four segments, a good enough approximation will be achieved when $k = 2$ (see Fig. 2).

The quality of the approximations is measured by the sampling-based Euclidean errors. The error plots are shown in Figs. 1 and 2. Additionally, the maximum errors between the circle and the approximating curves with various segments are listed in Table 1. The values in the column $k = 1$ decrease by about $1/4^2$, those in the column $k = 2$ by about $1/4^3$, those in the column $k = 3$ by about $1/4^4$, and so on. Therefore, one can always obtain a polynomial approximation within any given tolerance and with the constraint on the maximum degree of the polynomial by combining subdivision.

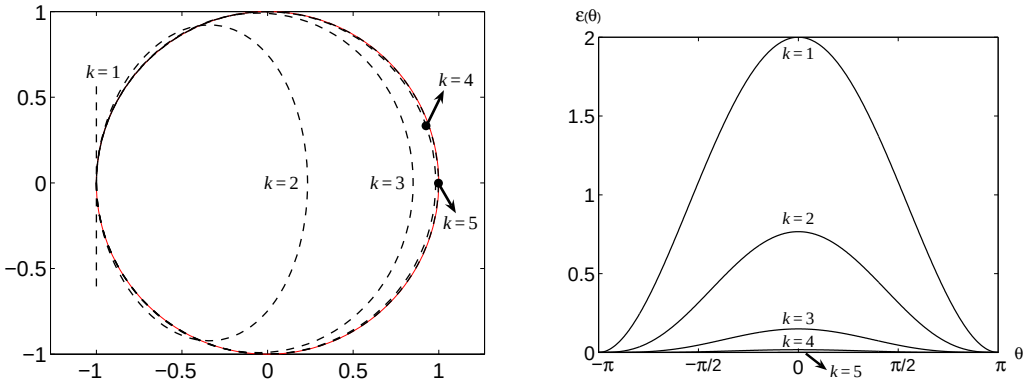


Fig. 1. Successive order-k polynomial approximations to the unit circle.

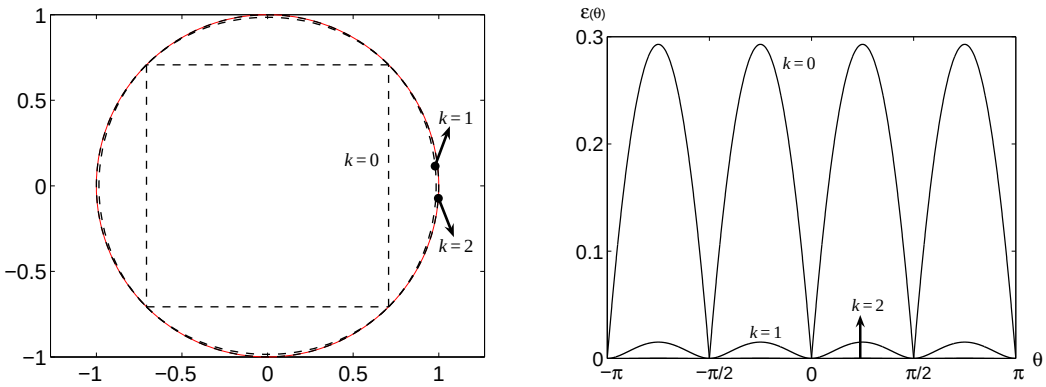


Fig. 2. Successive order-k polynomial approximations to the quarter-circular arcs.

Table 1. The maximum errors for the successive order-k (k = 1,2,…,5) Hermite interpolants to the unit circle with i-segments, i =1,2,4,8,16,32.

#Segemnts	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
1	2	7.663e−01	1.495e−01	1.759e−02	1.385e−03
2	2.146e−01	1.825e−02	8.230e−04	2.294e−05	4.341e−07
4	1.521e−02	3.154e−04	3.494e−06	2.404e−08	1.127e−10
8	9.808e−04	5.052e−06	1.393e−08	2.390e−11	2.795e−14
16	6.177e−05	7.942e−08	5.470e−11	2.346e−14	3.065e−17
32	3.868e−06	1.243e−09	2.139e−13	1.994e−17	3.752e−20

Summary

In this paper, we have constructed a simple method for the polynomial approximation to circular arcs by using the s-power series. Theoretical analysis and experiments show the effectiveness of the proposed method. Furthermore, the error bounds of approximations decrease exponentially after subdivision.

The major limitation of our method is the low accuracy of approximations. For a specific degree $n = 2k+1$, the resulting polynomial curve is actually the order-k Hermite interpolant. So, our method can not do better than those methods that make use of geometric data (tangents, curvatures, etc.). However, we can improve the accuracy by increasing the degree of curves or by subdividing the circular arc into several segments.

Acknowledgements

This work is supported by the Natural Science Foundation of China (61003186), the Natural Science Foundation of Zhejiang Province (Y6090211), and the Education Department of Zhejiang Province (Y201017555, Y201017322).

References

[1] G. Farin: *Curves and Surfaces for CAGD* (Morgan Kaufmann, San Francisco 2001).

[2] M. Goldapp: Comput. Aided Geom. Design Vol. 8 (1991), p. 227

[3] L. Fang: Comput. Aided Geom. Design Vol. 15 (1998), p. 843

[4] S.H. Kim, Y.J. Ahn: Comput. Aided Design Vol. 39 (2007), p. 490

[5] M.S. Floater: Comput. Aided Geom. Design Vol. 12 (1995), p. 617

[6] M.S. Floater: Comput. Aided Geom. Design Vol. 14 (1997), p. 135

[7] L. Piegl, W. Tiller: Comput. Aided Design Vol. 35 (2003), p. 601

[8] W. Yang, X. Ye, in: *Proceedings of CAD/Graphics*, Beijing (2007).

[9] J. Sánchez-Reyes: ACM Trans. Graph. Vol. 16 (1997), p. 319

[10]J. Sánchez-Reyes: ACM Trans. Graph. Vol. 19 (2000), p. 27

[11]J. Sánchez-Reyes: Comput. Aided Geom. Design Vol. 18 (2001), p. 587

[12]J. Sánchez-Reyes, J.M. Chacón: Comput. Aided Geom. Design Vol. 22 (2005), p. 103

[13]P.J. Davis: *Interpolation and Approximation* (Dover Publications, New York 1975).

doi:10.4028/www.scientific.net/AMM.50-51.683

Positive Approximation for Positive Scattered Data

Xiaolei Zhang [1, a], Jinming Wu[1,b]

[1]Department of Information and Computer Science, Zhejiang Gongshang University, Hangzhou, 310018, China

[a]zhangxl0411@yahoo.com.cn [b]wujm97@yahoo.com.cn

Keywords: Positive approximation; Positive scattered data; Gaussian function; CSRBF.

Abstract. The curve and surface fitting problem is very important in CAD and CAGD. However, it is important to construct a suitable function to interpolate or approximate which satisfies the underlying constraints since we have some additional information that is confined to interpolation or approximation. In this paper, we discuss the positive approximation for positive scattered data of any dimensionality by using radial basis functions. The approach is presented to compute positive approximation by solving a quadratic optimization problem. Numerical experiments are provided to illustrate the proposed algorithm is flexible.

Introduction

The curve and surface fitting problem is very important in CAD and CAGD. The problem of scattered data interpolation or approximation is of great significance in various areas of scientific research including scientific visualization, computer graphics, geometric modeling, numerical analysis, etc. Particularly, when the data are arising from some complex function or from some scientific phenomena, it is importantto incorporate the inherited features of the data. It gives an insight and guide to understand some physical phenomenon relevant to the data.

As is known to all, we often have some additional information that we wish to confine to interpolation or approximation. As it is known to all, positivity is a very important aspect of shape. There are many physical situations where entities only have meaning when their values are positive. For example, in a probability distribution the representation is always positive.

Smoothness is not the main requirements for pleasing visual display of the data when we construct interpolating or approximating function. Ordinary spline schemes, although smoother, are not helpful for the interpolation of the shaped data. Severely misguided results will violate the inherited features of the data. Thus, it is important to construct a suitable function to interpolate or approximate which satisfies the constraint of positivity in order to control the data features..

Many researchers have done a lot of significant work on it mainly by using piecewise cubic splines or piecewise algebraic curves (see [1–6] and references therein). The problem we are addressing is the positive approximation of positive scattered data.In this paper, we present a new technique to deal with positive approximation for positive scattered data of any dimensionality by using radial basis functions such as Gaussian function and compactly supported radial basis functions.

Radial basis functions

A radial basis function (RBF) is a relatively simple multivariate function generated by a univariate function. Due to its simple form and good approximation behavior, RBF has become an effective tool for multivariate scattered data interpolation during the last two decades [7-13].

For any given scattered data $(x_i, f_i) \in R^s \times R, i = 1, \cdots, m$ where points x_i are pairwise distinct, the so-called RBF interpolation is to use a function $\phi : R^s \to R$ to construct the interpolant in the form of $s(x) = \sum_{j=1}^{m} \lambda_j \phi(\|x - x_j\|)$ satisfying

$$s(x_i) = \sum_{j=1}^{m} \lambda_j \phi(\|x_i - x_j\|) = f_i,\ i = 1, \cdots, m.$$

The positive definiteness of ϕ guarantees that the above interpolation problem possesses a unique solution and refers to ϕ as a classical RBF. Most-widely used radial basis functions are mainly the following two types: One type of RBFs is globally supported, infinitely differentiable, and contain shape parameter c which affect both accuracy of the solutions and conditioning of the interpolation matrix such as multiquadric function, thin plate spline and Gaussian function, etc. Another type of RBFs is compactly supported radial basis functions (CSRBF) with parameter radius α which affect both the possible approximation behavior and best possible stability [11]. For simplify, from now on we just choose the radial basis functions as Gaussian function $\phi_c(r) = e^{-cr^2}$ and CSRBF $\phi_\alpha(r) = (1 - r/\alpha)_+^4 (4r/\alpha + 1)$ for later discussion.

Positive approximation for positive data

The problem we are addressing is as follows: Given a set of m scattered data $x_i, x_i \in R^s$ with associated data values $f_i, i = 1, \cdots, m$, where $f_i \geq 0$, we seek an optimal function $s(x)$ to approximate the set of data points (x_i, f_i) in least-squares sense satisfying $s(x) \geq 0$.

We first construct the function $s(x)$ as follows

$$s(x) = \sum_{i=1}^{N} \lambda_i \phi(\|x - c_i\|)$$

where $\phi(\|x - c_i\|)$ can be a Gaussian function $\phi_c(r)$ or CSRBF $\phi_\alpha(r)$ centered on c_i. The centers can be randomly sampled among the scattered data or obtained by orthogonal least square learning algorithm or found by clustering the samples and choosing the cluster means as the centers.

It is pointed out that the shape parameter c and radius α can be different for each Gaussian function $\phi_c(\|x - c_i\|)$ or CSRBF $\phi_\alpha(\|x - c_i\|)$ centered on c_i.

Therefore, the positive approximation for positive scattered data is reduced to solve the following optimization problem

$$\begin{cases} \min\limits_{\lambda_1, \cdots, \lambda_N} TotalError, \quad TotalError = \sum\limits_{i=1}^{m} (s(x_i) - f_i)^2 \\ subject\ to \quad s(x) \geq 0 \end{cases}$$

Note that all $\phi(\|x - c_i\|)$ are positive and $s(x)$ is a linear combination of them. So, we can restrict each λ_i to be positive in order to guarantee the function to be everywhere positive. This sufficient condition is the kernel idea of our proposed method.

Therefore, the problem of positive approximation is transformed into a quadratic optimization problem subjected to linear constraints

$$\begin{cases} \min\limits_{\lambda_1,\cdots,\lambda_N} TotalError \\ subject\ to \quad \lambda_1 \geq 0,\cdots,\lambda_N \geq 0 \end{cases}$$

Without loss of generality, we illustrate it with a very simple example in 1D [14]. The data set in Tab.1 shows the percentage of oxygen in the flue gas, as coal burns in a furnace. The oxygen percentage is inherently positive and we therefore require the approximation to preserve this property.

Time(mins)	0	2	4	10	28	30	32
Oxygen(%)	20.8	8.8	4.2	0.5	3.9	6.2	9.6

Table 1: Percentage of oxygen in flue gas.

Fig. 1 and Fig. 2 show the positive approximating curve applied to the above set of data by using Gaussian function $\phi_{c_i}(\|x-c_i\|)$ and CSRBF $\phi_{\alpha_i}(\|x-c_i\|)$, respectively. Here, the centers c_i are chosen the same as scattered data x_i, the shape parameters c_i and radius α_i should be chosen appropriately.

The TotalErrors for Fig.1 and Fig.2 are 0.04 and 0.63, respectively.

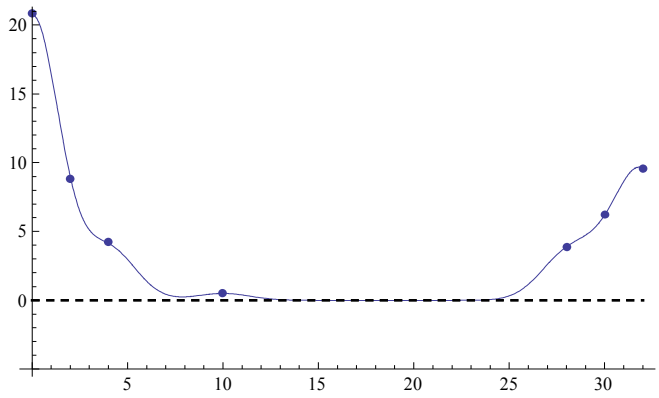


Figure 1 Positive approximating curve for the above data using Gaussian function

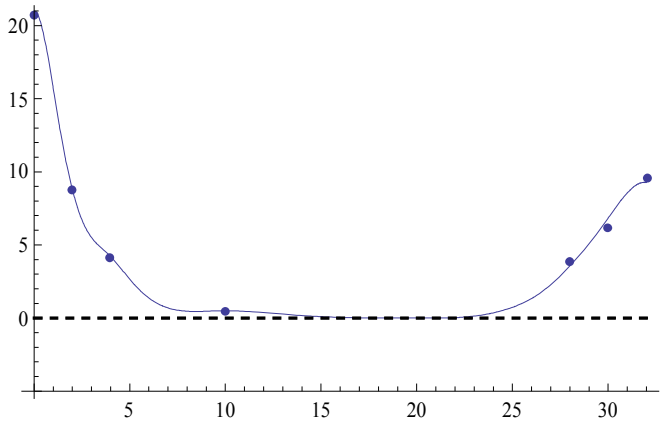


Figure 2 Positive approximating curve for the above data using CSRBF

It is proposed to consider RBF combinations in which all λ_i are positive. That is a much stronger requirement than what is necessary for the result to be everywhere positive. Imposing the strong for the requirement might simplify the discussion, but it might degrade the quality of the end result.

Numerical experiment

In order to illustrate the proposed method for clarity, we limit our example to data in one dimension instead of in higher dimension. Whereas, RBF is only in higher dimension that the outstanding features of RBF really manifest themselves.

We illustrate this technique on 2D and use the following positive function $f(x,y)$ (see Fig. 3):

$$f(x,y)=1.5-0.2e^{-8(x-0.2)^2-6(y-0.5)^2}+0.5e^{-10(x-0.6)^2-8(y-0.4)^2}$$

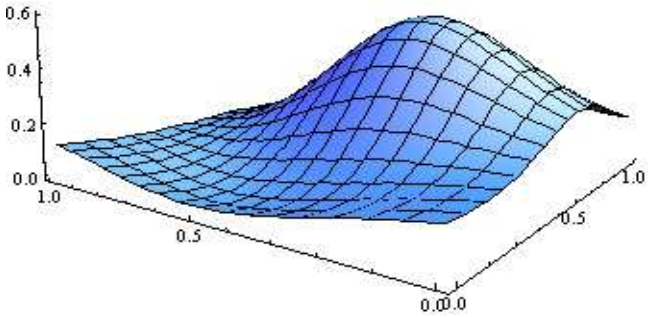


Figure 3 The approximated function

The function was sampled on a randomly generated set of 100 points in the square $[0,1]\times[0,1]$. The positive approximating surface by using CSRBF $\phi_{\alpha_i}(\|x-c_i\|)$ is shown in Fig.4. Here, the centers are found by clustering method.

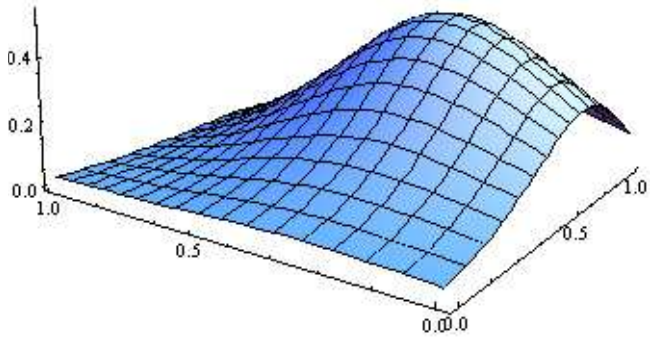


Figure 4 Positive approximating surfaces

It is pointed out that the optimal number of $\phi(\|x-c_i\|)$ and the position of centers c_i are hard to present a flexible approach till now. This fact will remain to be our future work.

Summary

This paper discusses the problem of positive approximation for positive scattered data by using radial basis function. We illustrate the proposed algorithm is flexible from several numerical examples. In fact, we can use B-spline base to deal with positive approximation. Meanwhile, the positive interpolation for positive data using radial basis functions is much more important and remains to be our future work.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 61003138), the Natural Science Foundation of Zhejiang Province (Nos. Y7080068 and Y6090211) and the Foundation of Department of Education of Zhejiang Province (Nos. Y200802999 and Y200907579).

References

[1] J.W. Schmidt and W. Hess: BIT Vol.28 (1988), p. 340

[2] S. Butt and K.W. Brodlie: Computer and Graphics Vol.17(1993), p. 55

[3] A.R.M. Piah, A. Saaban, A.A. Majid: Journal of Fundamental Sciences Vol.2(2006), p. 63

[4] M.Z.Hussaina, M. Sarfrazb: J. Comput. Appl. Math. Vol.218(2008), p. 446

[5] C.G.Zhu, R.H.Wang: Mathematical Problems in Engineering,Vol.2007(2007), p. 11

[6] C. L. Bajaj and G. Xu: Computer Aided Geometric Design Vol.16(1999), p. 557

[7] M.Buhmann, *Radial Basis Function: Theory and Implementations*, (Cambridge University Press, Cambridge 2000.)

[8] Z.M.Wu: Advances in Mathematics Vol.27(1998), p. 202

[9] Z.M.Wu: Advances in Computational Mathematics, Vol.4(1995), p. 283

[10] H.Wendland: Advances in Computational Mathematics, Vol.4(1995), p. 389

[11] M.S.Floater, A.Iske: J.Comput. Appl. Math. Vol.73(1996), p. 65

[12] D.Lazzaro，L.B.Montefusco: J.Comput.Appl. Math. Vol.140 (2002), p. 521

[13] Y.Ohatake, A.Belyaev, H.P.Seidel: Graphical Models, Vol.67(2005), p. 150

[14] K. W. Brodlie, M. R. Asim, K. Unsworth: Computer Graphics Forum Vol.24 (2005), p.809

doi:10.4028/www.scientific.net/AMM.50-51.688

Status of Electromagnetic Compatibility Research and Technology Application in Some Areas

Hong Zhao, Yang Li, Shunli Zheng, Lijun Yu

Dalian Institute Of Product Quality Supervision & Inspection, Dalian, Liaoning, China

zhaohong@dlzjs.org

Keywords: EMC research, EMC technology application, Electromagnetic environment effect

Abstract. Although still in its primary phase up to now, the research about the application of EMC technology in engineering field has become more and more noticeable. In this paper the main subjects for EMC research are introduced, then some hot research topics concerning the EMC technology application are described, involving the application in such areas as the military, industrial electronics, power system and customer electronics. At last the general trend of EMC technology development is discussed, using the case of EMC technology development in the military area for reference.

Introduction

The study of Eletromagnetic Compatibility (EMC) has close correlation with the life and work of the people. Zhang Linchang has defined EMC as this: EMC is a science that studies a variety of electrical or electronic equipments (and living things, in a broad sense), which aim to operate smoothly and coexist with limited resources of space, time and spectrum, and not to degrade under such a condition [1]. EMC is based on the fundamental theories and technologies of electrical and electronic sciences, and concerns the solution of each theoretical or technical problem arising from natural or man-made electromagnetic interference (EMI). The final purpose of EMC research is to ensure the EMC performance of each system or subsystem [2].

The content of EMC study is much comprehensive. Within most of the modern industries there are EMC problems to solve, as in the industries of power, telecommunication, transportation, spaceflight, computery, medical equipments and weapons. In this paper the main research subjects of EMC science are first introduced, then hot research topics within some of the EMC technology application areas are described, involving the war field electromagnetic environment effects [3], EMC control and management for industrial electronics [4,5], electromagnetic susceptibility (EMS) performance of electrical or electronic equipments in nuclear power plants or power systems [6,7], and EMI suppression for home appliances [8]. The trend of EMC technology development will be discussed at last.

Research subjects for EMC Science

In industrial production and human life, the harm of EMI has been recognized more and more widely. For example, it has been forbidden to use mobile communication devices in the civil aircraft cabin, because electromagnetic disturbance from these devices can interfere the normal operation of the antennas, sensors and other sensitive onboard components, threatening the safety of the aircraft.

The American Sugarman, Ellen depicted in his *Warning-the Electricity Around you May be Hazardous to your Health: How to Protect Yourself From Electromagnetic Fields*, that many epidemiology researches have revealed the fact: exposure to a magnetic field over 2 milligausses may add to the risk of catching a cancer. Consequently it is meaningful and necessary to study EMC problems.

The EMC research scope is much extensive, both theoretical problems and technical ones are involved. Generally the research subjects can be sorted into five categories [1].

Sources of electromagnetic disturbances. Only if $\mathrm{d}i/\mathrm{d}t \neq 0$ or $\mathrm{d}v/\mathrm{d}t \neq 0$ under whatever circumstances, electromagnetic noises will be generated, and these noises make up the major part of electromagnetic disturbance. The difficulties in the research about electromagnetic noise sources include the following.

The first is about the generation mechanisms of electromagnetic noises, which are too many and varied.

The second is the inherent conflict between EMC control and the development trends of some technical areas. For example, one trend in computer industry is that the CPU clock frequency goes higher and higher, but the higher the noise frequency goes, the more difficult the noise control gets.

The third difficult problem up to now, is the digital modeling for the physical phenomenon of the generation of electromagnetic noise.

Propagation characteristics of electromagnetic noises. The research method for the propagation characteristics is to build mathematical models according to electromagnetic field theory. The difficulties in this research are as follows.

The first is that in most of the mathematical models, both near field effects and far field effects of electromagnetic waves have to be considered at the same time, because the frequency bands of electromagnetic noises can be very wide.

The second, both the source and the channel of a noise have to be modeled in the same one.

The third is that the boundary conditions of the models are usually complicated because of the need for the models to be practical, and the idealization of the boundary conditions can be difficult in some degree.

Immunity characteristics of electronic devices and systems. The main concerns are about the responses of the noise receivers to electromagnetic noises, and about how to improve the electromagnetic immunity. This kind of research has involved many technical areas such as telecommunication, navigation, radar, broadcasting, television, information technology, remote control and remote sensing. The excessive dependence on the human experience is the difficult point in this research.

Test equipments, test methods and statistical methods. Electromagnetic noises are not marked by normal sinusoidal voltages, but the voltages taking on various shapes and different frequency spectrums, including pulse voltages. Therefore there are strict requirements about the performance of test equipments and test sites, about the test methods, the statistics or evaluation matter, etc.

Since the coefficients or performance parameters of test equipments and test sites may drift from the original with the lapse of time, it is strongly recommended to perform routine checks (also called “spot checks”) of the test equipments and test sites, say, once per week or twice per month. Table 1 is part of a check list on the performance of some anechoic chamber, the judgement standard in this check list is that the difference between spot check value and corresponding original value should be within ± 3 dB.

EMC analysis, prediction and EMC design. EMC design must rely on EMC analysis and prediction. Key points in EMC anslysis and prediction are the building of mathematical models and

the programming for the computation and analysis of EMI that is within a system or between systems.

NO	Frequency [MHz] / Date	30	dev.	100	dev.	200	dev.	300	dev.	400	dev.	500	dev.	600	dev.	700	dev.	800	dev.	900	dev.	1000	dev.	Judge
Original	2007-5-1	63.9	0	97.7	0	90.2	0	84.7	0	92.5	0	91.7	0	90.2	0	89.9	0	89.3	0	85.1	0	82.7	0	
1	2009-1-4	64	0.1	98.2	0.5	90.2	0.05	84.3	-0.5	92.1	-0.4	91.9	0.19	90.6	0.32	90.3	0.45	89.9	0.58	85.3	0.14	82.6	-0.1	OK
2	2009-1-14	64	0.09	98.2	0.47	90	-0.2	84.4	-0.3	91.9	-0.6	91.7	-0	90.4	0.16	90.2	0.31	89.7	0.41	85.3	0.17	82.6	-0.1	OK
3	2009-2-7	63.5	-0.4	97.7	-0	89.6	-0.6	83.6	-1.1	91.9	-0.6	91.8	0.07	90.5	0.26	90	0.11	89.5	0.21	85	-0.1	82.3	-0.4	OK
4	2009-2-24	63.6	-0.3	97.8	0.07	89.8	-0.4	83.8	-0.9	91.6	-0.9	91.6	-0.1	90.2	-0	89.9	0.01	89.2	-0.1	84.2	-0.9	82	-0.7	OK
5	2009-3-2	63	-1	97.3	-0.4	90	-0.2	84.8	0.11	91.6	-0.9	91.4	-0.3	89.8	-0.4	89.5	-0.4	89	-0.3	84.3	-0.8	82.1	-0.6	OK
6	2010-3-6	62.4	-1.5	96	-1.8	88.6	-1.5	85.2	0.5	90.8	-1.8	90.2	-1.5	88.7	-1.5	88.1	-1.8	87.7	-1.6	83.5	-1.6	81	-1.6	OK
7	2010-3-26	62.4	-1.5	96.4	-1.3	89.3	-0.9	85.6	0.89	91.3	-1.2	90.6	-1.1	89.4	-0.8	88.3	-1.6	87.6	-1.7	85.2	0.08	81.7	-1	OK
8	2010-4-8	62.3	-1.6	96.8	-0.9	89.5	-0.7	85.4	0.67	91.4	-1.1	90.8	-0.9	89.1	-1.1	88.6	-1.3	88	-1.3	84.6	-0.5	81.8	-0.9	OK
9	2010-4-15	63.7	-0.2	97.6	-0.1	89.9	-0.3	84.7	-0	92.3	-0.2	91.3	-0.4	89.9	-0.3	89.5	-0.4	88.7	-0.6	84.3	-0.9	82.1	-0.6	OK
10																								

data

Table 1 Check list on the performance of some anechoic chamber

Some hot research topics in EMC technology application

For EMC technology there are many application areas, such as weapon system, industrial electronics, power system and home appliances, see Fig. 1. In each area, there can be several levels of concrete applications, as levels of system, subsystem, interconnect and component, etc. The challenging EMC technology areas for all these applications can be restricted to three main themes: EMC Design, EMC Standardization and EM (electromagnetic) Safety [9]. Hot research topics in some of the application areas of EMC technology are presented as the following.

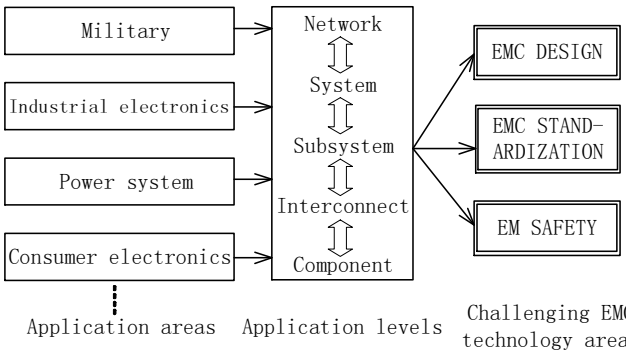


Fig. 1 EMC view on electronic systems [9]

War field electromagnetic environment effects. As military equipment technologies develop, the signal spectrum may be featured by the contribution of various kinds of electrical equipments, and with the utilization plans for electrical equipments changing at times, the characteristics of electromagnetic environment have become more and more complicated.

In a future war, the effective management, control and flexible utilization of electromagnetic environment, will become a key factor in winning the initiative. Listed below will be the key points of the research with respect to war field electromagnetic environment effects [3].

1. Analysis techniques for electromagnetic environment.
2. Design techniques for electromagnetic safety of weapon system.
3. Simulation techniques for (strong)electromagnetic environment.
4. Experimention, test and evaluation of electromagnetic environment effects.
5. Techniques of control and management.

EMC control and management for industrial electronics. The EMC research for electrical or electronic devices is still in its primary phase up to now. The research in the last years has been focused on such issues as techniques for modeling and suppression of EMI from power converter and motor transmission, parasitic effects of EMI filters, layout optimization for PCBs and numerical analysis techniques for EMI [4].

For the major electrical systems abroad such as the power, control and communication installations for the railroads, subways and airports, strict EMC control and management plans are established in the initial phase of the design. The plans generally contain the following elements [5].

1. EMI analyses.
2. Determination of EMC standards to meet.
3. Determination of technical indicators of EMC design.
4. Establishment of EMC design principles.
5. EMC management and organizing structure.
6. EMC implementing plan.
7. EMC test plan.

EMS performance of the equipments in power systems. Nowadays for the digital instrument and control systems in nuclear power plants, test of the immunity for EMI has become a focus for the attention of the engineers. To improve EMC performance of the entire system, strengthened design against EMI is necessary for such components as power supply, communication network, input/output module and CPU module, and some measures are needed for the correction of electromagnetic environment of the system. The measures may be proper grounding, normative wiring and dust proofing, etc. [6]

The full-scale digitization of protection devices in power systems has become an irreversible trend. How to improve the capacity of anti-electromagnetic-interference and guarantee the stable operation of power system, has always been a hot research topic for microprocessor-based protection technology, and the anti-interference capacity respectively for the hardware and software of each protection device can be taken into consideration [7].

EMI suppression for home appliances. According to operation principle and configuration, household electromotors can be sorted into commutator-motors and induction-motors. Home appliances with electromotors of the former kind can be characterized by sources of electromagnetic disturbances badly severe. As the anti-disturbance measures, proper filtering, shielding and grounding should be considered in the early design phase [8].

Trend of EMC technology development

To meet the market needs for products that are safe to use, reliable in quality and perfect in EMC performance, new methods and new means should be adopted to suppress EMI from electrical components, interconnects and (sub)systems, meanwhile improve their immunity for EMI. The future electromagnetic environment needs to be controlled by legislation and standardization with new test methods, frequency bands and limits [9]. Zhao Gang has discussed the development directions of EMC technology for weapon equipments in the information age [10], which can be of great use for reference by other EMC application areas, and can be generalized as below.

1. Application of “prediction analysis method” will be featured in EMC design and control for electrical equipments in this century.

2. Techniques for the verification and evaluation of and shielding from electromagnetic environment effect are hot research topics in current days.

3. Sound management and utilization of radio frequency spectrum resources are becoming

critical points for EMC research in the information age.

4. The exploration and application of new techniques, new materials and new devices are the trends of EMC technology development.

References

[1] Zhang Linchang. Developing electromagnetic compatibility undertaking in our country[J]. Transactions of China electrotechnical society, 2005, 20(2): 23~28

[2] Ma Weiming. Several problems in the research for EMC prediction of independent system[J]. Electromagnetic compatibility technology, 2005, (2): 1~4

[3] Hou dongyun. Electromagnetic environment status and EME/EMC technology development[J]. Electromagnetic compatibility technology, 2007, (4): 1-5

[4] Qian Zhaoming, Chen Henglin. State of art of electromagnetic compatibility research on power electronic equipment[J]. Transactions of China electrotechnical society, 2007, 22(7): 1~11

[5] Li Erping, Chen Degui. EMC control and management of intelligent electrical switchgear and controlgear[J]. Low voltage apparatus, 2001, (2): 16~18

[6] Huang Wenjun, Yu Haoyang, Ao Chunbo. EMC test and design of I&C system in nuclear power plants[J]. Nuclear power engineering, 2008, 29(3): 85~88,107

[7] Peng Honghai, Zhou Youqing, Wang Hongtao, etc. Anti-interference technology for microprocessor-based protection[J]. High voltage engineering, 2007, 33(10): 49~53

[8] Liu Ping, Sha Fei. Improving the electromagnetic compatibility of home appliances with commutator electromotor[J]. Journal of northern jiaotong university, 2002, 26(6): 64~68

[9] Van Doorn M. EMC technology roadmapping: a long-term stragegy[C]. In: Proc. of IEEE International Symposium on EMC'06, 2006: 156~159

[10] Zhao Gang. Development trend of weaponry furnishment EMC technology in the informational age[J]. Ship electronic engineering, 2007, 27(1): 20~22,36

©(2014) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.693

Blind JPEG Steganalysis Using Features Derived From Multi-domain

Wenqiong Yu[1, a], Zhuo Li [2, b], Lingdi Ping[2]

[1]Mathematics and Computer Science, Sanming University, Fujian Sanming 365004,China

[2]College of Computer Science and Technology, Zhejiang University, Zhejiang Hangzhou 310027,China

[a]email: ywq-2@163.com , [b]email: lizhuo84@gmail.com

Keywords: Steganalysis; Co-occurrence matrix; Feature vector; Multi-domain; LIBSVM.

Abstract: In this paper, a new blind steganalytic scheme is proposed to effectively detect JPEG steganography. In DCT domain, the co-occurrence matrices are used to capture the intra-block and inter-block correlations among the quantized coefficients. In spatial domain, we estimate the invariant components of the decompressed raw pixels between the cover and stego images by wiping selected intermediate frequency AC coefficients, and collect eighteen statistics in total. Combining the DCT and spatial features leads to be more effective for classification. The experimental results have demonstrated that the proposed method is stable, and outperforms the recently reported steganalysis in attacking the advanced JPEG steganography.

Introduction

Steganalysis focuses on discovering the presence of hidden messages which are embedded in the cover medium such as digital images imperceptibly. Since JPEG is the most commonly used image format on the Internet, researches in the field of JPEG steganalysis has attracted more and more attention now [1-5, 7].

Fu et al. [1] presented a universal JPEG steganalytic scheme in which 200 features were calculated directly from the quantized DCT coefficients. The Markov empirical transition matrices were used to exploit the correlations between DCT coefficients in both intra-block and row scanning inter-block sense. In [2], Shi et al. proposed to calculate the difference JPEG 2-D arrays of quantized block DCT coefficients along horizontal, vertical and diagonal directions firstly, and then apply the Markov process to model these difference 2-D arrays as features. Chen et al. [3] extended [2] and presented another scheme. The authors use the difference mode 2-D arrays to construct inter-block features. Combined with the intra-block features collected in [2], a 486- dimensional feature vector was obtained.

In this paper, we propose a novel blind steganalytic scheme to effectively detect the advanced JPEG steganography. Features are calculated both from the DCT domain and the decompressed spatial domain. Firstly, the second order statistics, i.e., co-occurrence matrices, are used to capture both the intra- and inter-block correlations between the quantized DCT coefficients. Then, an estimation technique is developed to remove the gross representation of the stego image in the spatial domain. In this way, minor embedding changes to the stego image in the spatial domain can be easily explored. The experimental results are presented to demonstrate that the proposed scheme outperforms the other blind JPEG steganalysis recently reported. The rest of this paper is organized as follows. In the next section, we explain how to generate the intra- and inter-block features in the DCT domain. In section 3, we describe the feature construction procedure in spatial domain, and then the details of experiments are illustrated in section 4. At last, the paper is concluded in Section 5.

Features in DCT Domain

A co-occurrence matrix is a matrix that is defined to be the distribution of co-occurring pixel values at a given offset over an image. Given a grayscale image, the co-occurrence matrix can reflect the

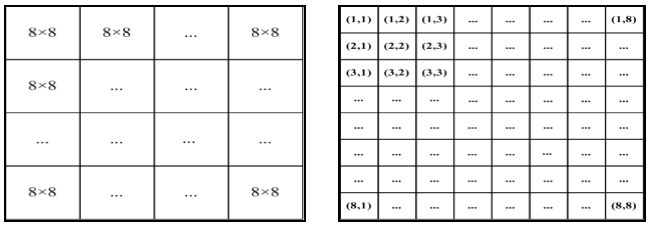
distribution of a pair of pixels. Mathematically, suppose a grayscale image I has the dimension of M×N and the gray-level of Ng, a gray-level co-occurrence matrix C is defined as

$$C(i,j)=\#\{(x,y),(x+\Delta x,y+\Delta y)\in(M\times N)\mid I(x,y)=i,I(x+\Delta x,y+\Delta y)=j\} \quad (1)$$

where I(x,y) denotes the pixel value at location (x,y) in image I, and #{X} indicates the number of elements in the set X.

Intra-block Features

Without loss of generality, we suppose the test JPEG image has the dimension of M×N. Then the JPEG coefficient matrix can be divided into N_B=[M/8]×[N/8] non-overlapping DCT blocks (as Fig. 1(a) shows). Each block has a dimension of 8×8 (as Fig. 1(b) shows). In order to describe the statistics conveniently, we use $dct_{r,c}(i, j)$ to represent the quantized DCT coefficient value at location (i, j) in an 8×8 DCT block, where the DCT block is in row r and column c in the JPEG coefficient matrix. Obviously $1 \leqq i,j \leqq 8$, $1 \leqq r \leqq [M/8]$ and $1 \leqq c \leqq [N/8]$.



(a) JPEG coefficient matrix (b) An 8×8 DCT block

Fig.1. Quantized DCT coefficients.

The intra-block correlation states that although the quantized DCT coefficients are decorrelated by the DCT transform to some extent, two coefficients within one DCT block are probably not independent. A coefficient still has correlation with its neighbours in the same block. As a result, we calculate the co-occurrence matrices of the neighbouring coefficient pairs along horizontal, diagonal, vertical, and minor diagonal directions and the offset is 1, for each 8×8 DCT block. And then we take the arithmetic average of co-occurrence matrices generated from all N_B blocks as the intra-block statistics. The co-occurrence matrix in horizontal can be calculated by

$$C_1(m,n)=\frac{1}{N_B}\sum_{r=1}^{[M/8]}\sum_{c=1}^{[N/8]}\sum_{i=1}^{i=8}\sum_{j=1}^{j=7}\left(\varphi(dct_{r,c}(i,j),m)\bullet\varphi(dct_{r,c}(i,j+1),n)\right) \quad , \quad \varphi(x,y)=\begin{cases}1, & if\ x=y\\ 0, & elsewise\end{cases} \quad (2)$$

Similar to C_1, the other three co-occurrence matrices in diagonal, vertical and minor diagonal directions are calculated as follows.

$$C_2(m,n)=\frac{1}{N_B}\sum_{r=1}^{[M/8]}\sum_{c=1}^{[N/8]}\sum_{i=1}^{i=7}\sum_{j=1}^{j=7}\left(\varphi(dct_{r,c}(i,j),m)\bullet\varphi(dct_{r,c}(i+1,j+1),n)\right) \quad (3)$$

$$C_3(m,n)=\frac{1}{N_B}\sum_{r=1}^{[M/8]}\sum_{c=1}^{[N/8]}\sum_{i=1}^{i=7}\sum_{j=1}^{j=8}\left(\varphi(dct_{r,c}(i,j),m)\bullet\varphi(dct_{r,c}(i+1,j),n)\right) \quad (4)$$

$$C_4(m,n)=\frac{1}{N_B}\sum_{r=1}^{[M/8]}\sum_{c=1}^{[N/8]}\sum_{i=1}^{i=7}\sum_{j=1}^{j=7}\left(\varphi(dct_{r,c}(i+1,j),m)\bullet\varphi(dct_{r,c}(i,j+1),n)\right) \quad (5)$$

To reduce the computational complexity, we take the arithmetic average of above four co-occurrence matrices (C_1~C_4) as the intra-block co-occurrence matrix C_{intra}.

$$C_{intra}=\frac{1}{4}(C_1+C_2+C_3+C_4) \quad (6)$$

Inter-block Features

A quantized DCT coefficient in one block has somewhat strong correlation with the coefficient of the same position in one of the neighbouring blocks [6]. As Fig. 1(a) shows, the JPEG coefficient matrix is divided into N_B non-overlapping 8×8 DCT blocks. For a DCT coefficient in one block B and the coefficient of the same position in one of block B's neighbouring blocks in horizontal, diagonal, vertical and minor diagonal directions, we calculate the co-occurrence matrices of this coefficients pair. As a result, we will get 64 co-occurrence matrices in a direction because of there exist 64 coefficients in one DCT block. And then, we take the arithmetic average of the 64 co-occurrence matrices as the inter-block statistics. The inter-block co-occurrence matrices associated with the horizontal, diagonal, vertical and minor diagonal are given by

$$C_5(m,n)=\frac{1}{64}\sum_{i,j=1}^{8}\sum_{r=1}^{[M/8]}\sum_{c=1}^{[N/8]-1}\left(\varphi(dct_{r,c}(i,j),m)\bullet\varphi(dct_{r,c+1}(i,j),n)\right) \quad (7)$$

$$C_6(m,n)=\frac{1}{64}\sum_{i,j=1}^{8}\sum_{r=1}^{[M/8]-1}\sum_{c=1}^{[N/8]-1}\left(\varphi(dct_{r,c}(i,j),m)\bullet\varphi(dct_{r+1,c+1}(i,j),n)\right) \quad (8)$$

$$C_7(m,n)=\frac{1}{64}\sum_{i,j=1}^{8}\sum_{r=1}^{[M/8]-1}\sum_{c=1}^{[N/8]}\left(\varphi(dct_{r,c}(i,j),m)\bullet\varphi(dct_{r+1,c}(i,j),n)\right) \quad (9)$$

$$C_8(m,n)=\frac{1}{64}\sum_{i,j=1}^{8}\sum_{r=1}^{[M/8]-1}\sum_{c=1}^{[N/8]-1}\varphi(dct_{r,c+1}(i,j),m)\bullet\varphi(dct_{r+1,c}(i,j),n) \quad (10)$$

Similar to the intra-block statistics, the average of above four matrices (C_5~C_8) is the inter-block co-occurrence matrix C_{inter} in our scheme. Since the elements in C_{intra} and C_{inter} have a very large range, we apply a threshold T=4 introduced in [1-3] to reduce the feature dimension and computational complexity. Let C_{intra} and C_{inter} denote the intra-block and inter-block co-occurrence matrices, respectively, the calibrated features are calculated as follows.

$$C_{intra}^{(c)}=C_{intra}(I_{ori})-C_{intra}(I_{cal}) \quad \text{and} \quad C_{inter}^{(c)}=C_{inter}(I_{ori})-C_{inter}(I_{cal}) \quad (11)$$

where I_{ori} is the testing image and I_{cal} is the calibrated image. In summary, there are 162 features extracted in the DCT domain for each JPEG image.

Features in Spatial Domain

The basis of steganalysis is that there exists difference between the images before and after data hiding, and the difference is detectable [7]. As we know, minor changes in the DCT domain result in amplifying changes in the spatial domain. Hence, we try to find some embedding changes occurred in spatial domain.

Image Calibration

As we know, most JPEG steganographic schemes manipulate the quantized DCT coefficients to embed messages. In fact, they mainly embed bits using the non-zero AC coefficients. As we observe the whole image database (illustrated in section 4), nearly 95% elements in average fall into the interval [-31, 0) and (0, 31]. So we can surmise that embedding changes mainly occur to the non-zero DCT coefficients in range of [-31, 0) and (0, 31], and the other non-zero AC coefficients rarely alert. Then if we wipe these variable coefficients in range of [-31, 0) and (0, 31] in the DCT domain, the remaining coefficients can be thought to be the invariant component during embedding messages. Decompress the test JPEG image and the wiped image into spatial domain, and then the difference

between them will be more sensitive to the embedding changes in spatial domain caused by embedded data. We call the difference as variable statistics.

High Order Differentials

In this subsection, we use high order differentials to capture the correlations between the elements in the variable statistics. Suppose that the variable statistics, denoted as B, has the size of M×N, and B(m, n) denotes the value at location (m, n). The 1-st order partial differential at location (m, n) in column direction can be calculated as

$$P_1^{(c)}(m,n) = B(m,n+1) - B(m,n) \tag{12}$$

and the partial differential in row direction is

$$P_1^{(r)}(m,n) = B(m+1,n) - B(m,n) \tag{13}$$

.

Hence, the 1-st order total differential at location (m, n) can be written as below

$$D_1(m,n) = \sqrt{\left(P_1^{(r)}(m,n)\right)^2 + \left(P_1^{(c)}(m,n)\right)^2} \tag{14}$$

In order to reduce the computational complexity, in our scheme we calculate it using

$$D_1(m,n) = \left|P_1^{(r)}(m,n)\right| + \left|P_1^{(c)}(m,n)\right| \tag{15}$$

In the same way, we can obtain the 2-nd and 3-rd order partial differentials and total differentials of B. Without loss of generality, B itself can be considered as the 0-th order total differential thereafter. In summary, we calculate six objects in total, including the 0-th, 1-st and 2-nd order total differentials and the 1-st, 2-nd and 3-rd order partial differentials. For each object, we use the histogram and co-occurrence matrix to capture the correlations between the elements.

Take the 0-th order total differential B for example, the histogram can be calculated as

$$H(v) = \sum_{i=1}^{M}\sum_{j=1}^{N}\varphi(B(i,j),v) \tag{16}$$

By replacing B(i, j) with D_1(i, j) and modifying the ranges of i, j, we can get the histogram of the 1-st order total differential. Similarly, we can calculate the histograms of other objects.

The neighboring co-occurrence matrices in B can be calculated as belows. For each object, we calculate the co-occurrence matrices in horizontal and vertical individually.

$$C^{(c)}(m,n) = \sum_{i=1}^{M}\sum_{j=1}^{N-1}\left(\varphi(B(i,j),m)\bullet\varphi(B(i,j+1),n)\right) \tag{17}$$

$$C^{(r)}(m,n) = \sum_{i=1}^{M-1}\sum_{j=1}^{N}\left(\varphi(B(i+1,j),m)\bullet\varphi(B(i,j),n)\right) \tag{18}$$

In summary, there are 6 histograms and 12 co-occurrence matrices obtained based on the high order differentials.

High order statistical moments

We use high order statistical moments to calculate features form the histograms and co-occurrence matrices obtained in section 3.2. In particularly, the 1-dimensional Discrete Fourier Transform of a histogram H can be referred to as the histogram characteristic function, which can be denoted as $f_k \triangleq \left|DFT_1(H[d])\right|$. Then the first and second order statistical moments can be calculated using Eq.19.

Since DFT is central symmetric, we only need to calculate in range of [1, N/2] for a DFT sequence with length N.

$$M_1^n = \sum_{k=1}^{N/2} k^n \bullet |f_k| \Big/ \sum_{k=1}^{N/2} |f_k| \tag{19}$$

where n = (1, 2), $|f_k|$ represents the magnitude corresponding to the frequency k, and N is the maximum of all magnitudes.

For a co-occurrence matrix C, the 2-dimensional Discrete Fourier Transform can be considered as its characteristic function denoted as $f_{r,c} \triangleq |DFT_2(M[d_1, d_2])|$, and the high order statistical moments in each dimension can be calculated using Eq.20.

$$M_2^n = \sum_{r=1}^{N_r/2} \sum_{c=1}^{N_c/2} (r^n, c^n) \bullet |f_{r,c}| \Big/ \sum_{r=1}^{N_r/2} \sum_{c=1}^{N_c/2} |f_{r,c}| \tag{20}$$

where n = (1, 2), *Nr* and *Nc* represent the maximum magnitudes in the direction r and c, respectively.

In summary, for each histogram, two features will be calculated by Eq.19. And for each matrix, four features will be calculated by Eq.20. As a result, there are 60 features extracted in the spatial domain.

Experimental Results

Image Database and JPEG Steganography Tools

Image database #1 consists of 1610 raw colour images of different dimensions downloaded from Greenspun [8]. **Image database #2** consists of 1096 sample images (768×512 or 512×768 pixels) included in the CORELDRAW Version 10.0 software CD#3 [9]. **Image database #3** includes 1338 uncompressed colour images (512×384 or 384×512 pixels) downloaded from UCID image database [10].

To evaluate the performance of our scheme, we focus on attacking five widely used JPEG steganography, i.e., Jphide[11], Outguess[12], Steghide[13], MB1[14] and MB2[15]. We use three kinds of embedding rates, 0.05, 0.10 and 0.20bpnc(bits per non-zero DCT coefficients) for each steganography. The embedded messages are random bit streams generated using Matlab2007b, we use LibSVM[16] to construct the classifiers.

Performance Comparisons with existing JPEG blind steganalysis

We use image database #1 in this experiment. In this subsection, experimental results are presented to demonstrate the performances of our proposed scheme. We compare the performance of our method with four effective universal algorithms FU200 [1], SHI324 [2] and CHEN486 [3]. For each steganography, we randomly select 500 cover images and the corresponding 500 stego images as the training set, and the rest images are used for testing. The experimental results reported in Table 1 are the arithmetic average of 20 times of tests. *TPR* and *TNR* represent the true positive rate and the true negative rate, respectively. *AR* is the final accuracy rate, which is calculated by *AR=(TPR+TNR)/2*. As we can see, our proposed method provides much better accuracy rate than that of FU200, SHI324 and CHEN486 in detecting the JPEG steganography. Especially when then embedding rate is low (0.05bpnc), the performance of our proposed method is prominent.

Table 1 Performance comparisons with recent steganalysis schemes (in the unit of %)

	bpnc	FU200			SHI324			CHEN486			Proposed222		
		TNR	*TPR*	*AR*	*TNR*	*TPR*	*AR*	*TNR*	*TPR*	*AR*	*TNR*	*TPR*	*AR*
JPHS	0.05	85.50	14.50	**50.00**	36.67	63.60	**50.14**	37.66	62.52	**50.09**	72.52	51.80	**62.16**
	0.10	81.71	26.22	**53.96**	40.63	63.87	**52.25**	34.14	71.89	**53.02**	73.42	61.17	**67.30**
	0.20	83.06	75.68	**79.37**	71.17	80.54	**75.86**	66.13	74.32	**70.23**	93.15	89.10	**91.13**
OG2	0.05	85.68	25.59	**55.63**	65.05	60.36	**62.70**	85.05	77.84	**81.44**	98.11	96.67	**97.39**
	0.10	88.83	57.12	**72.97**	85.05	92.07	**88.56**	97.12	97.39	**97.25**	98.74	99.37	**99.05**
	0.20	98.02	96.94	**97.48**	98.38	99.46	**98.92**	99.37	99.37	**99.37**	99.46	98.65	**99.05**
SH	0.05	85.14	20.63	**52.88**	62.88	49.55	**56.22**	67.66	63.42	**65.54**	74.95	79.46	**77.21**
	0.10	85.50	22.61	**54.05**	58.29	60.18	**59.23**	76.13	72.25	**74.19**	86.13	86.76	**86.44**
	0.20	83.42	34.23	**58.83**	66.76	68.47	**67.61**	90.18	84.59	**87.39**	93.06	95.68	**94.37**
MB1	0.05	84.95	26.85	**55.90**	72.79	43.15	**57.97**	82.70	27.48	**55.09**	80.27	83.51	**81.89**
	0.10	90.09	57.75	**73.92**	83.15	71.98	**77.57**	82.16	71.08	**76.62**	96.85	95.14	**95.99**
	0.20	93.42	98.11	**95.77**	96.76	98.11	**97.43**	96.13	99.28	**97.70**	98.11	99.10	**98.60**
MB2	0.05	81.62	31.53	**56.58**	71.17	52.25	**61.71**	76.67	41.17	**58.92**	85.41	84.59	**85.00**
	0.10	89.10	48.56	**68.83**	81.62	86.94	**84.28**	77.03	88.47	**82.75**	96.13	96.22	**96.17**
	0.20	92.43	95.68	**94.05**	98.02	98.92	**98.47**	97.66	99.64	**98.65**	98.83	98.83	**98.83**

Experiments on different image databases

To give more convictive results, we take the same experiments on image database #2 and #3, in which the JPEG images are with different resolutions, different quality factors, and different sources. For each image database, we randomly select 1/2 of images to train and the rest for testing. Experimental results are reported in Table 2, where the results are the arithmetic average of 20 times of tests. The rightmost three columns list the average, standard deviation and variance of final accuracy rates on three different image databases. As we can see, although the testing JPEG images are with different resolutions, different quality factors and different source, the proposed method can provide stable detection accuracy in detecting all these five JPEG steganographic schemes: JPHS, OG2, SH, MB1 and MB2. Especially when the embedding rate is high, such as 0.20bpp, the performance of our method is very stable.

Table 2 Experiments on three different image databases (in the unit of %)

	bpnc	Image database #1			Image database #2			Image database #3			Average	Standard Deviation	Variance
		TNR	*TPR*	*AR*	*TNR*	*TPR*	*AR*	*TNR*	*TPR*	*AR*	*AR*	*AR*	*AR*
JPHS	0.05	72.52	51.80	**62.16**	81.42	54.97	**68.20**	76.68	53.75	**65.21**	**65.19**	3.02	9.12
	0.10	73.42	61.17	**67.30**	80.73	65.10	**72.91**	73.26	57.97	**65.62**	**68.61**	3.82	14.57
	0.20	93.15	89.10	**91.13**	92.31	91.48	**91.89**	90.01	87.77	**88.89**	**90.64**	1.56	2.43
OG2	0.05	98.11	96.67	**97.39**	98.51	97.89	**98.20**	92.09	91.50	**91.79**	**95.79**	3.49	12.18
	0.10	98.74	99.37	**99.05**	100.00	99.31	**99.65**	99.41	98.97	**99.19**	**99.30**	0.31	0.10
	0.20	99.46	98.65	**99.05**	100.00	99.71	**99.85**	100.00	99.68	**99.84**	**99.58**	0.46	0.21
SH	0.05	74.95	79.46	**77.21**	78.94	79.10	**79.02**	70.84	76.53	**73.69**	**76.64**	2.71	7.36
	0.10	86.13	86.76	**86.44**	89.93	89.85	**89.89**	88.90	88.63	**88.77**	**88.37**	1.76	3.10
	0.20	93.06	95.68	**94.37**	93.12	96.89	**95.01**	93.88	94.90	**94.39**	**94.59**	0.36	0.13
MB1	0.05	80.27	83.51	**81.89**	86.34	88.56	**87.45**	79.73	78.12	**78.93**	**82.76**	4.33	18.71
	0.10	96.85	95.14	**95.99**	98.14	97.05	**97.60**	94.70	92.14	**93.42**	**95.67**	2.11	4.44
	0.20	98.11	99.10	**98.60**	100.00	99.60	**99.80**	99.73	98.73	**99.23**	**99.21**	0.60	0.36
MB2	0.05	85.41	84.59	**85.00**	89.40	89.76	**89.58**	83.47	81.69	**82.58**	**85.72**	3.56	12.64
	0.10	96.13	96.22	**96.17**	99.24	97.41	**98.32**	97.41	94.85	**96.13**	**96.87**	1.25	1.57
	0.20	98.83	98.83	**98.83**	100.00	99.31	**99.65**	99.76	99.71	**99.73**	**99.40**	0.50	0.25

Conclusions

In this paper, we have proposed a blind steganalysis scheme to effectively detect JPEG steganography. Not only the embedding changes occurred to the DCT coefficients but also the changes occurred in the spatial domain are analysed and captured. Since many advanced JPEG steganography can

preserve margin histograms of cover image during embedding, but cannot preserve such high-order statistics, we use the second order statistics, i.e., co-occurrence matrices, to capture the correlations between the coefficients in DCT domain. Besides, a new method to estimate the invariant components in spatial domain is introduced. In this way, the variable components probably changed during data embedding are calculated and more sensitive to the embedding changes. The experimental results bear out the effectiveness of the proposed method in detect advanced JPEG steganography compare with some steganalysis schemes reported recently.

Acknowledgments

This work was supported by National Technology Support Projects (2008BA21B03); Major Science Technology Projects of Zhejiang Province (2007C11088); Science and Technology Projects Plans of Zhejiang Province (2008C21077); Scientific Research Fund of Sanming University (2009B0901/G).

References

[1] D.D. Fu, Y.Q.Shi, D.K. Zou and G.R. Xuan: JPEG steganalysis using empirical transition matrix in block DCT domain. *IEEE 8th International Workshop on Multimedia Signal Processing* (2006), p.310-313.
[2] Y.Q. Shi, C.H. Chen and W. Chen: A Markov process based approach to effective attacking JPEG steganography. In: *Proceedings of the 8th Information Hiding Workshop*(2007), p.249-264.
[3] C.H. Chen, Y.Q. Shi: JPEG image steganalysis utilizing both intrablock and interblock correlations. *IEEE International Symposium on Circuits and Systems* (2008), p.3029-3032.
[4] X.Y. Luo, D.S. Wang and P. Wang: A Review on Blind Detection for Image Steganography. *Signal Processing*, vol.88 (2008), p.2138-2157.
[5] X.Y. Luo, D. Wang, W. Hu and F. Liu: Blind Detection for Image Steganography: A System Framework and Implementation. *International Journal of Innovative Computing, Information and Control*, vol.5, (2009), p.433-442.
[6] C. Tu and T.D. Tran: Context based entropy coding of block transform coefficients for image compression. *IEEE Transaction on Image Processing*, vol.11, no.11, (2002), p.1271-1283.
[7] F.J. Huang and J.W. Huang: Calibration based universal JPEG steganalysis. *Science in China Series F-Information Sciences*, vol.52, no.2, (2009), p.260-268.
[8] Greenspun Image Library. Available: http://philip.greenspun.com.
[9] CorelDraw Image Library. Available: http://www.corel.com.
[10] G. Schaefer and M. Stich: UCID - An Uncompressed Colour Image Database, Proc. SPIE, *Storage and Retrieval Methods and Applications for Multimedia* (2004), p. 472-480.
[11] JPHide&Seek. Available: http://linux01.gwdg.de/~alatham/stego.html.
[12] Outguess. Available: http://www.outguess.org/
[13] Steghide. Available: http://steghide.sourceforge.net/
[14] P.Sallee. Model based steganography. In: *Proceedings of the 2nd International Workshop on Digital Watermarking*, 2004, pp.154–167
[15] P. Sallee. Model-based methods for steganography and steganalysis. *International Journal of Image and Graphics*, 2005, pp.167-189
[16] C.C.Chang and C.J.Lin. LIBSVM: A Library for Support Vector Machines. (2001) http://www.csie.ntu.edu.tw/~cjlin/libsvm

doi:10.4028/www.scientific.net/AMM.50-51.700

Chinese Text Classification with a KNN Classifier Using an Adjusted Feature Weighting Method

Weiran Lin[1,a], Zhihui Wu[2], Lichao Feng[3] and Waibin Huang[1,b]

[1] School of Computer and Computing Science, Zhejiang University City College, Hangzhou 310015, China

[2] Applied Science College, Harbin University of Science and Technology, Harbin, China

[3] College of Science, Hebei Polytechnic University, Tangshan, China

[a] Linweiran1982@163.com,[b] (Corresponding author)waibin@163.com

Keywords: KNN; feature weight; text classification

Abstract. KNN algorithm is used for Chinese text classification in this paper. First, TF-IDF is chosen as the feature weighting method. To the characteristics of corpus used in this paper, TF-IDF is adjusted to a new method. At last, experimental result shows the accuracy of KNN text classifier can be improved with the adjusted feature weighting method.

1 Introduction

Since late 1990s, the explosive growth of Internet resulted in a huge quantity of documents available on-line. Technologies for efficient management of these documents are being developed continually. One of representative tasks for efficient document management is text categorization, called also classification: given a set of training examples assigned each one to some categories, to assign new documents to asuitable category.

There are several classification schemes that can be potentially used for text categorization. One of these existing schemes work well in the text categorization is KNN classification. KNN classification is an instance-based learning algorithm that has shown to be very effective for a variety of problem domains. This algorithm has also been used in text classification.A k-nearest-neighor classification based on feature extraction and a vector proximity or distance. KNN classification is a kind of memory-based learning in which every training instance is stored along with its category. To classify an object, the k nearest training examples to the object being classified are found. Each of the k nearest neighbors votes for its training category. The resulting classification scores are the result of voting. It is possible to weight the votes by proximity[1,2].

In this paper, LJParser 1.0.0.1 is used to segment Chinese into words.Then the feature extractor based on document frequency (DF) removes words which provides less information for text categorization[3,4,5].First,TF-IDF[6] is chosen as the feature weighting method. To the characteristics of corpus provided by Songbo Tan[7], TF-IDF is adjusted to a new method. At last, experimental result shows the accuracy of KNN text classifier can be improved with the adjusted feature weighting method.

2 Chinese Text Classification Based on KNN

The first step is to segment Chinese into words.Traditional text classification use VSM (Vector Space Model), which views documents as vectors in high dimensional spaces, to represent documents. In this vector model, each document is a vector and its element corresponds to words in the whole document set. Unlike western languages, Chinese is written without spaces between words. To build a vector, it is first necessary to determine word boundaries. In this paper, LJParser 1.0.0.1 is used to segment Chinese into words.

The second step in building a classifier is to build a vector which consistes of all the relevant words found in the train set of documents. The relevant words can be defined to be those which occurred

more than 5 times but less than 1790 times in the training set which contains 1800 documents. This would eliminate the very rare words that occurred and also the very common words that occur in almost every text document. these words would provide little information and it would be best to eliminate them. Then, the whole training set can be viewed as a matrix where each row is a document and its columns are the relevant words. The values in the matrix can be binary, 1 for presence of the word and 0 for absence of the word. They can also be the within-document word frequency (TF), Inverse Document Frequency (IDF), or TF-IDF which is the combination of TF and IDF Document Frequency (IDF), or TF-IDF which is the combination of TF and IDF[6].

The last step is to train the classifier. KNN classification is one of the most fundamental and simple classification methods.To classify an unknown document d_0 , the KNN classifier ranks the document's neighbors among the training documents, and use the class labels of k most similarity neighbors to predict the class of the input document. To measure the similarity efficiently, we make use of Euclidean distance as follows:

$$Dis(\vec{d}_1,\vec{d}_2)=\sqrt{\sum_{t=1}^{V}(d_{1t}-d_{2t})^2} \quad \text{.....(1)}$$

Then

$$Sim(\vec{d}_1,\vec{d}_2)=\frac{1}{1+Dis(\vec{d}_1,\vec{d}_2)} \quad \text{.....(2)}$$

where V denotes the dimension size of document vector $\vec{d}_1$, $\vec{d}_2$.

The classes of these neighbors are weighted using the similarity of each neighbor to d_0 as follows:

$$score(\vec{d}_0,C_i)=\sum_{\vec{d}_j\in KNN(\vec{d}_0)} Sim(\vec{d}_0,\vec{d}_j)\delta(\vec{d}_j,C_i) \quad (3)$$

where $KNN(\vec{d}_0)$ indicates the set of K-nearest neighbours of document $\vec{d}_0$, $\delta(\vec{d}_j,C_i)$ stands for the classification for document $\vec{d}_j$ with respect to class C_i ,that is,

$$\delta(\vec{d}_j,C_i)=\begin{cases}1 & \vec{d}_j\in C_i\\ 0 & \vec{d}_j\notin C_i\end{cases} \quad (4)$$

Consequently, the decision rule in KNN classification can be written as:

$$\begin{aligned}C&=\arg\max_{C_i}\left(score(\vec{d}_0,C_i)\right)\\&=\arg\max_{C_i}\left(\sum_{\vec{d}_j\in KNN(\vec{d}_0)} Sim(\vec{d}_j,\vec{d}_0)\delta(\vec{d}_j,C_i)\right)\end{aligned} \quad (5)$$

3 Experiment Design and Experimental Result

Algorithms are coded in Java and running on a Pentium-4 machine with single 2.0 GHz CPU. LJParser 1.0.0.1 is used to segment Chinese into words.

3.1 Corpora and Cross-Validation

The corpora with positive and negative evaluation of the hotel contains 2,000 documents provided by Songbo Tan [7].It includes:

1000 negative documents: *neg0,neg1,neg2……neg999*

1000 positive documents: *pos0,pos1,pos2……pos999*.

In this experiments, *ten-fold* Cross-Validation is adopted. The fold(1-9) is defined as follows:

Fold-i:

pos1i,pos2i……pos9i,pos1i0……pos1i9,pos2i0……pos9i0,pos9i1,pos9i2……pos9i9

neg1i,neg2i……neg9i,neg1i0……neg1i9,neg2i0……neg9i0,neg9i1,neg9i2……neg9i9.

Fold-i will be used for test and the remaining parts for train.

3.2 Feature Selection and Weighting Method

This paper, Document Frequency is employed as feature selection method. The features can be defined to be words which occurred less than 1798 times in the training set which contains 1800 documents.

TF-IDF is utilized as feature weighting method. The formula for calculating TF-IDF can be written as follows:

$$W(t,\vec{d}) = tf(t,\vec{d}) * \log(\frac{D}{n_t} + 0.01) \qquad (6)$$

where D is the total number of training documents, n_t is the number of documents containing the word t. $tf(t,\vec{d})$ indicate the occurrences of word t in document $\vec{d}$.

Formula (2) is used to measure the similarity of document vectors.

3.3 The Adjusted Feature Weighting Method

According to formula(6), more number of documents contain the word t,the word will be more unimportance. Let n_{1t} be the number of positive documents containing the word t and n_{2t} be the number of negative documents containing the word t,Then

$$n_t = n_{t1} + n_{t2} \qquad (7)$$

Analyzing the characteristics of corpus used in this paper, the number of positive documents containing emotinal words may be large. Simultaneously, the number of negative documents containing these words may be small. From formula (7), we could get that the relatively large number of documents contain these emotinal words, although they play an importance role in text categorization.

To the characteristics of corpus, TF-IDF could be adjusted to adapt to the characteristics of corpus.

1) Training corpus : the feature weighting method could be adjusted as follows:

$$W^*(t,\vec{d},C) = tf(t,\vec{d}) * \log(\frac{D}{n_t} + 0.01 + cf(t,C)) \qquad (8)$$

where C is category of document $\vec{d}$, $cf(t,C)$ indicate the occurrences of word t in category C.

2) Testing corpus: the feature weighting method could be adjusted as follows:

$$W^{**}(t,\vec{d}) = tf(t,\vec{d}) * \log(\frac{D}{n_t} + 0.01 + \frac{tf(t,\vec{d})}{cNum}) \qquad (9)$$

where $cNum$ is the number of categories .

3.4 Experimental Result

Table 1 the accuracy of KNN text classifier with TF-IDF and the adjusted TF-IDF

Accuracy / Fold to test	TF-IDF	The adjusted TF-IDF
fold-1	0.601010101	0.737373737
fold-2	0.636363636	0.737373737
fold-3	0.611111111	0.757575758
fold-4	0.616161616	0.797979798
fold-5	0.646464646	0.782828283
fold-6	0.671717172	0.797979798
fold-7	0.636363636	0.762626263
fold-8	0.585858586	0.752525253
fold-9	0.631313131	0.742424242

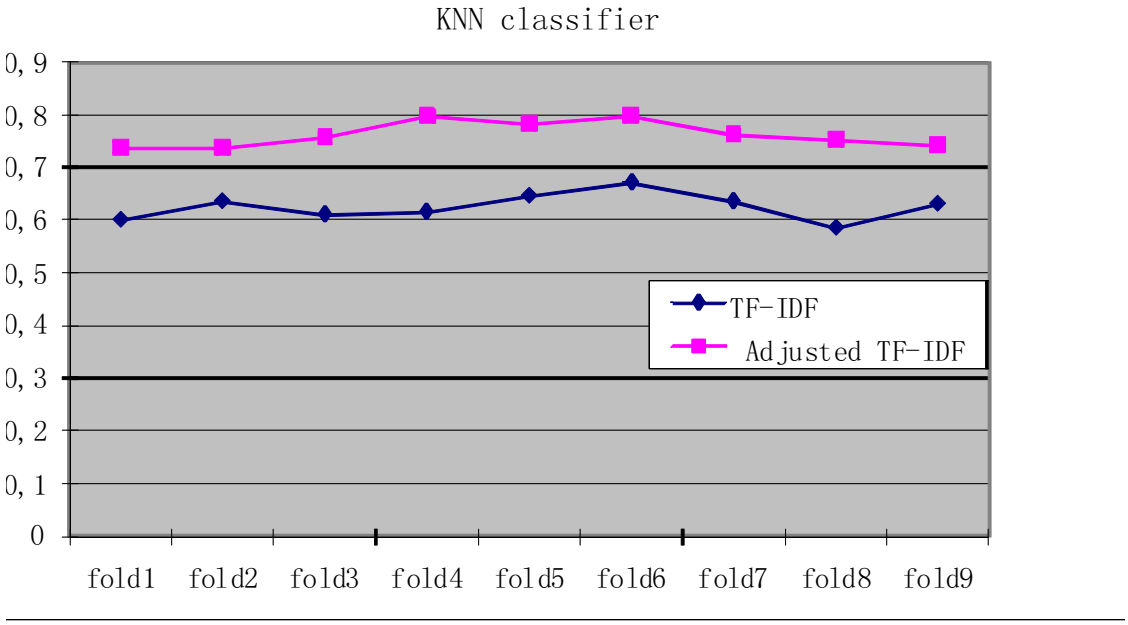


Fig. 1 Curves Basing on Table 1

From Fig.1, experimental result shows the accuracy of KNN text classifier can be improved with the adjusted feature weighting method.

4 Conclusion

To the characteristics of corpus used in this paper, TF-IDF is adjusted to a new method. Experimental result shows the accuracy of KNN text classifier can be improved with the adjusted feature weighting method. The future work is to apply the adjusted feature method to Naive Bayesian classifier and SVM classifier etc.

References

[1] Songbo Tan: An effective refinement strategy for KNN text classifier. Expert Systems with Applications Vol.30-2(2006), p.290-298.
[2] Yu-Long Qiao ,Jeng-Shyang Pan , Sheng-He Sun : Improved K nearest neighbor classification algorithm. Circuits and Systems Vol.2(2005),p.1101-1104.
[3] Huang Wei,Liu Yi ,Gao Bing ,Yang Ke-wei. Study on Method of Word Segmentation in Feature Selection in Chinese Text Categorization, Proceedings of Third International Conference on Knowledge Discovery and Data Mining, (2010).
[4]Jung-Yi Jiang ,Shie-Jue Lee .A Weight-based Feature Extraction Approach for Text Classification, Proceedings of Second International Conference on Innovative Computing, Information and Control, (2007).
[5] Jain, A.,Zongker, D.: Feature selection: evaluation, application, and small sample performance. Pattern Analysis and Machine Intelligence Vol.19-2(1997), p.153-158.
[6] Yanling Li ,Jing Yuan ,Xia Ye : Method for feature word weight calculating, Proceedings of IEEE International Conference on Intelligent Computing and Intelligent Systems, (2009).
[7] Information on http://www.searchforum.org.cn/tansongbo/corpus-senti.htm .

Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.704

Existence of positive solution for multi-point Sturm-Liouville boundary value problem

Xianrui Meng[1,a], Nana Li[2,b] and Yuxia Tong [1,c]

[1] College of Science, Hebei United University, Tangshan, P.R.China

[2] Tangshan College,Tangshan, *P.R.China*

[a]xianruimeng@yahoo.com.cn, [b]lalazinana@126.com, [c]tongyuxia@126.com

Keywords: boundary value problem; positive solution; cone

Abstract: Multi-point boundary value problem is studied in this paper. With the condition that nonlinear term f is superlinear or sublinear, it is proved that there exists at least one positive solution to multi-point Sturm-Liouville boundary value problem by using the fixed-point theorem concerning cone expansion and compression of norm type.

Introduction

Multi-point boundary value problems (BVPs)for ordinary differential equations arise in a variety of areas of applied mathematics and physics(see[1]); also,many problems in the theory of elastic stability can be handled by multi-point problems in [2].The existence of solutions of second order multi-point boundary value problems has been studied by many authors using the nonlinear alternative of Leray-Schauder fixed point theorem in cones (see[3,4,5]).There are very few works on the multi-point boundary value problem for higher order ordinary differential equation.

Recently, the existence of solutions of fourth order four-point boundary value problems has been studied by reference[6] using the fixed point theorem in cones.

$$\begin{cases} y^{(4)}(t)-f\left(t,y(t),y''(t)\right)=0, 0\le t\le 1 \\ y(0)=y(1)=0 \\ ay''(\xi_1)-by'''(\xi_1)=cy''(\xi_2)+dy'''(\xi_2)=0 \end{cases} \tag{1}$$

We are going to investigate more general fourth nonlinear ordinary differential equation with the following four-point boundary value conditions

$$\begin{cases} u^{(4)}(t)-f\left(t,u(t),u''(t)\right)=0, 0\le t\le 1 \\ \alpha u(\mu_1)-\beta u'(\mu_1)=\gamma u(\mu_2)+\delta u'(\mu_2)=0 \\ au''(\xi_1)-bu'''(\xi_1)=cu''(\xi_2)+du'''(\xi_2)=0 \end{cases} \tag{2}$$

We assume $f:[0,1]\times[0,+\infty)\times(-\infty,0]\to[0,+\infty)$ is continuous, $\alpha,\beta,\gamma,\delta,a,b,c,d$ are nonnegative constants and $0\le\xi_1<\xi_2\le 1, 0\le\mu_1<\mu_2\le 1,$ furthermore , $\alpha,\beta,\gamma,\delta,a,b,c,d$, ξ_1,ξ_2 ,

μ_1, μ_2 satisfy the following conditions:

$\rho_1 := \alpha\delta + \beta\gamma + \alpha\gamma(\mu_2 - \mu_1) > 0$， $\rho_2 := ad + bc + ac(\xi_2 - \xi_1) > 0$， $-a\xi_1 + b \geq 0$ *and* $c(\xi_2 - 1) + d \geq 0$

The way of processing in this paper different from paper [6]. With the condition that f is superlinear or sublinear , we proved that the boundary value problem (2) has at least one positive solution.

Preliminaries and Lemmas

The key tool in our approach is the following fixed point theorem.

lemma1: Let X be a Banach space,and $P \subset X$ be a cone in X .Assume Ω_1, Ω_2 are open subsets of X with $0 \in \Omega_1, \overline{\Omega_1} \subseteq \Omega_2$ and let $T : P \cap (\overline{\Omega_2} \setminus \Omega_1) \to P$ be a completely continuous operator such that either:

1) $\| Tu \| \leq \| u \|, u \in P \cap \partial\Omega_1$ *and* $\| Tu \| \geq \| u \|, u \in P \cap \partial\Omega_2$

2) $\| Tu \| \geq \| u \|, u \in P \cap \partial\Omega_1$ *and* $\| Tu \| \leq \| u \|, u \in P \cap \partial\Omega_2$

Then T has a fixed point in $P \cap (\overline{\Omega_2} \setminus \Omega_1)$.

Let $G_1(t,s)$ is Green function of boundary value problem

$$\begin{cases} -u''(t) = 0, 0 \leq t \leq 1 \\ \alpha u(\mu_1) - \beta u'(\mu_1) = \gamma u(\mu_2) + \delta u'(\mu_2) = 0 \end{cases}$$

obviously

$$G_1(t,s) = \frac{1}{\rho_1} \begin{cases} (\alpha(s - \mu_1) + \beta)(\gamma(\mu_2 - t) + \delta), s \leq t, \mu_1 \leq s \leq \mu_2 \\ (\alpha(t - \mu_1) + \beta)(\gamma(\mu_2 - s) + \delta), t \leq s, \mu_1 \leq s \leq \mu_2 \end{cases} \tag{3}$$

Denote $X = C[0,1]$, which equipped with the norm $\| v \| = \max_{t \in [0,1]} | v(t) |$, then $(X, \| \cdot \|)$ is a Banach space .Let

$$K = \{ v \in X : v(t) \geq 0, t \in [0,1] \}, \qquad P = \left\{ v \in K : \min_{\eta_1 \leq t \leq \eta_2} v(t) \geq \frac{1}{4} \| v \| \right\}$$

Where $\eta_1 = \xi_1 + \frac{1}{4}(\xi_2 - \xi_1), \qquad \eta_2 = \xi_2 - \frac{1}{4}(\xi_2 - \xi_1)$

Then K and P is obviously a cone in X .We define the operator $S : K \to X$ by

$$(Sv)(t) = \int_{\mu_1}^{\mu_2} G_1(t,s) v(s) ds, t \in [0,1] \tag{4}$$

Lemma 2： Suppose the boundary value problem of integral differential equation

$$\begin{cases} v''(t) + f(t, (Sv)(t), -v(t)) = 0, 0 \leq t \leq 1 \\ av(\xi_1) - bv'(\xi_1) = cv(\xi_2) + dv'(\xi_2) = 0 \end{cases} \tag{5}$$

has positive solution，then the boundary value problem (2) has positive solution too.

Proof: Assume $v(t)$ is a positive solution of the boundary value problem(5)，let $u(t)=(Sv)(t),t\in[0,1]$，then we can proof $u(t)$ is a positive solution of boundary value problems (2) easily.

Let $G_2(t,s)$ is Green function of boundary value problem

$$\begin{cases}-v''(t)=0,0\le t\le 1\\ av(\xi_1)-bv'(\xi_1)=cv(\xi_2)+dv'(\xi_2)=0\end{cases} \tag{6}$$

Then $G_2(t,s)=\dfrac{1}{\rho_2}\begin{cases}\big(a(s-\xi_1)+b\big)\big(c(\xi_2-t)+d\big),s\le t,\xi_1\le s\le\xi_2\\ \big(a(t-\xi_1)+b\big)\big(c(\xi_2-s)+d\big),t\le s,\xi_1\le s\le\xi_2\end{cases}$

Lemma3： $G_2(t,s)$ hold the following properties(see[6])：

1） $0\le G_2(t,s)\le G_2(s,s),t\in[0,1],s\in[\xi_1,\xi_2]$，　　2） $G_2(t,s)\ge\dfrac{1}{4}G_2(s,s),t\in[\eta_1,\eta_2],s\in[\xi_1,\xi_2]$

We define the operator $T:P\to X$ by

$$Tv(t)=\int_{\xi_1}^{\xi_2}G_2(t,s)f\big(s,(Sv)(s),-v(s)\big)ds,t\in[0,1] \tag{7}$$

obviously，a fixed point of operator T is positive solutions of boundary value problem (5).

Lemma 4： $T:P\to P$ is a completely continuous operator(see[6]) .

Main result

Denote $f^0=\lim\limits_{-y\to0^+}\max\limits_{t\in[\xi_1,\xi_2]}\sup\limits_{x\in[0,+\infty)}\dfrac{f(t,x,y)}{-y}$，　$f_\infty=\lim\limits_{-y\to+\infty}\min\limits_{t\in[\eta_1,\eta_2]}\inf\limits_{x\in[0,+\infty)}\dfrac{f(t,x,y)}{-y}$

$f_0=\lim\limits_{-y\to0^+}\min\limits_{t\in[\eta_1,\eta_2]}\inf\limits_{x\in[0,+\infty)}\dfrac{f(t,x,y)}{-y}$，　$f^\infty=\lim\limits_{-y\to+\infty}\max\limits_{t\in[\xi_1,\xi_2]}\sup\limits_{x\in[0,+\infty)}\dfrac{f(t,x,y)}{-y}$

Theorem1: Assume that f is superlinear，i.e., $f^0=0$ and $f_\infty=+\infty$，then the boundary value problem(2) has at least one positive solution.

Proof: since $f^0=0$，then for any ε satisfying $\varepsilon\int_{\xi_1}^{\xi_2}G_2(s,s)ds\le1$,we can choose $h_1>0$ such that

$$f(t,x,y)\le\varepsilon(-y)，\ t\in[\xi_1,\xi_2]，\ x\in[0,+\infty)，\ -y\in[0,h_1] \tag{8}$$

Set $\Omega_1=\{v\in P:\|v\|<h_1\}$,For $\forall v\in\partial\Omega_1$，from (8) and lemma 3 we can get

$$Tv(t)=\int_{\xi_1}^{\xi_2}G_2(t,s)f\big(s,(Sv)(s),-v(s)\big)ds\le\varepsilon\int_{\xi_1}^{\xi_2}G_2(s,s)v(s)ds\le\varepsilon\int_{\xi_1}^{\xi_2}G_2(s,s)\|v\|ds\le\|v\|$$

Which implies

$$\|Tv\|\le\|v\|，\ v\in\partial\Omega_1 \tag{9}$$

next，since $f_\infty=+\infty$，then for any ε^* satisfying $\dfrac{1}{4}\varepsilon^*\int_{\eta_1}^{\eta_2}G_2\left(\dfrac{1}{2}(\eta_1+\eta_2),s\right)ds\ge1$,there exists $h_2>h_1$ such that

$$f(t,x,y)\geq \varepsilon^{*}(-y)\quad t\in[\eta_1,\eta_2],x\in[0,+\infty),\quad -y\in\left[\frac{1}{4}h_2,+\infty\right) \tag{10}$$

Set $\Omega_2=\{v\in P:\|v\|<h_2\}$, For $\forall v\in\partial\Omega_2$,we know that $\min\limits_{\eta_1\leq t\leq\eta_2}v(t)\geq\frac{1}{4}\|v\|=\frac{1}{4}h_2$

Thus , from(10) we have

$$Tv\left(\frac{1}{2}(\eta_1+\eta_2)\right)=\int_{\xi_1}^{\xi_2}G_2\left(\frac{1}{2}(\eta_1+\eta_2),s\right)f\left(s,(Sv)(s),-v(s)\right)ds$$

$$\geq\varepsilon^{*}\int_{\eta_1}^{\eta_2}G_2\left(\frac{1}{2}(\eta_1+\eta_2),s\right)v(s)ds\geq\frac{1}{4}\varepsilon^{*}\int_{\eta_1}^{\eta_2}G_2\left(\frac{1}{2}(\eta_1+\eta_2),s\right)\|v\|ds\geq\|v\|$$

Which implies

$$\|Tv\|\geq\|v\|,\quad v\in\partial\Omega_2 \tag{11}$$

Hence from(9) , (11) and Lemma 1,the operator T has a fixed point u in $P\cap\left(\overline{\Omega_2}\setminus\Omega_1\right)$, u is a positive solution of boundary value problems (2).

Theorem2: Assume that f is sublinear，i.e., $f_0=+\infty$ and $f^{\infty}=0$， then the boundary value problem(2) has at least one positive solution.

Proof: since $f_0=+\infty$，then for any ε satisfying $\frac{1}{4}\varepsilon\int_{\eta_1}^{\eta_2}G_2\left(\frac{1}{2}(\eta_1+\eta_2),s\right)ds\geq 1$,we can choose $h_3>0$ such that

$$f(t,x,y)\geq\varepsilon(-y),\ t\in[\eta_1,\eta_2],\ x\in[0,+\infty),\ -y\in[0,h_3] \tag{12}$$

Set $\Omega_3=\{v\in P:\|v\|<h_3\}$，For $\forall v\in\partial\Omega_3$，using method of [6], we can get

$$\|Tv\|\geq\|v\|,\quad v\in\partial\Omega_3 \tag{13}$$

We consider two cases:

Case(i):Suppose f is bounded,i.e.,there exists a positive constant L such that $f(t,x,y)\leq L$.Take $h_4\geq\max\left\{L\int_{\xi_1}^{\xi_2}G_2(s,s)ds,2h_3\right\}$,For $\forall v\in P$ and $\|v\|=h_4$,from lemma 3 we can get

$$Tv(t)=\int_{\xi_1}^{\xi_2}G_2(t,s)f\left(s,(Sv)(s),-v(s)\right)ds\leq L\int_{\xi_1}^{\xi_2}G_2(s,s)ds\leq h_4=\|v\|$$

Case(ii): Suppose f is unbounded,then define a function $f^{*}:[0,+\infty)\to[0,+\infty)$ by

$$f^{*}(r):=\max\{f(t,x,y):\xi_1\leq t\leq\xi_2,0\leq x\leq Mr,0\leq -y\leq r\},\quad M=\max_{\xi_1\leq t\leq\xi_2}\int_{\mu_1}^{\mu_2}G_1(t,s)ds$$

It is easy to see that $f^{*}(r)$ is nondecreasing and $\lim\limits_{r\to+\infty}\frac{f^{*}(r)}{r}=0$, then for any ε^{*} satisfying $\varepsilon^{*}\int_{\xi_1}^{\xi_2}G_2(s,s)ds\leq 1$,we can choose $h_4>h_3$ such that

$$f^{*}(r)\leq\varepsilon^{*}r,r\in[h_4,+\infty) \tag{14}$$

Then,from defined f^{*} and (14),we can get

$$f(t,x,y)\leq f^{*}(h_4)\leq\varepsilon^{*}h_4,\qquad \xi_1\leq t\leq\xi_2,0\leq x\leq Mh_4,0\leq -y\leq h_4 \tag{15}$$

On the other hand,for $\forall v\in P$ and $\|v\|=h_4$,we know that

$$0\leq(Sv)(t)=\int_{\mu_1}^{\mu_2}G_1(t,s)v(s)ds\leq h_4\max_{\xi_1\leq t\leq\xi_2}\int_{\mu_1}^{\mu_2}G_1(t,s)ds=Mh_4,\qquad t\in[\xi_1,\xi_2] \tag{16}$$

From lemma 3 and (15)、(16)，we can get

$$Tv(t)=\int_{\xi_1}^{\xi_2}G_2(t,s)f\left(s,(Sv)(s),-v(s)\right)ds\leq\varepsilon^{*}h_4\int_{\xi_1}^{\xi_2}G_2(s,s)ds\leq h_4=\|v\|$$

Therefore,in either case,we can set $\Omega_4 = \{v \in P : \| v \| < h_4\}$,such that

$$\| Tv \| \leq \| v \|, v \in \partial\Omega_4 \tag{17}$$

Hence from(13) , (17) and Lemma 1,the operator T has a fixed point u in $P \cap \left(\overline{\Omega_4} \setminus \Omega_3\right)$, u is a positive solution of boundary value problems (2).

Acknowledgments

This work was supported by Natural Science Foundation of Hebei Province of China (A2010000910) and Science Research Foundation of Hebei Polytechnic University (No. z200916).

References

[1] M. Moshiinsky: Sobre los problems de condiciones a la frontiera en una dimension de caracteristicas discontinuas. Bol.Soc.Mat.Mexicana Vol. 7 (1950), p. 1-25.

[2] S. Timoshenko: Theory of Elastic Stability. McGraw-Hill,New York(1961).

[3] A.K. Rahmat, M. Rafique: Existence and multiplicity results for some three-point boundary value problems.Nonlinear Anal. Vol. 66 (2007), p.1686-1697.

[4] W.H. Jiang, Y.P. Guo: Multiple positive solutions for second-order m-point boundary value problems. J Math Anal Appl.Vol.327(2007), p.415-424.

[5] H. Feng, W.G. Ge: Triple symmetric positive solutions for multipoint boundary value problem with one dimensional p-Laplacian,Mathematical and Computer Modelling Vol.47(2008) p.186-195.

[6] Y.L. Zhong, S.H. Chen and C.P. Wang: Existence results for a fourth-order ordinary differential equation with a four-point boundary condition. Appl Math Lett.Vol. 21(2008), p. 465-470.

doi:10.4028/www.scientific.net/AMM.50-51.709

New Harmonious Development of The College Sports Game Status and Strategic Thinking

Jianyong Di [1], Yanmei Yang [2], Kailei Xi [1] and Xiaofeng Xu [3]

[1] College of Light Industry, Hebei Polytechnic University, Tangshan, China
[2] College of Science, Hebei Polytechnic University, Tangshan, China
[3] Department of Sports, Hebei Polytechnic University, Tangshan, China

dijianyong60@126.com

Keywords：College; Game; Tactics;

Abstract. This paper introduces game theory to analysis of the development of college sports in the emerging imbalance. In the new social conditions, the paper is to explore the deep interaction of decision-making behavior in college educational resource allocation. Aim at exploring to make college physical education resources optimize the mode of cooperation, and promote the harmonious development of college sports.

1. Introduction

In the new social environment, from the point of the stage on the effectiveness of college sports , strengthen their students' health is the main purpose , but the current situation on the effect of college sports enhancing student body is not satisfactory, which may be relevant with our methods of operation and guiding ideology [1]. Ago, we emphasized on the reality of increasing physical benefits, but ignore how to mobilize the students actively, how to encourage students to form the habit of physical exercise. Now the education sector has reached a consensus, education will pay more attention to teaching students life-long education and learning [2]. With the changes in school physical education guidelines, the overall objective of school physical education has been, strengthen their students' health, promote the full development of students, and the awareness of students engaged in sports, interests, habits, and abilities. Improve physical literacy to lay a good foundation for life-long physical education.

Game refers to the process that individuals or organizations, under certain rules and certain environmental conditions，rely on the information available, start from its maximum benefits to choose their own actions or strategies and implement , make the expected utility or benefits to achieve the maximum[3]. Game Theory is a science that analysis of game behavior and game decision-making, is a theory that study the behavior of the main decision-making under the direct interaction, how to make decisions, and that how to achieve balance[4]. This paper introduces game theory to analysis of the development of college sports in the emerging imbalance. In the new social conditions, the paper is to explore the deep interaction of decision-making behavior in college educational resource allocation. Aim at exploring to make college physical education resources optimize the mode of cooperation, and promote the harmonious development of college sports.

2. The Main Problems of College Sports

2.1 The Deviation of Physical Ideas

First, misunderstanding of enhancing students in college physical teaching

"Enhancing Students" is one of the important objective of physical education, which is the founding of New China, relevant Comrade Mao introduce the guidelines of develop sports and enhance people's health. For a long period of time, the main goal of physical education did not suffer from doubt. But with the physical education reform, physical education produced a number of issues in the practical work.

Usually within 90 minutes of PE curriculum, up to about 40% of the time for physical exercise, a student will have 40 minutes. Coupled with the capacity, sports facilities and other factors, students have less time in the classroom training. Each semester has 10 hours, how can we enhance physique in such short time.

Second, students drop in status, whether due to physical education is not achieve the sports teaching objectives of "physical enhancement". Students' monitoring results in 2001 tell us that, in addition to a slight increase in height and weight than in the past, students' endurance, strength and lung capacity are even worse. Whether the physical education has gone wrong? there are many reasons due to student physical health problems; different regions have different problems and different reasons. Mainly due to students lack of physical exercise, including time and intensity is not enough. Among these are lack of school space, timing and content of arranging sports activities, but also the student's lack of hard training. He also emphasized that the social accelerated rhythm, academic pressure and social competition increase, lack of sleep and mental stress also become reasons that affect student's health. Clearly, the decline of student physical health status is not due to physical education.

2.2 Confusion and Reflection of Physical Education Goals

"To teach students basic knowledge, basic technology and basic skill" (hereinafter referred to impart three bases) are the most commonality goal for physical education and other disciplines. However, due to physical science's characteristics in the teaching, as we study the physical theory is not thorough; there are also doubts in objective understanding of the physical education.

First of all, the "basis" and "basic "of sports science based on scientific content in logic and human's requirements. Sports science was not clearly defined in both areas. Consequently, the "basis" and "basic" of sports science has been unclear. Second, is a realistic way of thinking that consider the "basis" and "basic "of sports science from social sports and life-long sports practice's needs. However, under the mindset that we regard the "practical" as a "pragmatic", under the conditions that the thinking in lifelong sports not to be accepted the research and work in this area also appeared quite lacking. Therefore, the methodology that we determine the "basis" and "basic "of sports science from social sports and life-long sports practice's needs has not yet established.

2.2 Due to college sports practice cannot keep touch in the guiding ideology of lifelong physical, the students dislike physical education.

Our school physical education still have the phenomenon that of great importance to the competitive sports and curricular contempt, neglect extracurricular contempt. Teaching content part from the actual situation is difficult to grasp for students, is not conducive to students interested in sports and difficult to develop exercise habits. is not conducive to strengthening students either. Divided education and development into several disparate parts is not conducive to student learning, is a heavy physical burden on students.

College Sports did not lay a foundation for life-long sport mainly as follows: students in a poor

sense of physical training, weak health awareness and students lack of physical exercise habits and abilities. This situation mainly due to our school sports make too much emphasis on students' short-term benefit in the school, neglect life-long physical exercise social function and value for students entering society .We also emphasized on students form physical exercise habits, but was not comprehensive, profound in theory or not be a system in practice. Teach students in passive mechanical imitation. Largely suppressed the development of student individual, ignore the individual creativity of students and the development of physical capabilities.

3. The factors affecting the development of university sports and the inherent contradictions and countermeasures of game.

3.1 Analysis on higher education resources in college sports uneven distribution.

With the development of the community, sports have taken an increasingly significant role to promote human development. People pay more and more attention to sports. The sports have a crucial role to promote community and modernization of students It is an important means of personality to promote successful students.

From the current situation of the development in higher education, only a small part of the universities put emphasis on college sports. Most of the universities neglect college sports. This makes college sports in an embarrassing situation that would like to pursue development but have no resources to support, and it made college sports far behind the development of higher education. College is the last stop for people to accept school physical education. The lagging development of college sports, directly affect the students' formation of the concept of sports, sports consciousness, and physical exercise habits.

College sports are an integral part of higher education. Both of them share a common goal serve the people with the all-round development. But the two also constraints, constitute the game of the two sides. According to the views of Game Theory, the main issues to be addressed now, is to find a win-win situation led to collaboration between higher education and college sports.

3.2 Analysis on college sports' internal factors

College sports consist of sports classroom teaching, sports extracurricular activities, sports associations, various sports competitions and so on. Now in China, College sports are over-emphasis on classroom teaching, and neglect other components. It is not enough to form students' awareness of life-long sport and good exercise habits with weekly classroom instruction. It requires physical education teaching, sports extra- curricular activities and other components work together, for students to form a good habit of physical activity and good physical and mental state and to achieve the requirements for overall development.

Sports classroom teaching, sports extracurricular activities and sports associations are interrelated constraints and promote each other. Game Theory tells us if we take too much emphasis on one aspect will affect the other two. This cannot form a good win-win model，affecting the development of college sports either.

4. Conclusions and Recommendations

1) In recent decades our school sports study more in the issue of specific micro-level, but it is not enough in the ideas of school sports and other macro- level, meson-level problems. School sports' reform must be to develop deep. Here, we should study school sports trends and reform measures

mainly from the perspective of lifelong physical education. Establish life-long physical education as the main ideas of school sports' reform and development.

2) School sports implement the guiding ideology of life-long sports, reflects the life-long participation in sports activities to develop students ability and habit. Life-long physical idea is the best choice to guide other school sports' ideology. Regard life-long physical education as the dominant ideology of school sports' reform and development, which is the inevitable evolution of the concept of school sports, is also to promote the needs of school sports' reform and fitness. If we want school sports open up a new situation, we must uphold the guiding ideology of life-long sports.

3) Handle the relationship between school sports and competitive sports correctly. School sports and competitive sports are not antagonistic or incompatible. School sports' purpose is to enhance students' body, improve students' health, and let students learn basic knowledge, technology and skills Competitive sport is a useful supplement to school sports, competitive sports is the cultural treasures of mankind sports.

4) With the further reform of school physical education, new curriculum standards will be implemented nationwide. Physic and health curriculum requires students not only enhance physical fitness, have basic sports knowledge and sports skills, but students are expected to form the habit of physical exercise, healthy and positive lifestyles , optimistic and cheerful attitude towards life.

References

[1] Y.Z. Lu, sports sociology, Higher Education Press (2009), p.153-168

[2] W.X. Yang, Q. Chen, sports theory, Higher Education Press (2006), p. 85-104

[3] Y.Z. Lu, Introduction to Humanities and Social Sciences Sports (Advanced Course), Higher Education Press (2003) p.385-401

[4]J.R. Li， The competing and mutual influencing situation and tactical research between athletic star and advertising market, Journal of WuHan University of science and engineering, Vol.21, no.10 (2008).

doi:10.4028/www.scientific.net/AMM.50-51.713

The unbiased Estimator based on the function of poisson population

Zhendong Li, Cuixian Sun and WANG Cheng

Basic Education Department, Tangshan College, P.R. China, 063000

tsclzd97@sina.com

Keywords: Poisson population; unknown parameter; function; unbiased estimator

Abstract: The unbiased estimator of unknown parametric function based on possion's population is discussed, and the expression of two classes of estimable function are given. Applying Maclaurin's series, it's proved that the unbiased estimator of the functions are exist, and utilizing the induction method , derived out the generalized expression of unbiased estimator. By means of comparing the unbiased estimator of the two classes of estimable functions, it's found that the estimator cannot be direct constructed with parametric estimator.

0 Introduction

Let X ~Poisson(λ)， $g(x)$ is a real valued function, when $\hat{\theta}$ is the unbiased estimator of the unknown parameter θ. Generally speaking, $g(\hat{\theta})$ isn't the unbiased estimator of $g(\theta)$, thus, when $E(\hat{\theta})=\theta$, we don't get $E(g(\hat{\theta}))= g(\theta)$.the definition of unbiased estimator for the unknown parameters θ and $g(\theta)$ can be found in reference[1]. In the paper, we assumed that the population distribution is Poisson distribution and by the Maclaurin series expansion, we obtain that the unbiased estimator of the two expression of $g(\theta)$.

1 Several unbiased estimators of function

Let X ~Poisson(λ), and it's probability distribution is $P(X=x)=\frac{\lambda^x}{x!}e^{-\lambda}$， $x=$ 0, 1, 2…, the unknown parameter $\theta=\lambda>0$, Sampling the random samples $(X_1, X_2,\cdots, X_n)$ from population X, the probability distribution of X_i is $P(X_i=x)=\frac{\lambda^x}{x!}e^{-\lambda}$， $x=$ 0, 1, 2…, we make the estimator of $g(\theta)$ based on the random samples $(X_1, X_2,\cdots, X_n)$. And it's unbiased estimator is $\hat{g}(X)$. $\hat{g}(X)$ is the function of the random samples, so it is a random variable . so $\hat{g}(X)=f(X_1, X_2,\cdots, X_n)$.

Using $(x_1, x_2, \cdots, x_n)$ denotes the values of random samples $(X_1, X_2, \cdots, X_n)$. When the samples is given, it's a value of the estimator of $g(\theta)$.

1.1 for the case $g(\theta) = g(\lambda) = e^{k\lambda}$

In bibliography[2], the Maclaurin expansion of the function e^x is given by

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots, \quad (-\infty < x < +\infty)$$

Using the Maclaurin expansion of the function, we have $g(\theta) = g(\lambda) = e^{k\lambda}$.

Theorem 1. Let $g_1(\theta) = g_1(\lambda) = e^{\lambda}$, then the unbiased estimator of $g_1(\lambda)$ is 2^{X_i}.

Proof: By expectation of function of random variable[3] and Maclaurin expansion of e^x, We obtain

$$E(2^{X_i}) = \sum_{x=0}^{\infty} 2^x P(X_i = x) = \sum_{x=0}^{\infty} 2^x \frac{\lambda^x}{x!} e^{-\lambda} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(2\lambda)^x}{x!} = e^{-\lambda} e^{2\lambda} = e^{\lambda},$$

Thus, 2^{X_i} is the unbiased estimator of $g_1(\theta) = g_1(\lambda) = e^{\lambda}$.

Corollary 1. Let $g_2(\theta) = g_2(\lambda) = e^{2\lambda}$, then the unbiased estimator of $g_2(\lambda)$ is 3^{X_i}.

Proof: From Th (1) it follows that

$$E(3^{X_i}) = \sum_{x=0}^{\infty} 3^x P(X_i = x) = \sum_{x=0}^{\infty} 3^x \frac{\lambda^x}{x!} e^{-\lambda} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(3\lambda)^x}{x!} = e^{-\lambda} e^{3\lambda} = e^{2\lambda}.$$

Combining Th (1) with Corollary (1) ,we have the following:

Theorem 2. Let $g_3(\theta) = g_3(\lambda) = e^{k\lambda}$,then the unbiased estimator of $g_3(\lambda)$ is $(k+1)^{X_i}$,where k is a positive integer.

Proof: By Th (1) with Corollary (1), we have

$$E((k+1)^{X_i}) = \sum_{x=0}^{\infty} (k+1)^x P(X_i = x) = \sum_{x=0}^{\infty} (k+1)^x \frac{\lambda^x}{x!} e^{-\lambda}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{((k+1)\lambda)^x}{x!} = e^{-\lambda} e^{(k+1)\lambda} = e^{k\lambda}.$$

1.2 for the case $g(\theta) = g(\lambda) = e^{-k\lambda}$

Theorem 3. Let $g_4(\theta) = g_4(\lambda) = e^{-\lambda}$, then the unbiased estimator of $g_4(\lambda)$ is

$$\hat{g}(X)=\hat{e}(X_i)=\begin{cases}1, & X_i=0\\ 0, & X_i\neq 0\end{cases}$$

Proof: Set estimator $\hat{e}(X_i)=f(X_i)$,so

$$E(\hat{e}(X_i))=\sum_{x=0}^{\infty}f(x)P(X_i=x)=\sum_{x=0}^{\infty}f(x)\frac{\lambda^x}{x!}e^{-\lambda}$$

$$=e^{-\lambda}\cdot 1\cdot\frac{\lambda^0}{0!}+e^{-\lambda}\sum_{x=1}^{\infty}0\cdot\frac{\lambda^x}{x!}=e^{-\lambda}\quad .$$

Corollary 2. Let $g_5(\theta)=g_5(\lambda)=e^{-2\lambda}$, then the unbiased estimator of $g_5(\lambda)$ is

$$\hat{t}(X_i)=\begin{cases}1, & X_i \text{ is even.}\\ -1, & X_i \text{ is odd.}\end{cases}$$ (Note that number 0 is an even)

Proof: Set $\hat{t}(X_i)=\varphi(X_i)$, from Th (3) it follows that

$$E(\hat{t}(X_i))=\sum_{x=0}^{\infty}\varphi(x)P(X_i=x)$$

$$=\sum_{x=0}^{\infty}(-1)^x\frac{\lambda^x}{x!}e^{-\lambda}=e^{-\lambda}\sum_{x=0}^{\infty}\frac{(-\lambda)^x}{x!}=e^{-\lambda}e^{-\lambda}=e^{-2\lambda}\quad .$$

Combining Th (3) with Corollary (2), we conclude the following theorem:

Theorem 4. Let $g_6(\theta)=g_6(\lambda)=e^{-k\lambda}$, then the unbiased estimator of $g_6(\lambda)$ is

$\hat{m}(X_i)=(-(k-1))^{X_i}$, where $k\geq 2$, 0 is even.

Proof: Define estimator $\hat{m}(X_i)=\rho(X_i)$, so

$$E(\hat{m}(X_i))=\sum_{x=0}^{\infty}\rho(x)P(X_i=x)$$

$$=\sum_{x=0}^{\infty}(-(k-1))^x\frac{\lambda^x}{x!}e^{-\lambda}=e^{-\lambda}\sum_{x=0}^{\infty}\frac{(-(k-1)\lambda)^x}{x!}=e^{-\lambda}e^{-(k-1)\lambda}=e^{-k\lambda}\quad .$$

2 **When $\hat{\theta}$ is the unbiased estimator of the unknown parameter θ, $g(\hat{\theta})$ is not the unbiased estimator of $g(\theta)$.**

When X ~Poisson(λ),and it's probability distribution is $P(X=x)=\frac{\lambda^x}{x!}e^{-\lambda}$, $x=0, 1, 2\ldots$, the unknown parameter $\theta=\lambda>0$, sampling the random samples $(X_1, X_2,\cdots, X_n)$ from the

population X, we can proved [4] $\overline{X}$ and $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ are all the unbiased estimator of λ.

In generally speaking ,for arbitary real number α，0≤α≤1，$\alpha\overline{X}+$（1-α）S^2 is the unbiased estimator of λ.thus $\hat{\theta}=\hat{\lambda}=\overline{X}$ （S^2 or $\alpha\overline{X}+$（1-α）S^2）.

Since $\overline{X}$ is the unbiased estimator of λ [5]. then $\hat{\theta}=\hat{\lambda}=\overline{X}$. think about , $g_7(\theta)=g_7(\lambda)=\lambda^2$ is the function of λ, whether $g_7(\hat{\theta})=g_7(\hat{\lambda})$ is the unbiased estimator or not? In other words, its unbiased estimator is $(\hat{\theta})^2=(\hat{\lambda})^2=\overline{X}^2$ or not?

Lemma 1. Suppose that X ~Poisson(λ),sampling the random samples（X_1，X_2,…，X_n）from the population X, then $n\overline{X}=\sum_{i=1}^{n}X_i\sim P(n\lambda)$.

Proof: since X_1 ,X_2 i.i.d. and $X_1\sim P(\lambda)$, $X_2\sim P(\lambda)$, then probability distribution function of $Z=X_1+X_2$ is $P(z)=\sum_{x=0}^{z}P_{X_1}(x)P_{X_2}(z-x)=\frac{(2\lambda)^z}{z!}e^{-2\lambda}$, where $z=0$、1、2、…, that is $Z=X_1+X_2\sim P(2\lambda)$.

By mathematical induction, we deduce that $n\overline{X}=\sum_{i=1}^{n}X_i\sim P(n\lambda)$.

Theorem 5. Let $g_7(\theta)=g_7(\lambda)=\lambda^2$, then the unbiased estimator of $g_7(\theta)$ is $\frac{1}{n}\sum_{i=1}^{n}X_i(X_i-1)$, not $\overline{X}^2$.

Proof: By Lemma 1, we have

$$E(\overline{X}^2)=E(\frac{1}{n^2}(n\overline{X})^2)=\frac{1}{n^2}E((n\overline{X})^2)=\frac{1}{n^2}[D(n\overline{X})+(E(n\overline{X}))^2]$$

$$=\frac{1}{n^2}[D(\sum_{i=1}^{n}X_i)+(E(\sum_{i=1}^{n}X_i))^2]=\frac{1}{n^2}(n\lambda+(n\lambda)^2)=\frac{1}{n}\lambda+\lambda^2\neq\lambda^2.$$

Hence $(\hat{\theta})^2=(\hat{\lambda})^2=\overline{X}^2$ is not the unbiased estimator of $g_7(\theta)$.

$$E(\frac{1}{n}\sum_{i=1}^{n}X_i(X_i-1))=E(\frac{1}{n}\sum_{i=1}^{n}X_i^2-\frac{1}{n}\sum_{i=1}^{n}X_i)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^2)-\frac{1}{n}\sum_{i=1}^{n}E(X_i)$$

$$=\frac{1}{n}\sum_{i=1}^{n}(D(X_i)+(E(X_i))^2)-\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}(\lambda+\lambda^2)-\frac{1}{n}\sum_{i=1}^{n}\lambda=\lambda^2.$$

So $\frac{1}{n}\sum_{i=1}^{n}X_i(X_i-1)$ is the unbiased estimator of $g_7(\theta)$.

In theorem 6, we obtained the unbiased estimator of $g_1(\theta)=g_1(\lambda)=e^{\lambda}$ is 2^{X_i}, the conclusion can't get from $g_1(\hat{\theta})=e^{\overline{X}}$, for $g_1(\hat{\theta})=e^{\overline{X}}$ is not the unbiased estimator. proof as follows.

Theorem 6. Let $g_1(\theta)=g_1(\lambda)=e^{\lambda}$, then $g_1(\hat{\theta})=e^{\overline{X}}$ is not the unbiased estimator of $g_1(\theta)$.

Proof: $E(g_1(\hat{\theta}))=E(g_1(\hat{\lambda}))=E(e^{\overline{X}})=E(e^{\frac{1}{n}(n\overline{X})})=\sum_{z=0}^{\infty}e^{\frac{z}{n}}P(n\overline{X}=z)$

$$=\sum_{z=0}^{\infty}e^{\frac{z}{n}}\frac{(n\lambda)^z}{z!}e^{-n\lambda}=e^{-n\lambda}\sum_{z=0}^{\infty}e^{\frac{z}{n}}\frac{(n\lambda)^z}{z!}\neq e^{\lambda}$$

3 Conclusion

We obtain the unbiased estimator of two functions $g(\theta)=g(\lambda)=e^{k\lambda}$ and $g(\theta)=g(\lambda)=e^{-k\lambda}$, given the conditions that X ~Poisson(λ) and $\theta=\lambda$. And analysis the unbiased estimator of the unknown parameters θ and $g(\theta)$.

References

[1] X.R. Chen: Probability and statistics. Beijing, China Science Press.(2000) p.175

[2] Mathematics Department of Tongji university: Higher mathematics (sixth edition). Beijing, higher education press.(2008)p.280

[3] H.F. Shen: Probability and statistics, Beijing, higher education press.2003.p146-147

[4] J.F. Bu, C.X. Sun: Existence of unbiased estimator for function f(p) under B(n,p) population. Pur and Applied Mathematics,2(2007),p.201-204

[5] J.F. Bu, C.X. Sun: Unbiased Estimator of Parameter in Poisson Distribution. Journal of Tangshan College, 6(2009),p.8-9

© (2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.718

The Equivalence of Mann and Implicit Mann Iterations for Uniformly Pseudocontractive Mappings in Uniformly Smooth Banach Spaces

Cheng Wang, Zhiming Wang

Department of Basic Science, Tangshan College, Tangshan,063000 P.R.China

149960953@qq.com

Keywords: Mann iteration, implicit Mann iteration, uniformly pseudocontractive map; generalized Lipschitzian map, uniformly smooth Banach space

Abstract. In this paper, suppose E is an arbitrary uniformly smooth real Banach space, and D is a nonempty closed convex subset of E. Let T be a generalized Lipschitzian and uniformly pseudocontractive self-map with $F(T)\neq\phi$. Suppose that $\{u_n\}_{n=1}^{\infty}$, $\{x_n\}_{n=1}^{\infty}$ are defined by Mann iteration and implicit Mann iteration respectively, with the iterative parameter $\{a_n\}_{n=1}^{\infty}$ satisfying certain conditions. Then the above two iterations that converge strongly to fixed point of T are equivalent.

1. Introduction

Let E^* be the dual space of real Banach space E and D be a nonempty closed convex subset of E. Suppose $J:E\to 2^{E^*}$ is the normalized duality mapping defined by

$$J(x)=\left\{f\in E^*:<x,f>=\|x\|^2=\|f\|^2\right\},\forall x\in E,$$

where $<\cdot,\cdot>$ denotes the generalized duality pairing. The single-valued normalized duality mapping is denoted by j .

Definition 1.1[1] A mapping $T:D\to D$ is called uniformly pseudocontractive if for any $x,y\in D$, there exist $j(x-y)\in J(x-y)$ and a strictly increasing continuous function $\Phi:[0,+\infty)\to[0,+\infty)$ with $\Phi(0)=0$ such that

$$<Tx-Ty,j(x-y)>\leq\|x-y\|^2-\Phi(\|x-y\|).$$

Definition 1.2[2] The mapping $T:D\to D$ is called generalized Lipschitz map if there exists a constant $L>0$ such that

$$\|Tx-Ty\|\leq L(1+\|x-y\|),\qquad(1.1)$$

for all $x,y\in D(T)$. Clear, if map T either is Lipschitz or has a bounded range, then T must be a generalized Lipschitz map. Conversely, the following example indicates that the class of

*Project supported by Hebei province science and technology foundation(No. 072056197D).

generalized Lipschitz map neither is Lipschitz nor has the bounded range.

Example 1.3 (see [2]). Let $E=(-\infty,+\infty)$ with the usual norm. Define $T:E\to E$ by

$$Tx=\begin{cases}\dfrac{-3x}{2}+1, & if \quad x\in(-\infty,-1),\\ \dfrac{-3x}{2}+\sqrt{-x}, & if \quad x\in(-1,0),\\ \dfrac{-3x}{2}, & if \quad x\in[0,+\infty).\end{cases}$$

Definition 1.4 Let $T:D\to D$ be a mapping. For any given $u_1\in D$, define the sequence $\{u_n\}_{n=1}^{\infty}\subset D$ by the iterative schemes

$$u_{n+1}=(1-a_n)u_n+a_nTu_n, n\geq 1, \tag{1.2}$$

which is called Mann iteration, where $\{a_n\}_{n=1}^{\infty}$ is a real sequence in $[0,1]$ satisfying some certain conditions. Another, we assume that there exists $(I-tT)^{-1}$, for all $t\in(0,1)$. For any given $x_1\in D$, define the sequence $\{x_n\}_{n=1}^{\infty}\subset D$ by the iterative schemes

$$x_{n+1}=(1-a_n)x_n+a_nTx_{n+1}, n\geq 1, \tag{1.3}$$

which is called implicit Mann iterations.

The aim of this paper is to prove the equivalence between the convergence of above two iterations when T is generalized Lipschitz and uniformly pseudocontractive mapping in real uniformly smooth Banach spaces. For this, we need the following Lemmas.

Lemma 1.5[2] Let E be a real Banach space and let $J:E\to 2^{E^*}$ be a normalized duality mapping. Then

$$\|x+y\|^2\leq\|x\|^2+2<y,j(x+y)>,$$

for all $x,y\in E$ and each $j(x+y)\in J(x+y)$.

Lemma 1.6[3] Let $\{\rho_n\}_{n=1}^{\infty}$ be a nonnegative real numbers sequence satisfying the condition

$$\rho_{n+1}\leq(1-\theta_n)\rho_n+o(\theta_n),$$

where $\theta_n\in(0,1)$ with $\sum_{n=1}^{\infty}\theta_n=\infty$. Then $\rho_n\to 0$ as $n\to\infty$.

2. Main Results

Theorem 2.1 Let E be a real uniformly smooth Banach space and D be a nonempty closed convex subset of E, let $T:D\to D$ be a generalized Lipschitzian and uniformly pseudocontractive

map. Let q be a fixed point of T in D. Suppose that $\{u_n\}_{n=1}^{\infty}$ and $\{x_n\}_{n=1}^{\infty}$ are defined by (1.2) and (1.3) respectively, with the iterative parameter $\{a_n\}_{n=1}^{\infty}$ satisfying: $a_n \to 0$ as $n \to \infty$; $\sum_{n=0}^{\infty} a_n = \infty$. Then the following two assertions are equivalent:

(i) Mann iteration (1.2) converges to the fixed point of T;

(ii) implicit iteration (1.3) converges to the fixed point of T.

Proof. By the definition of uniformly pseudocontractive map, we know that the fixed point of T is unique. If Mann iteration (1.2) converges to the fixed point $q \in F(T)$, i.e., $\|u_n - q\| \to 0$ as $n \to \infty$, then we want to prove that $\|x_n - q\| \to 0$ as $n \to \infty$. Without loss of generality, we assume that $\|u_n - q\| \le 0$ for all $n \ge 0$.

Applying (1.1), (1.2) and Lemma1.5, we have

$$\begin{aligned}
&\|x_{n+1} - u_{n+1}\|^2 \\
&= \|(1-a_n)(x_n - u_n) + a_n(Tx_{n+1} - Tu_n)\|^2 \\
&\le (1-a_n)^2 \|x_n - u_n\|^2 + 2a_n < Tx_{n+1} - Tu_n, j(x_{n+1} - u_{n+1}) > \\
&\le (1-a_n)^2 \|x_n - u_n\|^2 + 2a_n < Tx_{n+1} - Tu_{n+1}, j(x_{n+1} - u_{n+1}) > \\
&+2a_n < Tu_{n+1} - Tu_n, j(x_{n+1} - u_{n+1}) > \\
&\le (1-a_n)^2 \|x_n - u_n\|^2 + 2a_n(\|x_{n+1} - u_{n+1}\|^2 - \Phi(\|x_{n+1} - u_{n+1}\|)) \\
&+2a_n \|Tu_{n+1} - Tu_n\| \cdot \|x_{n+1} - u_{n+1}\| \\
&\le (1-a_n)^2 \|x_n - u_n\|^2 + 2a_n(\|x_{n+1} - u_{n+1}\|^2 - \Phi(\|x_{n+1} - u_{n+1}\|)) \\
&+2a_n L(1 + \|u_{n+1} - u_n\| \cdot \|x_{n+1} - u_{n+1}\|.
\end{aligned} \tag{2.1}$$

Observe that

$$\begin{aligned}
&\|u_{n+1} - u_n\| \\
&= \|a_n(Tu_n - u_n)\| \\
&\le a_n(\|Tu_n - Tq\| + \|u_n - q\|) \\
&\le a_n(L + (1+L)\|u_n - q\|) \\
&\le a_n(1+2L).
\end{aligned} \tag{2.2}$$

Substituting (2.2) into (2.1), we obtain

$$\begin{aligned}
&\|x_{n+1} - u_{n+1}\|^2 \\
&\le (1-a_n)^2 \|x_n - u_n\|^2 + 2a_n(\|x_{n+1} - u_{n+1}\|^2 - \Phi(\|x_{n+1} - u_{n+1}\|)) \\
&+2a_n L(1 + a_n(1+2L)) \cdot \|x_{n+1} - u_{n+1}\| \\
&\le (1-a_n)^2 \|x_n - u_n\|^2 + 2a_n(\|x_{n+1} - u_{n+1}\|^2 - \Phi(\|x_{n+1} - u_{n+1}\|)) \\
&+a_n L(1 + a_n(1+2L)) + a_n L(1 + a_n(1+2L))\|x_{n+1} - u_{n+1}\|^2,
\end{aligned} \tag{2.3}$$

Without loss of generality, we assume that $0 < 1 - 2a_n - a_n L(1 + a_n(1+2L)) < 1$. Then (2.3) implies

that

$$\begin{aligned}&\|x_{n+1}-u_{n+1}\|^2\\&\le\frac{(1-a_n)^2}{1-2a_n-a_nL(1+a_n(1+2L))}\|x_n-u_n\|^2+\frac{a_nL(1+a_n(1+2L))}{1-2a_n-a_nL(1+a_n(1+2L))}\\&-\frac{2a_n}{1-2a_n-a_nL(1+a_n(1+2L))}\Phi(\|x_{n+1}-u_{n+1}\|)\end{aligned}\tag{2.4}$$

Since $2a_nk_n+2a_n\gamma_nL+a_nL\|u_{n+1}-u_n\|\to 0$ as $n\to\infty$, then there exists N such that

$$2a_nk_n+2a_n\gamma_nL+a_nL\|u_{n+1}-u_n\|\le\frac{1}{2},\forall n>N.i.e.,1>1-2a_nk_n-2a_n\gamma_nL-a_nL\|u_{n+1}-u_n\|\ge\frac{1}{2}(n>N).$$

Thus, we have

$$\begin{aligned}&\|x_{n+1}-u_{n+1}\|^2\\&\le\|x_n-u_n\|^2+2a_n\frac{\alpha_n+(k_n-1)+\gamma_nL+L\|u_{n+1}-u_n\|}{1-2a_nk_n-2a_n\gamma_nL-a_nL\|u_{n+1}-u_n\|}\|x_n-u_n\|^2+2a_n\frac{L\gamma_n\|u_n-q\|^2+L\|u_{n+1}-u_n\|}{1-2a_nk_n-2a_n\gamma_nL-a_nL\|u_{n+1}-u_n\|}\\&-\frac{2a_n}{1-2a_nk_n-2a_n\gamma_nL-a_nL\|u_{n+1}-u_n\|}\Phi(\|x_{n+1}-u_{n+1}\|)\\&\le\|x_n-u_n\|^2+\frac{2a_nB_n}{1-2a_nk_n-2a_n\gamma_nL-a_nL\|u_{n+1}-u_n\|}\|x_n-u_n\|^2+\frac{2a_nC_n}{1-2a_nk_n-2a_n\gamma_nL-a_nL\|u_{n+1}-u_n\|}\\&-\frac{2a_n}{1-2a_nk_n-2a_n\gamma_nL-a_nL\|u_{n+1}-u_n\|}\Phi(\|x_{n+1}-u_{n+1}\|)\\&\le\|x_n-u_n\|^2+4a_nB_n\|x_n-u_n\|^2+4a_nC_n-2a_n\Phi(\|x_{n+1}-u_{n+1}\|)\end{aligned}\tag{2.5}$$

where $B_n=\alpha_n+(k_n-1)+\gamma_nL+L\|u_{n+1}-u_n\|, C_n=L\gamma_n\|u_n-q\|^2+L\|u_{n+1}-u_n\|.$

Set $\inf_{n\ge N}\frac{\Phi(\|x_{n+1}-u_{n+1}\|)}{1+\|x_{n+1}-u_{n+1}\|^2}=\lambda.$ Then $\lambda=0.$ If it is not the case, assume that $\lambda>0.$ Let $0<\gamma<\min\{1,\lambda\}$, then $\frac{\Phi(\|x_{n+1}-u_{n+1}\|)}{1+\|x_{n+1}-u_{n+1}\|^2}\ge\gamma,$ i.e., $\Phi(\|x_{n+1}-u_{n+1}\|)\ge\gamma+\gamma\|x_{n+1}-u_{n+1}\|^2\ge\gamma\|x_{n+1}-u_{n+1}\|^2.$

Thus

$$\begin{aligned}\|x_{n+1}-u_{n+1}\|^2&\le\frac{1+4a_nB_n}{1+2a_n\gamma}\|x_n-u_n\|^2+\frac{4a_nC_n}{1+2a_n\gamma}\\&=(1-a_n\frac{2\gamma-4B_n}{1+2a_n\gamma})\|x_n-u_n\|^2+\frac{4a_nC_n}{1+2a_n\gamma}.\end{aligned}\tag{2.6}$$

By $a_n,B_n\to 0$ as $n\to\infty$, we choose $N_1>N$ such that $\frac{2\gamma-4B_n}{1+2\alpha_n\gamma}>\gamma$, for all $n>N_1$.

It follows from (2.6) that

$$\|x_{n+1}-u_{n+1}\|^2\le(1-a_n\gamma)\|x_n-u_n\|^2+\frac{4a_nC_n}{1+2a_n\gamma},$$

for all $n > N_1$. By Lemma 1.6, then $\|x_{n+1} - u_{n+1}\| \to 0$ as $n \to \infty$, which is a contradiction and so $\lambda = 0$. Consequently, there exists an infinite subsequence such that $\|x_{n_j+1} - u_{n_j+1}\| \to 0$ as $j \to \infty$. Next, we want to prove that $\|x_{n_j+m} - u_{n_j+m}\| \to 0$ as $j \to \infty$ by induction. Let $\forall \varepsilon \in (0,1)$, choose $n_j \to N$ such that $\|x_{n_j+1} - u_{n_j+1}\| < \varepsilon$, $B_{n_j+1} < \dfrac{\Phi(\varepsilon)}{8(1+\varepsilon^2)}$, $C_{n_j+1} < \dfrac{\Phi(\varepsilon)}{8}$. First we want to prove $\|x_{n_j+2} - u_{n_j+2}\| < \varepsilon$. Suppose it is not this case. Then $\|x_{n_j+2} - u_{n_j+2}\| \geq \varepsilon$, this implies that $\Phi(\|x_{n_j+2} - u_{n_j+2}\|) \geq \Phi(\varepsilon)$.

Using the formula (2.5), we obtain the following estimates

$$\begin{aligned} &\|x_{n_j+2} - u_{n_j+2}\|^2 \\ &\leq \|x_{n_j+1} - u_{n_j+1}\|^2 + 4a_{n_j+1}B_{n_j+1}\|x_{n_j+1} - u_{n_j+1}\|^2 \\ &+4a_{n_j+1}C_{n_j+1} - 2a_{n_j+1}\Phi(\|x_{n_j+2} - u_{n_j+2}\|) \\ &< \varepsilon^2 - a_{n_j+1}\Phi(\varepsilon) \leq \varepsilon^2 \end{aligned} \tag{2.7}$$

which is a contradiction. Hence $\|x_{n_j+2} - u_{n_j+2}\| < \varepsilon$. Assume that it holds for $m = k$. Then by the argument above, we easily prove that it holds for $m = k+1$. Hence, we obtain $\|x_n - u_n\| \to 0$ as $n \to \infty$. Owing to $\|u_n - q\| \to 0$ as $n \to \infty$. From the inequality $0 \leq \|x_n - q\| \leq \|x_n - u_n\| + \|x_n - q\|$, we get $\|x_n - q\| \to 0$ as $n \to \infty$. This completes the proof.

References

[1] S.S.Chang, Some results for asymptotically pseudocontractive mappings and asymptotically nonexpansive mappings, proc.Amer.math.Soc.vol.129(2001) p.845-853

[2] Z.Q. Xue, On the convergence between the main iteration and ishikawa iteration for the generalized lipschitzian and Φ-strongly pseudocontractive mapping, Bull.Korean Math.Soc. vol.45 (2008), p. 635-644

[3] X. Weng, Fixed point iteration for local strictly pseudocontractive mapping. proc. Amer. math. Soc, vol.113

[4] E. U. Ofoedu, Strong convergence theorem for uniformly L-Lipschitzian asymptotically pseudocontractive mapping in real Banach space, J. Math. Anal. Appl, vol.321,(2006) p.722-728,

[5] Z. Q. Xue, H. Y. Zhou and Y. J. Cho, Iterative solutions of nonlinear equations for m-accretive operators in Banach spaces, J. Nonlinear and Convex Analysis, 1(3) (2000), p.313-320

doi:10.4028/www.scientific.net/AMM.50-51.723

Pricing Option on Jump Diffusion and Stochastic Interest Rates Model*

Bo Peng[1,a] , Zhihui Wu[2,b]

[1]Institute of Mathematics, Zhejiang Wanli University, Ningbo 315100, China

[2]Applied Science College, Harbin University of Science and Technology, Harbin 150080, China

[a]email: pb_1020@163.com, [b]email: wuzhihui19820905@163.com

Keywords: jump diffusion; martingale; European bi-direction option; Stochastic Interest Rates

Abstract: This paper assumed that the stock price jump process for a special kind of renewal jump process, that is incident time interval for independent and subordinate to Gamma distribution random variable sequence. We obtain the European bi-direction option pricing formulas on jump diffusion model under the stochastic interest rates by simply mathematical induce by means of martingale method.

Introduction

As we know, Black and Scholes in 1973 published famous papers entitled "The Pricing of Options and Corporate Liabilities", and then deduced the famous Black-Scholes option pricing formula [1]. In 1976, jump-diffusion model was introduced by Merton, and he derived jump diffusion model of option pricing formula [2]. Since then, more and more scholars began research in this topic [3-6]. This paper assumes the process is more general than Poisson jump process - a special kind of renewal jump process. And on the basis consider the random lending rate, which satisfies Ito stochastic differential equations. We know that in actual market transactions, risk-free interest rate adjustment will be impact on the stock market to a certain extent. Thus, we assume the random factors affecting interest rates and stock price are related. Moreover, by means of martingale method, we derived the model of bi-direction European option pricing formula.

Definition 1. $(T_i)_{i\geq 0}$ i.i.d~ $\Gamma(a,\lambda)$ $(a>0,\lambda>0)$, let $\tau_n=\sum_{i=1}^{n}T_i$, then the counting process $N_t=\sup\{n:\tau_n\leq t,t\geq 0\}$ is a special kind of renewal process.

Lemma 2. If $(N_t)_{t\geq 0}$ is a special kind of renewal process as definition 1 says, then

$$P(N_t=n)=\frac{\lambda^{na}}{\Gamma(na)}\int_0^t x^{na-1}e^{-\lambda x}dx-\frac{\lambda^{(n+1)a}}{\Gamma((n+1)a)}\int_0^t x^{(n+1)a-1}e^{-\lambda x}dx,\quad n=0,1,... \qquad (1)$$

when $a\in \mathrm{N}^+$,

$$P(N_t=n)=\sum_{s=1}^{a}\frac{(\lambda t)^{(n+1)a-s}}{((n+1)a-s)!}e^{-\lambda t} \quad , \qquad n=0,1,... \qquad (2)$$

especially $a=1$,

*The work was supported by Zhejiang Province Natural Science Foundation (Grant No.Y6100021) of China and Ningbo Natural Science Foundation (Grant No.2009A610082) of China.

$$P(N_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad , \qquad n = 0,1,... \tag{3}$$

Lemma 3. If $X \sim N(\mu, \sigma^2)$,we have $E(e^X I_{\{X \geq a\}}) = e^{\mu + \frac{\sigma^2}{2}} \Phi(\frac{-a + \mu + \sigma^2}{\sigma})$, in which $\Phi(\cdot)$ is distribution function of standard normal distribution.

Financial market model assumptions

We consider now a two assets of the financial markets $(B(t), S(t))$: One for continuous transaction risk assets (stock), the price process $S(t)$ in σ - algebra flow $(F_t)_{t \geq 0}$ in the probability space (Ω, F, F_t, P) satisfies stochastic differential equations (SDE) as follows:

$$\frac{dS(t)}{S(t)} = \mu(t)dt - vd\sum_{n=0}^{\infty} nP_n(t) + \sigma(t)dW(t) + UdN_t \tag{4}$$

Here $\mu(t)$ is yield rate of stocks.

Let $\gamma(t) = \frac{\mu(t) - r(t)}{\sigma(t)}$, $\frac{dP^*}{dP}\Big|F_T = \exp\left\{-\int_0^T \gamma(t)dW(t) - \frac{1}{2}\int_0^T \gamma^2(t)dt\right\}$, then P^* is the equivalent martingale with P , we can know from Girsanov theorem $W_1(t) = W(t) + \int_0^t \gamma(s)ds$ is standard Brownian motion under P^* -measure, then(4)is

$$\frac{dS(t)}{S(t)} = r(t)dt - vd\sum_{n=0}^{\infty} nP_n(t) + \sigma(t)dW_1(t) + UdN_t \tag{5}$$

The other is risk-free asset (bonds). Its price process satisfies to differential equation

$$\frac{dB(t,T)}{B(t,T)} = r(t)dt + \delta(t,T)dW_2(t) \ , \quad B(T,T) = 1 \tag{6}$$

Here, $r(t)$ may be short-term random interest rate, $\sigma(t)$ is volatility rate of stocks price without jump diffusion, $\delta(t,T)$ is volatility rate of bonds, $W_1(t)$ and $W_2(t)$ are standard one-dimensional Brownian motions in the probability space (Ω, F, F_t, P), and $\rho(W_1(t), W_2(t)) = \rho$, $0 \leq |\rho| \leq 1$. N_t is renewal process as defined above; $U(U > -1)$ is random variable, that is the relative high jump in the stock price when renewal jump happened. Also, $\ln(1+U) \sim N(\ln(1+v) - \frac{1}{2}\omega^2, \omega^2)$, in which, $v = E(U)$. We assumed U , N_t , $W_i(t)$, $(i = 1,2)$ are

independent， $vd\sum_{n=0}^{\infty} nP_n(t)$ is the average increasing caused by renewal jump diffusion.

Option pricing formula under jump diffusion and stochastic interest rates model

Given stock price process $\{S(t), 0 \le t \le T\}$, this means that at time t the stock price is $S(t)$. Suppose K is agreed strike price, then at time t bi-direction option payoff is $f(S(T)) = |S(T) - K| = (S(T) - K)^+ + (K - S(T))^+$.

Theorem 4. Suppose $S(t)$, $B(t,T)$ satisfy SDE（5）and (6) respectively，we can obtain the pricing of European bi-direction option at time 0 $V(0, S(0))$ for expiry date T and strike price K:

$$V(0,S(0)) = \sum_{k=0}^{\infty} P_k(T)[S(0)(1+v)^k e^{-v\sum_{n=0}^{\infty} nP_n(T)}(2\Phi(d_1)-1) - KB(0,T)(2\Phi(d_2)-1)]$$

$$d_1 = \frac{\ln\frac{S(0)(1+v)^k}{KB(0,T)} - v\sum_{n=0}^{\infty} nP_n(T) + \frac{1}{2}(\int_0^T \theta^2(t)dt + k\omega^2)}{\sqrt{\int_0^T \theta^2(t)dt + k\omega^2}}, \quad d_2 = d_1 - \sqrt{\int_0^T \theta^2(t)dt + k\omega^2},$$

$$\theta^2(t) = \sigma^2(t) + \delta^2(t,T) - 2\rho\sigma(t)\delta(t,T) \tag{7}$$

Proof For the stock price process $S(t)$ and the bond price process $B(t,T)$ satisfy SDE(5)and (6) respectively, then let

$$Z(t) = \frac{S(t)}{B(t,T)}, \quad \beta(t) = \delta^2(t,T) - \rho\sigma(t)\delta(t,T)$$

From Ito formula, we have

$$\frac{dZ(t)}{Z(t)} = \beta(t)dt + \sigma(t)dW_1(t) - \delta(t,T)dW_2(t) + UdN_t - vd\sum_{n=0}^{\infty} nP_n(t) \tag{8}$$

$$\frac{B(t,T)}{B(0,T)}e^{-\int_0^t r(s)ds} = \exp\left\{\int_0^t \delta(s,T)dW_2(s) - \frac{1}{2}\int_0^t \delta^2(s,T)ds\right\}, \text{ let } \frac{dQ}{dP^*} = \frac{B(t,T)}{B(0,T)}e^{-\int_0^t r(s)ds},$$

According to Girsanov theorem we know $W_2^*(t) = W_2(t) - \int_0^t \delta(s,T)ds$ is standard Brownian motion under Q-measure, then let

$$\theta^2(t) = \sigma^2(t) + \delta^2(t,T) - 2\rho\sigma(t)\delta(t,T), \quad \tilde{W}_2(t) = \int_0^t \frac{\delta(s,T)dW_2^*(s) - \sigma(s)dW_1^*(s)}{\theta(s)}ds$$

Here, $W_1^*(s) = W_1(s) - \rho\int_0^t \delta(s,T)ds$. We can obtain that $\tilde{W}_2(t)$ is standard Brownian motion under Q-measure, the simplification of SDE (8) is

$$\frac{dZ(t)}{Z(t)} = -\theta(t)d\tilde{W}_2(t) + UdN_t - vd\sum_{n=0}^{\infty} nP_n(t) \tag{9}$$

We can obtain the solution of SDE (9) by means of Dolease-Dade index formula:

$$Z(T) = Z(0)\exp\left\{-\frac{1}{2}\int_0^T \theta^2(t)dt - v\sum_{n=0}^{\infty} nP_n(T) - \int_0^T \theta(t)d\tilde{W}_2(t) + \sum_{i=0}^{N_T}\ln(1+U_i)\right\} \tag{10}$$

For no-arbitrage theory and martingale method in option pricing, we can know:

$$V(0,S(0)) = E^{P^*}[e^{-\int_0^T r(t)dt}(S(T)-K)^+ \,|\, F_T] + E^{P^*}[e^{-\int_0^T r(t)dt}(K-S(T))^+ \,|\, F_T]$$

$$= C(0,S(0)) + P(0,S(0)) \tag{11}$$

Firstly, $C(0,S(0)) = B(0,T)E^Q[S(T)I_{\{S(T)\geq K\}} \,|\, F_T] - KB(0,T)E^Q[I_{\{S(T)\geq K\}} \,|\, F_T]$

$$= \text{I} - \text{II} \tag{12}$$

denoted $X = -\int_0^T \theta(t)d\tilde{W}_2(t) + \sum_{i=0}^{k}\ln(1+U_i)$, under Q-measure, we have:

$$X \sim N(\ln(1+v)^k - \frac{1}{2}k\omega^2, \int_0^T \theta^2(t)dt + k\omega^2)$$

$$\text{I} = S(0)\sum_{k=0}^{\infty} P_k(T)e^{-\frac{1}{2}\int_0^T \theta^2(t)dt - v\sum_{n=0}^{\infty} nP_n(T)} \cdot E^Q[e^X I_{\left\{X\geq \ln\frac{K}{Z(0)} + \frac{1}{2}\int_0^T \theta^2(t)dt + v\sum_{n=0}^{\infty} nP_n(T)\right\}}]$$

$$= S(0)(1+v)^k \sum_{k=0}^{\infty} P_k(T)e^{-v\sum_{n=0}^{\infty} nP_n(T)}\Phi(d_1) \tag{13}$$

$$\text{II} = KB(0,T)E^Q[I_{\left\{-\int_0^T \theta(t)d\tilde{W}_2(t) + \sum_{i=0}^{N_T}\ln(1+U_i)\geq \ln\frac{K}{Z(0)} + \frac{1}{2}\int_0^T \theta^2(t)dt + v\sum_{n=0}^{\infty} nP_n(T)\right\}}]$$

$$= KB(0,T)\sum_{k=0}^{\infty} P_k(T)\Phi(d_2) \tag{14}$$

$$C(0,S(0)) = \sum_{k=0}^{\infty} P_k(T)[S(0)(1+v)^k e^{-v\sum_{n=0}^{\infty} nP_n(T-t)}\Phi(d_1) - KB(0,T)\Phi(d_2)] \tag{15}$$

Similarly, we deduce:

$$P(0,S(0)) = \sum_{k=0}^{\infty} P_k(T)[KB(0,T)\Phi(-d_2) - S(0)(1+v)^k e^{-v\sum_{n=0}^{\infty} nP_n(T-t)}\Phi(-d_1)] \tag{16}$$

Combining the results in equation (11), (15), (16), the price of European bi-direction option is found to be the right of equation (7).

Remark 5. This paper considers the incident time interval for independent and identical subordinate to Gamma distribution $\Gamma(a,\lambda)$ random variable sequence. When $a=1$, $\Gamma(1,\lambda)$ is an exponential

distribution, in which case the renewal jump diffusion model is the famous Poisson jump diffusion. Also, when $r(t)$ and $\sigma(t)$ are constant, we have $\delta(t,T) \equiv 0$. Moreover, we have $B(t,T) = e^{-r(T-t)}$.

Acknowledgment
The author sincerely thanks Professor Du Xueqiao and anonymous referees.

References

[1] F.Black, M.Scholes: The Pricing of Options and Corporate Liabilities, edited by Journal of Political Economy (1973).

[2] R.C.Merton: Option pricing when underlying stock return are discontinuous, edited by Journal of Economy(1976).

[3] Margrabe W: The Value of Option to Exchange one Asset for another, edited by Journal of Finance, Vol.1 (1978), p.177-186.

[4] Martin Schweizer: Option heading for semi-martingales, edited by Stochastic processes and their Applications, Vol.37 (1991), p. 339-360.

[5] Knut K,Aase: Contingent claims valuation when the security price is combination of an ito process and a random point process, edited by Stochastic processes and their Applications, Vol.28 (1998),p. 185-220.

[6] Chan T: Pricing contingent claims on stocks driven by Levy processes, edited by Annals of Appl Prob, Vol.9(1999),p. 504-528.

[7] Marek Musiela, Marek Rutkowski: Martingale Methods in Financial Modelling, New York, Springer, (2007).

[8] Steven E.Shreve: Stochastic Calculus for Finance Continuous-Time Models, New York, Springer, (2007).

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.728

STUDY ON THE EFFECT FACTORS OF NON-PERORMANCE LOAN RATIO OF CHINESE COMMERICAL BANKS

Ping Li[1, a], Mingying zhuo[1, b], Lichao Feng[2, c], Rui zhang[1, d]

[1] School of mathematics and statistics, Huazhong university of science and technology, P.R.Cinna

[2] College of Science, Hebei Polytechnic University, NO.46 Xinhua West Street, Tangshan 063009, Hebei Province, China

[a]pingli.hust@gmail.com, [b]zmyl05020119@163.com, [c]flch19820520@yahoo.cn, [d]zhangrui20090110@163.com

Keywords: Non-performance loan ratio; principal component analysis; SPSS16.0

Abstract. Non-performance loan ratio is one of the important assessment criteria of the security of credit assets. It is also an important financial indicator to evaluate the general strength of commercial banks. Using principal component analysis method and statistical software SPSS16.0 and based on the non-performance loan ratio and relative data of some commercial banks in China in 2007, this paper provided a principal component analysis model for the non-performance loan ratio of China's commercial banks. The factors that affect the non-performance loan ratio were refined in this paper. Finally, the characteristics of effect factors of each bank were analyzed and compared in detail.

1. Introduction

As China's financial reform in full swing, the competition among commercial banks in China is becoming more and more intense. Non-performance loan (NPL) ratio, the proportion of NPLs accounting for the total loan, is an important financial assessment indicator of bank's overall competitiveness|. The level of NPL ratio of financial institutions directly related to the size of risk of recovery of loan. In accordance with its risk-based loans is divided into normal, special mention, substandard, doubtful and loss five categories (i.e. five levels), of which the latter three collectively known as non-performing loans.

The studies of ratio of NPLs is always the important task of financial world, In China, WANG Jihen and NIE Qingping analyzed the complex courses of the high ratio of NPLs in Chinese commercial banks [1,2]. YANG Pengpeng, YUAN Zhiping and NI Haijiang tested the four assumptions on the relation of NPL ratio and efficiency in commercial banks [3]. Xie Bing analyzed the macroeconomic factors of non-performing loan scale of commercial banks[4]. And at abroad, Tarron Khemraj and Sukrishnalall Pasha attempt to ascertain the determinants of non-performing loans in the Guyanese banking sector using a panel dataset and a fixed effect model similar to Jimenez and Saurian [5]. Guonan Ma and Ben S.C.Fung propose using asset management companies to resolve non-performing loans in China [6]. Anne Harrison analyzed how the non-performing loans affected FISIM [7]. Dan Rosen and David Saunders used analytical methods to hedge systematic credit risk with linear factor portfolios [8].

But, there is relative lack of the research about effect factors of NPL ratio at present. This paper aims to refine the factors affecting the ratio of NPLs of Chinese commercial banks, based on

statistical data and Principal component analysis method. Then a specific analysis and comparison on the effect factors of NPL's ratio are carry out for some banks.

2. Modeling

We would establish the model about NPL ratio of Chinese commercial banks basing on the data of 2007. the banks studied conclude Industrial and Commercial Bank of China, China Construction Bank, Bank of China, Shanghai Pudong Development Bank, HuaXia Bank, China Minsheng Bank , China Merchants Bank, Bank of Nanjing, Industrial Bank, Bank of Beijing, Bank of Communications, China CITIC Bank, Shenzhen Development Bank and Bank of Ningbo.

2.1 Principal Component Analysis Model

Principal Component Analysis rearranges the original interrelated variables into several extraneous aggregate variables, which reacts the most of information of the original variables. Its mathematical model is as follows:

$$\begin{cases} F_1 = a_{11}ZX_1 + a_{21}ZX_2 + \cdots\cdots + a_{p1}ZX_p \\ F_2 = a_{12}ZX_1 + a_{22}ZX_2 + \cdots\cdots + a_{p2}ZX_p \\ \cdots\cdots \\ F_p = a_{1m}ZX_1 + a_{2m}ZX_2 + \cdots\cdots + a_{pm}ZX_p \end{cases}$$

here $a_{1i}, a_{2i}, \cdots\cdots, a_{pi} (i = 1, \cdots\cdots, m)$ are eigenvectors of covariance matrix $\sum$ of original variables $X_1, X_2, \ldots, X_p$. $ZX_1, ZX_2, \cdots\cdots, ZX_p$ are the standardized values of original variables.

2.2 Principal Component Analysis of Effect Factors of NPLs Ratio of Commercial Banks

2.2.1 Variables Selecting

We select 8 variables as follows:

X_1—Rate of capital sufficiency(%); X_2 --Core capital adequacy ratio (%); X_3 --NPL Provisioning Coverage Ratio (%), X_4 --Normal amount, X_5 --Concerned amount, X_6 -- Secondary amount, X_7 -- Suspicious amount, X_8 --Loss amount.

2.2.2 Data Standardizing

Because of the differences of variables dimension, we must eliminate the affection of dimension before the calculation, so we standardize the original data using statistical software SPSS16.0.

2.2.3 Relativity of Variables Judging

According to the definition of principal component analysis, it is not necessary to do a principal component analysis if there is no relativity between variables. And if one of the variables is non-associative to others, it should be out of the principal component analysis.

2.2.4 Determining the Number of Principal Component

The number of extracted principal components is *m* means that there are *m* principal components whose characteristic values are greater than 1. In a way the characteristic value can be seen as the index reflecting the intensity of the principal components If it is less than 1, we can think that the principal component is inferior to its original variable.

2.2.5 Identifying and Naming the Principal Components

Run the program "Factor" in the statistical software SPSS16.0, take principal component analysis for the factors that affect non-performing loans ratio.

The output is as follows.

Correlation Matrix

		X1	X2	X3	X4	X5	X6	X7	X8
Correlation	X1	1.000	.984	.362	-.183	-.123	-.183	-.142	-.154
	X2	.984	1.000	.355	-.167	-.106	-.158	-.118	-.137
	X3	.362	.355	1.000	-.305	-.283	-.337	-.311	-.292
	X4	-.183	-.167	-.305	1.000	.981	.962	.994	.957
	X5	-.123	-.106	-.283	.981	1.000	.935	.987	.937
	X6	-.183	-.158	-.337	.962	.935	1.000	.969	.954
	X7	-.142	-.118	-.311	.994	.987	.969	1.000	.946
	X8	-.154	-.137	-.292	.957	.937	.954	.946	1.000

Table 1 Correlation Matrix

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.064	63.297	63.297	5.064	63.297	63.297
2	2.039	25.493	88.790	2.039	25.493	88.790
3	.734	9.175	97.966			
4	.084	1.050	99.016			
5	.049	.613	99.629			
6	.016	.199	99.828			
7	.013	.160	99.988			
8	.001	.012	100.000			

Extraction Method: Principal Component Analysis.

Table 2 Total Variance Explained

Table 2 shows that two principal components should be extracted from the variables. By calculating, we get the coefficients of the linear expression of every principal component. The coefficients of the variables of the first principal component are: -0.13,-0.12,-0.18, 0.44, 0.43, 0.43, 0.44, 0.43. Those of the second principal component are: 0.65, 0.66, 0.30, 0.09, 0.13, 0.08, 0.11, and 0.10.

The expression of the first principal component is

$$F_1 = -0.13ZX_1 - 0.12ZX_2 - 0.18ZX_3 + 0.44ZX_4 + 0.43ZX_5 + 0.43ZX_6 + 0.44ZX_7 + 0.43ZX_8. \quad (1)$$

The expression of the second principal component is

$$F_2 = 0.65ZX_1 + 0.66ZX_2 + 0.30ZX_3 + 0.09ZX_4 + 0.13ZX_5 + 0.08ZX_6 + 0.11ZX_7 + 0.10ZX_8. \quad (2)$$

The positive coefficients are larger than the absolute value of the negative coefficients in the formula (1), so the first principal component mainly reflect the information of factors X_4, X_5, X_6, X_7, X_8. And the positive coefficients are very close, meaning that these five factors have a

similar degree of affecting the non-performing loans ratio. We call F_1 as *the first principal component loans composition factor*. Because only ZX_1 and ZX_2 have lager positive coefficients in the formula (2), so the second principal component mainly reflects the information of the factor X_1 and X_2. We call F_2 as *second principal component capital adequacy ratio factor*.

We get the principal component value of each bank by calculation with data. The results show that the first principal component value and the value of non-performing loan rate can not correspond fully. Therefore, the first principal component does not fully reflect the effect factors of non-performing loan ratio of the factors.

Take the first principal component F_1 as the lateral axis, and the second one F_2 as vertical axis, we get a rectangular plane coordinate system. The principal component values of each bank are a pair of coordinates that is the point of the coordinate plane. The points of principal component values of 14 banks are shown in Figure 1. It can be seen from Figure 1, the principal component values of Industrial and Commercial Bank, Construction Bank and Bank of China, the three principal component values of state-owned banks, are close. This is mainly because they have the same system. Bank of Communications and China Merchants Bank's principal component points are closer to the points of three state-owned banks, mainly due to their natures and scales are relatively close to those of three state-owned banks. The locations of Bank of Nanjing, Bank of Beijing and Bank of Ningbo are very close to each other, and the first principal component affects their non-performing loan ratio less than the second principal component. The first principal component have little effect on the non-performing loans ratio of China Minsheng Bank, Shanghai Pudong Development Bank, HuaXia Bank, Shenzhen Development Bank, and Industrial Bank, indicating these banks have good risk assessment systems. So the rate of non-performing loans of these banks are mainly affected by the second principal component.

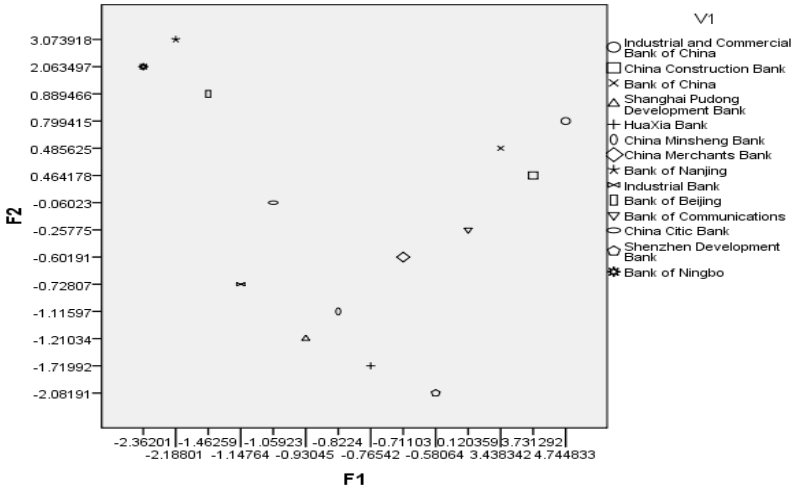


Figure1 Principal component coordinate chart of the banks in 2007

3. Conclusion

Based on the above analysis, we draw the following conclusions:

(1) Using principal component analysis method and based on relative actual data, we refined two principal components of the effect factors of NPL ratio for China's commercial banks. The first principal component, the loans composition factor, which collected the information on the levels of loan amount, is the main component factor affecting the non-performing loan ratio. And the second principal component, as the capital adequacy ratio factor, contains the information on

capital adequacy ratios, which is a secondary component factor.

(2) We introduce a principle component diagram in the analysis of the effect factors of the NPL ratio in China's commercial banks. With this diagram, we can classify the considered banks according to their principle component values. Furthermore, we provide a comparison and analysis for the differences between various classes of the banks and the causes of these differences.

(3) The Results and conclusions obtained of this paper are consistent with the actual, so they are reasonable and effective. As a useful reference and reference, they can be used in banks and other financial institutions.

References

[1] Jihen Wang: The analysis on courses to NPLs of state-owned commercial banks and advice on dealing with it, Business Studies.No.18(2006), p. 159-160.

[2] Qingping Nie: Advice on NPLs and reform policy of Chinese banks, Economic Science, No.3(2002).

[3] Pengpeng Yang, Yuan Zhiping and Ni Haijiang: The causality-test of nonperforming loans and efficiency in china's commercial banks, Academic Journal of Xi'an Jiaotong University (Social Sciences), Vol. 28(2008), p. 30-34.

[4] Bing Xie: Analysis on Macroeconomic Factors of Non-performing Loan Scale of Commercial Banks, The Theory and Practice of Finance and Economics, vol.30(2009), p. 23-25.

[5] Tarron Khemral and Sukrishnalall Pasha: The determinants of non-performing loans: an econometric case study of Guyana .

[6] Guonan Ma and Ben S.C.Fung: Using asset management companies to resolve non-performing loans in China, Industrialization.

[7] Anne Harrison: Non-performing Loans--Impact on FISIM, Fourth meeting of the Advisory Expert Group on National Accounts 30, Frankfurt(2006).

[8] Dan Rosen and David Saunders: Analytical methods for hedging systematic credit risk with linear factor portfolios, Journal of Economic Dynamics & Control33(2009), p. 37–52.

[9] Xiulin Yu and Xuesong Ren: Multivariate Statistical Analysis, China Statistical Publishing House, Beijing.

[10] Hongbing Zhang, Laixi Jia and Lu Li: A Treasury of Knowledge about SPSS, Publishing House of Electronics Industry, Beijing.

doi:10.4028/www.scientific.net/AMM.50-51.733

Research on Generalized Green Computing

Juanli Hu[1,a], Jiabin Deng[1,b],Jinbo Liu[1,c] Juebo Wu[2,d]

[1] No.25, boaiqi road, East District, Zhongshan Polytechnic, Zhongshan, China

[2] Shenzhen Angelshine Co., LTD, Shenzhen, China

[a]hjlfoxes@163.com, [b]hugodunne@yahoo.com.cn, [c]hugodunne24@gmail.com, wujuebo@gmail.com

Keywords: Generalized Green Computing; Data Storage; Application Virtualization; Cloud Computing

Abstract. Green, energy-saving, environmental protection and sustainable development have become the most popular buzzword and topic in science research. This paper presents the concept of generalized green computing and describes the limit and the range for it, including the research task and relationship. It analyzes and compares these researches on green computing. Problems of current topics are discussed, and finally future directions are proposed in this paper.

1 Introduction

By the end of 2009, the number of computer system all over the world (including desktop systems, portable computer systems, special embedded computer systems and server system, etc.) had been more than 310 million, and keeping rapid growth [1]. Take laptop for example, iSuppli (market research company) predicts that the sales of laptop are remaining optimistic in future years, and the sales to dealers of notebook in 2010 can reach 20.95 million with the growth by 22.5% compare with 2009. If the electric charge is $0.1 per KWH, the total cost of global computer users in the world could reach more than 800 million dollars last year.

The paper will analyze the influence on environment and energy saving and emission reduction caused by computer systems. With the development and current situation of green computing, it will make sure its direction and analyze the component of generalized green computing, including the discussion and summarization of the latest, effective green computing technology and solutions.

2. The background of green computing

Green computing is rising in pursuit of the computer system performance under the premise of rapid development and improving the environment, continuing to improve the quality of life. The background can be traced back on March 23 1950 when the convention of the world meteorological organization formally went into effect. In 1987, the UN report "Our Common Future" put forward the basic concept of Sustainable development, which admitted by the immediate environmentalists, economists and social activists.

In 1992, the U.S. Environmental Protection Agency launched Energy Star, a voluntary labeling program which is designed to promote and recognize energy-efficiency in monitors, climate control equipment, and other technologies [2]. The Energy Star is the milestone of when the computer system put into energy saving and emission reduction and it also formed the prototype of green computing for the future study [3].

3. Major methods of Green Computing

3.1 Physical Devices Improvement. Physical equipment is one of the most important composition factors in green calculation, which relates to energy conservation, emission reduction and performance. Fig. 1 shows the distribution of Desktop System. Central Processing Unit (CPU) is one of the most important hardware in the computer system. Many CPU firms take measures on several levels to power optimization such as layout, circuit, door, RTL, microstructure level, etc, and more and lower consumption products are introduced like Intel® Atom™ Processor N410 Series (1.6 GHz). Such products not only reduces the dominant frequency and heat, but also adopts Hyper-Threading where a single physical core simulates two cores and let each logical core use threads for parallel computing. In such way, multithreading operating system and software can be better compatible, lessen the idle time and improve the CPU efficiency. The biggest characteristic is to use the integration technology, integrated graphics and store controller: Integrated Intel® multimedia accelerator 3150 in combination with integrated storage controller can enhance performance in the response of the system and the ability, further reducing the energy consumption card instead of graphics [4].

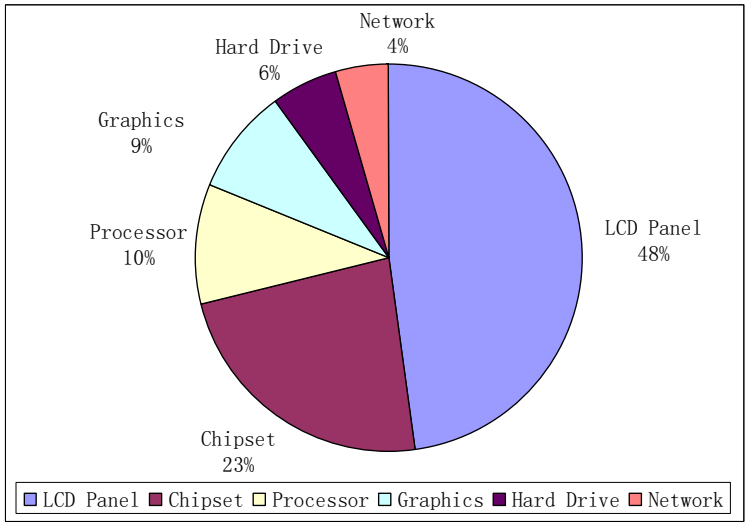


Figure 1. The energy distribution of Desktop

3.2 Data Center Improvement. Now the most pivotal problem of data center is the high consumption and high calorific value [5], as shown in Fig. 2.

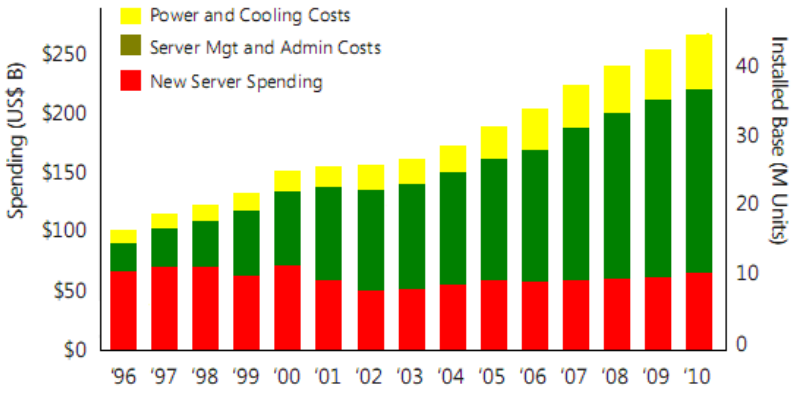


Figure 2. Rate of Server Management, Power and Cooling Cost Increase

According to a 2007 EPA report, U.S. data centers alone consumed 61 billion kWh in 2006- enough energy to power 5.8 million average households. Even under conservative estimates, IT energy consumption is projected to double by 2011. Reducing data center energy consumption is a pressing issue for the entire IT industry now and into the future. In this section, we argue that dense and real-time environmental monitoring systems are needed to improve the energy efficiency of IT facilities. Industry even emerges working at cooling in 1Watt and 2Watt.

3.3 Portable Computing. The hand-held mobile devices have been more extensive applications in nowadays such as mobile phone and PDA. The handheld mobile computing has gradually become the mainstream calculation mode of today. Currently, the handheld mobile equipments have got the supports from Intel, Texas Instruments and QUALCOMM (R) etc. The dominant frequency can be over 600MHz while RAM can achieve 256M~512M. By extension, the storage space can achieve 16Gb~32Gb or even more. Operating system and the corresponding software are more powerful enough to meet the daily requirement in simple offices, and the price also continues to reduce. Recently, according to the portable office and entertainment, each big mainstream companies have produced the corresponding tablet (such as iPad). Such products own the similar performance and function as ordinary laptops, but the energy and heat are well below the average laptops.

3.4 Recycling of Electronic Waste. Used-equipment dumping is perhaps the most legislated and litigated green IT issue. Many countries now regulate electronic-waste disposal, including computers, printers, video gaming systems, monitors, and so on. Seventy percent of all pollutional waste is e-waste, which is bulky, complicated to recycle, and most of them contain unsafe levels of heavy metals and other dangerous chemicals. The current US congressional session is considering significant legislation aimed at a uniform approach that would unscramble the various e-waste regulations already implemented in nearly half the states. Mostly to developing nations with more relaxed environmental regulations, and then have man-dated e-waste recycling, and all but California require manufacturers to pay for it. Electronic equipment manufacturers and distributors are concerned that the states' differing regulations could bring uncertainty into their business plans.

3.5 Software Power Savings Modes. With software functions increasing raiply, how to reduce the complexity of the algorithm and its system has become a key issue in many systems design, mainly including reducing the functional complexity of the interface between human and machine, running time and the hardware stress in order to energy conservation and emission reduction. At present, many software design functions are too complex in computer system. Meanwhile, upgrading frequently in many companies recently has become the new reason of positive leather disadvantages to launch. Besides, many software products, such as Windows Vista OS released by Microsoft Corporation in 2006, it is difficult to replace Windows XP OS due to the complex function, hardware, less friendly interface, high requirement and high energy resources share. Other software, such as Offices, have many functions that ordinary users rarely use, even never use them, so all of these increase the purchase cost, learning cost, and Microsoft's design and development costs, which leads to the social resources waste. Because the software is a special kind of product with bodiless, this wastes do not arouse people's attention. Therefore, some enterprises take the customization design through SAAS and PAA in order to save costs and reduce complexity of software development. At the same time, by virtualization, it can increase the use of hardware and balance resources between servers, aiming to achieve better shared load balance.

3.6 New Energy Sources Consumption. American ABI Research carried out a survey to the prospect of fuel cell using small-sized in the computer industry, and the result showed that there will be 13.5% of laptop power using small-sized fuel cell in 2012. Mini fuel cell generates

electricity by fuel. The lithium ion batteries were used early in early automobile industry as low emission energy, but it is being concerned by computer industry in notebook computers and mobile phones as instead of old powers. In recent years, lithium ion batteries become a new option of computer industry. Such products can prolong battery life as 2-3 times than ever before, or even extend to 10 times in the future.

Figure 3. LG's ecological laptop

In 2005, parts of laptops and PDAs have begun to use super-mini fuel cell, for example: LG introduced a notebook with ecological concept design called as e-Book in 2006, which use fuel cell and OLED display technology, winning "best design concept" award in 2006 as shown in Fig. 3.

The battery of e-Book takes the natural gas as methanol fuel, which are stored environmental blue notebook shaft in the transparent. It can save space and make notebook more beautiful. The future notebook of LG's will adopt OLED instead of LCD panel in order to lower power consumption.

ABI Research forecasts that 2,000 super-mini fuel cell will be produced by trial in North America and Japan next year. After establishing the necessary code and standard, the number of global super-mini fuel cell will reach 1.2 billion. ABI Research said that it is expensive to add fuel system or replace fuel box, so it may hinder the small-sized popularization. The organization of Intel notebook computer batteries ascension released "Laptops fuel cell technology standard", which gave the specification for current, the manufacture, the fuel control, heat, and other conditions.

4 CONCLUSIONS & FUTURE WORK

Green computing was presented in the construction of low-carbon economy and economical society advocating. It can create social value in today's society, and it has become a consensus in energy-saving, environmental protection and saving. Green computing has become an important trend for the computer industry, and the concept of the consumer health and energy requirements have become the new computer product update standard.

In this paper, based on the comparative study of green computing technology, we adopted a more uniform and abstract manner to put forward the concept of generalized green computing, and discuss the main research contents. This paper gave the clear boundaries and scope of the green computing, and straightened out the relevant research task and relationship. We presented some basic ideas and methods of green computing for the next research under different conditions. Such

all methods are the specific model of green computing and they laid a solid foundation as a tool.

References

[1] GUO Bing, SHEN Yan, SHAO Zi-Li. The Redefinition and Some Discussion of Green Computing. Chinese Journal of Computer, 2009, 32(12):2311~2319.

[2] Turgeon, Jason. 2009 Energy star benchmarking workshop proceedings. Journal of New England Water Environment Association, 2009, 43(2): 39-43

[3] Climate Savers Computing Initiative. http://www.climatesaverscomputing.org, 2008.

[4] Intel Ltd .Intel® Atom™ Processor Z540, Z530, Z520, Z510, and Z500 on 45 nm Process Technology Datasheet. http://download.intel.com/design/processor/datashts/ 319535.pdf, 2009

[5] Wang David. Meeting green computing challenges. Proceedings of International Symposium on High Density Packaging and Microsystem Integration 2007, 2007: 1-4

doi:10.4028/www.scientific.net/AMM.50-51.738

Study on the Collective Effect of Industrial Clusters

Zhang Ci

Light Industry College of Hebei Polytechnic University, China

yehe_ci@163.com

Keywords: Industrial cluster; Collective effect; policies & countermeasures

Abstract: Industrial clusters have emerged in many parts of China. But with the development of the cluster, some industrial cluster shows its unique cluster effect and lead to the development of local economy, while others die out. Based on the above, this article studies on the conditions of industrial cluster effect.

1 Introduction

Industrial cluster is a typical phenomenon in the process of world economic development, such as American silicon valley, India's software, the Italian ceramic tile industry etc. In China, many provinces and cities also appeared a certain scale of prototype industrial clusters or industrial cluster. Industrial cluster gradually become the focus of research and academic issues. However, the collective effect of industrial clusters on promoting industrial cluster development rarely is studied by expert systems. But this is a problem that is not allow to ignore. This article discusses the police to different types of industrial cluster by the conditions of t the collective effect of industrial clusters.

2 Industrial clusters effect

Industry cluster is a special economic form in which enterprises with close relation and relevant support agencies are gathered[1]. With the share of basic facilities, market & technical information among those enterprises, good industrial soft-environment and complete innovation system could be formed. Following that, the industrial development could be promoted and the optimization of regional industrial structure can be brought. So, industrial cluster is an important source of regional competitiveness. Industrial cluster are able to promote economic development, the key is its own peculiar cluster effect:

2.1External scale economy

One of the important characteristics of industrial clusters is the geographic concentration. It shows that enterprises with close relation and relevant support agencies gather in certain place. Those enterprises could share basic facilities, technical information &market internet by geographic concentration. They could also make use of some assistant enterprises which provide parts, intermediate product, processing scraps and producing service together to lower producing and transport cost. They can make face to face conversation and communicate information to lower information searching and traffic cost. And technical innovation can be promoted while knowledge diffusion quickened. Then, external economy of scale appears.

2.2Specialized Division

In industrial clusters, many enterprises gather together with close and flexible relations of specialized division and cooperation. AThey have clear division iAAAccordance with the upstream and

downstream of the industry chain, there are clear division between enterprises. With the industrial development, it further promote the productive specialization. Producing service in the core enterprise of the industrial cluster will be separated gradually. Instead, new producing service from new enterprises gathering around the core enterprise comes. Then, the internet of production, sale, service and information with complete framework and function will be formed for the gathering of enterprises, servicing institution and government. Furthermore, specialized production and service will be provided.

2.3Collective effect

In industrial cluster, fierce market competition and kinds of cooperation take place among enterprises. As this interactive correlation, high effective internet interaction and cooperation in training, finance, technology development, production design, marketing, export and assigning can realize among them so that disadvantages of those enterprises could be overcome. The economic benefit conflict and business obstacle of can be reduced. And the collective effect can be obtained and win-win can be achieved. In addition, coexisting of cooperation and competition is helpful for enterprises to build strategic alliance and partnership and make flexible specialized production.

2.4Overflow effect

In industrial cluster, formal and informal communion is very close among enterprises for the geographic concentration. It is helpful for the spreading of new idea, opinion, technology and knowledge. Follow that, knowledge overflow effect forms. And the research and innovation mechanism can be enhanced. With the gradual innovation of enterprises, the technology progress will come and the industrial upgrading and optimization will be achieved.

3 Producing conditions of industrial clusters' collective effect

When the out-of-order competition of shoes industry in China's Jinjiang, once thriving lamps and lanterns industry in China's Wenzhou almost die and the downturn of industrial clusters in yangtze delta area occurred, People found the competitive advantage of industrial cluster is not occur in all of the cluster, people begin to think what are the producing conditions of industrial clusters' collective effect?

Actually it was referred by Michael Porter in his "diamond system theory" that relevant and supporting industries are very important for industrial clusters. If no matching relevant and supporting industries, industrial clusters will stop developing for the lack of specialized senior producing elements [2]. The importance of relevant and supporting industries can be seen obviously.

3.1Correlation of industry clusters

The correlation of industry clusters is that enterprises in industrial clusters are in the same or similar industrial chains and have forward, backward and transverse relations. For the correlations in industrial chains, it's possible that enterprises in industrial clusters can use the same raw materials and sale way of productions, have potential scale and standard demands, lower the costs of producing, storing & transporting through combing producing, sale & purchasing, use same or similar management way and marketing plan as well as develop same or correlative production technology, function & manufacturing process to make the external scale economy. The gathering of manufacturers and suppliers or enterprises in industries with close relations will strengthen the communications of knowledge, technology, information and experiences. In addition, fierce

competition can make enterprises in industrial clusters bear great pressure. And this pressure can be transferred to great innovative motivation. So enterprises will develop new products and process so that the products quality and industrial level will be enhanced. At the same time, when innovation or upgrading of products' technology, process and function take place in the certain industry especially the core correlative industry, corresponding innovation or upgrading of products' technology, process and function will occur in industries of upstream and downstream industrial chains. And then, corresponding innovation or adjustment will also happen in the other correlative industries. This is so-called "chain innovation" [3].

3.2Support of industrial clusters

Support of industrial clusters is that enterprises in industrial clusters cooperate with high consistence in producing, marketing, purchasing, management and technology. For industrial clusters, cooperative gathering corresponding supporting industries should take place with the geographic gathering of enterprises. Only interactive support among enterprises occur, it happen cooperative effect of 1+1>2 which is shown as specialized division and cooperation of production. With the need of scale economy, enterprises' departments for producing service will be separated and gather around them to provide specialized matching producing service. In addition, many new local enterprises will grow up for the complete industry system of industrial clusters. On the one hand, deepening of specialized division will weaken the industrial isomorphism and excessive competition for gathering. On the other hand, specialized service will be brought with specialized production. Advanced producing service can lower the cost of regional transaction and meet the demand of enterprises' matching production. And the enterprises' producing and managing efficiency will be enhanced and the additional value of output be added. The enterprises' competition will be enhanced. Moreover, it's much easier that the industrial competitive advantages being nurturing. Support of industrial clusters is helpful for the communion and providing of knowledge among support industries. It make different industries in industrial cluster to be the source of knowledge or may be the receiver of knowledge and technology diffusion overflow. So bidirectional flow of knowledge and technology take place among supporting industries.

At present, industrial clusters occurred in many areas in China. But collective effect does not happen in many of them which just have get graphic concentration. So industrial clusters not only need the geographic concentration, but also need the cooperative gathering corresponding supporting industries. Therefore, the key factor of industrial clusters' collective effect is the correlation and support.

4 Policies

From the above conclusions ,we know that collective effect of industrial clusters need good industrial correlation and support and their better interaction. Therefore, for formed industrial clusters which don't have good industrial correlation and support simultaneity, the concept of clusters' correlation and support should be kept to promote the better development of industrial clusters

4.1Industrial clusters with good correlation and poor support

Generally, these industrial clusters are shown as combing of several large core enterprises and small ones which have forward or backward industrial correlations. For these industrial clusters, large core enterprises are in leading position. With market changes, large enterprises may develop in an omni-directional and monopolistic way to merger other small enterprises. In contrary, large enterprises don't have advantages of small ones easy to change. It's easy for large enterprises to quit the market or be in weak position for the market turbulence. Thus, excessive competition can take

place as all enterprises are in similar positions[4]. For avoid those phenomena, non-core and uneconomical business of large enterprises should be separated to form assistant producing enterprises to provide specialized matching services or large enterprises can be made through entering of external large enterprises or combing of enterprises.

4.2Industrial clusters with poor correlation and good support

These industrial clusters are generally composed of enterprises with close relations of enterprises such as classmates, home fellows, relatives etc. Special relations can make close relations among enterprises and make them form strong economic and benefit group for preferential policies or government support. But for these clusters don't have industrial systems with specialized division and cooperation, they don't have much higher cooperation and benefit goals and can't have sustainable development motivation. The kind of cluster enterprise cluster need to establish business association and industrial association, carry on the industrial division, form supporting industry[5].

4.3Industrial clusters with poor correlation and support

These industrial clusters are formed by uncorrelated gathering enterprises. The gathering enterprises are possibly due to preferential policies, cheap rents and so on. Actually, it can't be called industrial clusters. Collective effect can't work and maybe work in converse way. These clusters must be avoided in the early building time. Industrial clusters' planning should be adjust to local situation and combine with the local industry development, then form characteristics of industrial cluster.

5 Summary

The collective effect of industrial clusters depends on relevant and supporting industries' coordinated cluster. So according to industry chain, industrial cluster which is forming or has formed choose enterprise again through the limited enterprise conditions (such as business enterprise qualification and industry orientation, etc.).It is to make sure that enterprise have strong industrial relevance and good industrial supporting, so there form the continuous development of industry cluster of dynamic mechanism.

References:

[Pote M．E．Clusters and New Economics Competition[J]．Harvard Business Review，1998，(11)．[1] Michael Porter．Clusters and New Economics Competition[J]．Harvard Business Review，1998，(11)．
[2] Michael Porter．Theory of Competition [M]. Beijing: CITIC Publishing, 2003.
[3]WangJiCi. Innovative spatial: enterprise clusters and regional development [M].beijing: Beijing university press, 2001.
[4]Fritz O M，Mahringer H，Valderrama M T．A risk-oriented analysis of regional clusters[A]．In Stei • net M(ed)．Clusters and regional specialisation[C]．London：Published by Pion Limited，1998．
[5]Low M B，Abrahamson E．Movements，bandwagons，and clones：Industry evolution and the entre-preneurial process[J]．Journal of Business Venturing，1997，12(6)：435-457．

doi:10.4028/www.scientific.net/AMM.50-51.742

Research on Discipline Evolution in Local University Based on Ant Colony Algorithm

Wang Xiaoci[1,a],Chen Liang[2,b],Zheng Caiyun[3,c]

[1,3]College of Light Industry, Hebei Polytechnic University, Tangshan, Hebei, China

[2]Beijing Decision-making Consultant Center, Beijing, China

[a]Wangxiaociyy2001@163.com,[b]18931571560@189.cn,[c]qykq@qy.heut.edu.cn

Keywords: Local University; Ant Colony Algorithm; Innovation Element of Science and Technology; Discipline Evolution

Abstract: Local Universities are an important part of higher education in China. The discipline as the core element of higher education is the basic strategy unit which supports teaching, research and social services. The development level of the discipline represents higher education development. From the perspective of technological innovation in the development of disciplines, this paper studies the discipline evolution, which based on ant colony optimization theory. We constructed a mathematical model applied to the discipline evolution.

1 Introduction

With the popularization of higher education, the number of Local universities is increasing step by step. Local universities have become an important force in national higher education system, and they undertake training high quality personnel, high-level scientific research and production services for the local economy as the key universities. Overall, the local university discipline is composed of a number of research fields, and a research field is formed by one or some innovation elements of science and technology. The innovation elements of science and technology of local universities doesn't exist inherently, and its formation has to go through identification, planning, training, and maintaining, which is a complex process. The growth of disciplines is a constantly selection, communication and evolutionary process.

2 Disciplines and Innovation Element of Science and Technology

From the view of evolutionary, this paper presents the concept of innovation elements of science and technology, which is considered as the basic science and technology innovation activities of academic units. The innovation elements of science and technology of local universities is linked by vertical or horizontal issues. It is an Academic community which is formed by Interrelated and Coordinate Individuals that aim to complete the issues.

The innovation element of science and technology is the most basic cellular that is needed by Scientific Activities of local universities. Although it is subject to the issue organization as a link, which has a one-time feature, its core members do not dissolve with the completion of the issue. With the characteristics of the path dependence, the core members will continue to follow the direction of the issue in order to form Stable academic community. They will continue to engage in the research field relevant to the research. And in long-term cooperation, assistance and support to

gradually, they form a more stable basis of the research direction. Some stable disciplines direction evolved from the innovation element of science and technology will go through the continued integration, and finally form in the usual sense discipline which has stable organizational structure and forms. Some associated, similar or complementary disciplines through cross-disciplinary integration, which coexist in colleges, schools, or other major habitats, produce the usual sense of the discipline group.

Therefore, the innovation element of science and technology has the characteristics of step-by-step development. In the development process it will continue to release outside information element which is similar to the ants wandering pheromone. Across the frozen process of pheromone the innovation element of science and technology demonstrates its research results and research track.

3 The Ant Colony theory

3.1 The principle mechanism of ant colony algorithm.

The ant colony algorithm which simulates ant foraging behaviors was introduced as a new computational intelligence model. The algorithm is based on the following basic assumptions:

A. Ants communicate with each other through pheromones and environment. Each ant is only based on its local environment surrounding, and responds only to its impact on the surrounding local environment.

B. Ants response to the environment decided by internal models. Because ants are genetically modified organisms, the behavior of ants actually is adaptability of its genes, which represent that ants are the main reactive adaptation.

C. From the view of individual, each ant makes independent choice based solely on the environment. At the population level the behaviors of a single ant are random, but the process of colony formation is highly ordered by self-organizing group behavior.

From the above assumptions, we think that the ant colony algorithm is a stochastic search algorithm similar to other evolutionary algorithm models.

Ant colony algorithm seeks the optimal solution by the evolution of group which is composed of the candidate solutions. The process consists of two stages: adaptation and collaboration. In the adaptation stage, the candidate solution continuously adjusts its structure according to the accumulated information. The path is very likely to be chosen, which is passed by more ants and produced more information. In the collaboration phase, candidate solutions exchange information between each other in order to generate a better solution. The ant colony algorithm is initially random to select search path. With understanding of the solution space, the search has become a regular, and gradually approaches and eventually reaches the global optimal solution.

3.2 The description of ant colony behaviors

The ant colony algorithm achieves the purpose of solving combinatorial optimization problems by simulating the process of searching for food. The method is inspired by the research on the real Behaviors of ant colony [1]. After extensive research by bionics home, it is found that individual ants transfer information through a substance called pheromone [2]. In the moving process, the ants left the substance in the path, and they can sense the existence and strength of the pheromone. The ants is guided the direction of their movement by the pheromone, and they tend to move the direction of movement of material strength [3]. Therefore, collective actions of the ant colony composed of a large number of ants show a positive feedback of information: The path which is

passed by more ants is very likely to be selected by later ants [4].

4 The application of Ant colony algorithm in the discipline development

By constructing the mapping relationship of the innovation element of science and technology and discipline, we unify to consider the development of the innovation element of science and technology and the process of finding the optimal. This article thinks that the development of discipline is a chaotic search process. The process of establishing the optimal development direction is the gradual transition from the chaotic search to a transient chaos and until the process of convergence to the optimal path. In other words, the subject of the development process is a self-organization process which is based on the interaction of pheromone and chaotic.

We establish that there are n innovation elements of science and technology in discipline, and these elements are in the search space S, and they want to minimize a function $f: S \to R$.Every point s in the space S is a given feasible solution of the problem. This paper only considers the space being searched is continuous (That is $S = R^l$).The position of No. i Element expressed by an algebraic variable symbols $s_i = (z_{i1}, z_{i2}, ..., z_{il})$, $i = 1, 2, ..., n$.Naturally, each variable can be any finite-dimensional. In the moving process of innovation elements of science and technology, each element is subject to the discipline's impact. In the mathematical expression, the campaign strategy of a single element is a function related to their present position, the best location that it's own or is found by partner and organizational variables. That is:

$$z_{id}(t) = g(z_{id}(t-1), p_{id}(t-1), y_i(t)) \qquad (1)$$

(1) t means the current time step, and $t-1$ means the previous step.

(2) $z_{id}(t)$ represents the current state of d-dimensional variable of No.i Element, $d = 1, 2, ..., l$. l is the dimension of search space, and $y_i(t)$ is the present state of the organization variable

(3) $p_{id}(t-1)$ stands for the best location which is found by No. i Element or its Neighbors in step $t-1$.

(4) g is a nonlinear function.

The chaotic behaviors of single element adjust and let themselves be able to eventually move to the position of the best fitness value, which is achieved through the introduction of a continuous change of variable $y_i(t)$.In order to achieve the exchange of information between elements and let them be able to eventually move to the best position, we introduce $(p_{id}(t-1) - z_{id}(t-1))$.We choose p_{id} according to the widely used optimization theory. Thus, by the following model:

$$\begin{cases} y_i(t) = y_i(t-1)^{1+r_i} \\ z_{id}(t) = z_{id}(t-1)\exp((1-\exp(-ay_i(t)))(3-\psi_d z_{id}(t-1))) + \exp(-2ay_i(t)+b)(p_{id}(t-1) - z_{id}(t-1)) \end{cases} \qquad (2)$$

a is a sufficiently large normal number, and its value can be chosen as 200. b is a constant and $0 \le b \le 2/3$.ψ_d determines the search scope of the No. d element in the search space. r_i is a normal number of less than 1,which named as organization factor of No. i Element. $y_i(0) = 0.9999$.

As previously mentioned，innovation elements of science and technology exchange information by a direct or indirect means of communication. Along with the time evolution, the effective

communication causes organization's function between each other becomes stronger and stronger. Finally, all innovation elements of science and technology are able to walk toward optimal path of the research direction. Eq. 2 expressed chaos search process. $y_i(t)$ is used to control the chaos process. At first, the impact of $y_i(t)$ is slow. Over time, the organizational variables $y_i(t)$ impact strongly on the behavior of innovation elements of science and technology with the organizational factor r_i. In the end, the state of chaotic systems state in the optimal state or sub-optimal state.

In order to make each innovation element of science and technology have different organizational factor, we limit $r_i = 0.1 + 0.2rand$ and $rand$ is a random number between 0 and 1. ψ_d affects the search scope of the ant colony algorithm. If the parameter ψ_d is small, the search scope can be very big, vice versa. Its value needs to act according to the concrete optimized question to carry on the suitable choice.

Discovered after the massive value simulation, if $\psi_d > 0$, Eq. 2 may realize the search process in the positive number sector (in this sector $z_{id} \geq 0$). If $\psi_d < 0$, Eq. 2 may realize the search process in the negative number sector (in this sector $z_{id} \leq 0$). When optimal solution is located at positive number sector or negative number sector interior time, Eq. 2 may solve the value optimization question effectively. However, all elements of optimal solution are impossible simultaneously to be located at the positive number or in the negative number sector scope. They possibly are located in all real number sector scope. In order to solve the optimized algorithm's search scope, this article has given the following more general mathematical model:

$$y_i(t) = y_i(t-1)^{(1+r_i)}$$

$$z_{id}(t) = (z_{id}(t-1) + \frac{7.5}{\psi_{id}} * V_i)\exp((1-\exp(-ay_i(t)))(3-\psi_{id}(z_{id}(t-1) + \frac{7.5}{\psi_{id}} * V_i))) - \frac{7.5}{\psi_{id}} * V_i$$

$$+\exp(-2a\mu_i(t)+b)(p_{id}(t-1) - \theta_{id}(t-1)) \qquad (3)$$

V_i had decided the search region of innovation element i, and each element can be possible to search the different in the question space. If we had known the fixed search scope, this article may establish V_i an appropriate positive value. This article calls Eq. 3 generalized model of the ant colony algorithm. In this model, the initial point of single innovation element of science and technology is: $z_{id}(0) = -7.5/\psi_d * V_i * rand$, and $\psi_d > 0$ has decided system's search scope. If ψ_d is very big ,algorithm search scope will be very small, vice versa. We suppose the search zone are $[-\omega_d/2, \omega_d/2]$, then there is an approximate relations ψ_d between ω_d and $\omega_d = 7.5/\psi_d$.

In the above generalized model， we used two random numbers. The first random number was used in the initial point of innovation elements in the choice process, namely $z_{id}(0) = -7.5/\psi_d * V_i * rand$.

The chaos system is sensitive to initial condition. We establish the different value of the initial point in order that innovation elements of science and technology can search the different position of phase space with the time evolution. The second random number was used in the expression of organization factor r_i. At the first, innovation elements carry on the chaos search. Through some kinds of direct or indirect communication ways, they exchange information with the neighbor, which are about the best position. Along with the time, organization's function is getting stronger and stronger.

5 Conclusion

The ant colony Algorithm is a good method for seeking the optimal solution. The discipline is very important object of study in higher education domain. This article profits from the ant colony principle, and conducts the research of the discipline development evolution. We proposed concept of the innovation element of science and technology.

Standing the point of discipline scientific innovation, this article introduces the optimization of ant colony algorithm, and regards the discipline developing process as an unceasing optimization process. The early behavior of the innovation element of science and technology are chaotic, and there is no regular pattern at all. With the continuous development of their own and the strong exchange of information, the innovation element of science and technology may be more intensive research direction, and they ultimately form the research or discipline.

References:

[1]Wang Zhigang:The Study of Heuristic Algorithm in Distribution Network Planning, ,Zhengzhou: Zhengzhou University, 2003

[2]Varela F and Bourgine P: Proc of the First European Conf On Artificial Life[C].Paris，France：Elsevier Publishing，1991

[3] Michael JB K，Jean-Bernard B，Laurent K.Ant-like task and recruitment in cooperative robots[J].Nature，2000，406（31）：992-995

[4]Zhao Qingjiang: Study on Vehicle Routing Problem with Complicated Non-Ergodic Characteristics, Hangzhou：Zhejiang University，2003

doi:10.4028/www.scientific.net/AMM.50-51.747

Earthquake Loss Assessment Based on Industry Association

Xiaoguang Xue[a] Yan Xiong[b] Xuchao Song[c]

College of Light Industry, Hebei Polytechnic University, Tangshan 063000, Hebei, China

[a]xuexiaoguang7@yahoo.com.cn,[b]xiongyan19556@163.com,[c]songxuchao@126.com

Keywords: Earthquake, disaster risk assessment, Industrial Relations loss, input-output table

Abstract: current research on earthquake damage assessment summarizes In this paper, the composition of the input-output tables and internal relations were introduced, given the earthquake loss assessment methods which base on industry association.

1 Introduction

Earthquake would cause serious casualties and major economic losses, particularly the devastating earthquake. These two Indexs has become the Main index to determine the degree of earthquake, the earthquake occurred in recent years show that, with the social and economic development, building code revision and improvement of infrastructure construction and other seismic performance of a number of continuously improve the earthquake casualties declining, but the absolute value of economic losses increasing, this situation especially in economically developed countries. China is one of earthquake-prone countries, the economic losses caused by the earthquake was paid of more and more attention. Earthquake not only cause direct economic losses of buildings, lifeline facilities, indoor destruction etc.,but aslo cause temporarily stop or reduce production, the original balance and the economic system in the disturbances and reduction of external investment in indirect economic losses resulting. With strengthened of the growing economic links between regions, it has been widely recognized that earthquake caused by the indirect economic losses will also increase the importance of seismic indirect economic loss, earthquake quantification of indirect economic losses have become eager to know the people issue, the reduction of earthquake hazards is also one important scientific question. But accurately estimate the indirect losses of earthquake has been a problem which placed in front of the correct journalists.Because of the earthquake damage is involved in earthquake engineering,structural engineering, socio-economics and other areas of an interdisciplinary, to solve this problem requires the joint efforts of various parties. Indirect economic losses in the earthquake related loss is the result of coordination between the various industrial sectors relations from being undermined, the formation of local productive resources (including productivity resources) caused by the accumulation of stagnant and, in the economic system to describe this interaction affect inter-relationship between the input-output tables, this earthquake related loss in the current research on the basis of the summary, given the input-output table based on earthquake related loss assessment methods.

2 Research

The late eighties of last century, Kozuhiko, Kawashima and other indirect economic losses of earthquakes are discussed in detail, mainly by the indirect effects of the earthquake. Facilities, transport facilities ,the destruction of raw materials and commodities decision are taken into account.

The impact of re-investment direct loss, based on the relationship between application of the industrial sector (input-output analysis) of the 1983 earthquake .This is made by Nihonkai-chubu indirect damage assessments[1]. Gordon, MoorShinozuka and other scholars are take into account by the indirect economic loss of the earthquake. The city is mainly due to traffic and the impact caused by industrial production capacity, and developed a seismic evaluation of the overall model of indirect economic losses [2]. Taking into account the decline in industrial capacity caused by the earthquake, the supply and demand relationship between the various departments have changed intersectoral input-output calculation of indirect losses. Huang Yu-Xiang and other disasters said taking into account the economic losses, the related loss as a direct stop production cuts by the loss and the national economy caused by an indirect chain cut stop loss and the use of input-output tables of quantitative calculation of the loss. Wang Hai zi said from an economic point of view, the earthquake is discussed in detail the calculation of indirect economic loss model, focusing on each side from the input-output related loss assessment to establish the theoretical model. Hu Yu Yin, Yang got a large system of earthquake damage prediction theory. The theory said that the social system is a very complex large system, which is composed of many subsystems (such as command, energy, transportation, electricity, communications etc.). Forecast for economic losses taking into account the social system itself, the earthquake caused direct economic losses and the various subsystems of the indirect economic losses. Suppose the social system is composed of several subsystems, the model considered: the social system by the sub-total economic loss of the direct economic loss; each subsystem's own indirect economic loss; bring the subsystem to other subsystems external indirect economic economic losses such as loss of component parts. The third part of the method in calculating the economic loss to the basis of input and output method, application of modern economic theory, such as sectoral production functions or research methods. Zhong Jiang rong said other cause direct damage from the construction start stop cutting losses, the use of industrial sector analysis of the relationship between the input and output, thus the calculated interaction between the industrial sectors generate the second indirect economic losses.

3 Based on industry related earthquake loss estimation methods

From the impact of earthquakes on the economic system arising from related loss, you can input and output systems of the exogenous variables, such as the final output Y or N and change the initial investment in the impact to simulation. Suppose industries in the quake cut the damage caused by the stop loss TILi, the final product of the industrial sector, Yii under the direct losses in the kinds of value: $Yii=\beta i^* TILi$

Where: βi for the coefficients, $\beta i = Yi0/Xi0$ for the first i-sectors are not affected by the earthquake, the total product; Yi0 for the first i-sectors are not affected by the earthquake when the final product; Yii sectors for the first i- affected by the earthquake when the final product; TILi for the first stop i-cut loss of industry sectors. Other industries affected by the loss of i-sectors of value:

$Yij=(Yi0/Yj0)^* Yjj$

Where: Yij for the i-j-industries sectors affected by the loss of value; Yi0 for the first i-sectors are not affected by the earthquake when the final product; Yj0 one industry sector for the first j affected by the earthquake is not the final product when ; Yij for the first j-j-industries sectors are the first earthquake when the final product. Practice, the earthquake would have more direct economic losses caused by industrial sectors, which led to the stopping number of industries cut losses, and all associated industries will be affected departments. This earthquake i-sectors of the impact of the final product is :

$\overline{Y}_i$=max(Yi1,Yi2,…,Yij,…,Yin)

Industrial sector i total indirect economic loss: ILi= $\sum_{j=1}^{n} \overline{a}_{ij} * \overline{Y}_j$

Industry sectors associated with indirect economic loss: GILi = ILi-TILi

The total economic loss associated with the various departments for industry, only indirect economic losses associated with the sum to which all industry sectors associated with the total loss :GIL= $\sum_{i=1}^{n} GIL_i$

4 Conclusions

Earthquake related loss in indirect economic losses in the earthquake can not be neglected as a part of the paper losses associated with the seismic industry and summarizes current research on the composition of the standard input-output tables and internal relations are introduced. Given the input-output table based on earthquake related loss assessment methods. In this paper, about the assessment of earthquake related loss, there are two assumptions: First, assume that after the earthquake or before the earthquake ,the input-output coefficient table is unchanged. Before or after the earthquake level of production technology is constant or does not change ; Second ,assume that the production department of any inputs consumed in number and level of the sector's total output proportional relationship. In addition, the related loss assessment given by the papers is based on the input-output tables of the line. In fact, it can also be the direction of input-output tables to create columns. By defining the coefficient of direct supply to "total initial investment , intermediate inputs into the" basis of this relationship to build models and related loss assessment.

References:

[1]Jiangrong zhong. Prediction of urban earthquake economic loss[D].harbin: Institute of Engineering Mechanics, China Earthquake Administration,2009

[2]Kawashima K,Kanoh T. Evaluation of indirect economice effects caused by the 1983 Nihonkai-chubu, Japan earthquake[J]. EarthquakeSpectra,1990,6(4):739-756.

[3]Shinozuka M,Chang S E. Advances in earthquake loss estimation and application to Memphis,Tennessee[J].Earthquake Spectra,1997,13(4):739-758.

[4]Gordon P,Moore J E,Richardson H W,Shinozuka M.An integrated model of bridge performance,highway networks,and the spatial metropolitan economy:towards a general model of how losses due to earthquake impactson life lines affect the economy,NCEER-97-0005[A].Proceedings of the Workshop on Earthquake Engineering Frontiers inTransportation Facilities[C],National Center for Earthquake Engineering Research, Buffalo,N.Y.,Aug.29,1997.

[5]yuxiang huang,zongyue yang,yinghong shao. Measurement of indirect economic losses Disaster[J]. Disaster, 2004,9(3):7-11.

[6]haici wang. Earthquake damage assessment and mitigation of economic evaluation of investment projects[R].shanghai: Tongji University,2007

[7]yifan yuan. Report on Natural disaster damage assessment[R].harbin: Institute of Engineering Mechanics, China Earthquake Administration,2002

[8]qijia zou,tuo su,zhizhou ge. Socio-economic impact of the Tangshan earthquake[M]. Academic Books, 2000.

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.751

Study on Processing Mineralogy of Xuanhua Iron Ore

Xiuli Han[1,2,a], Changcun Li[1,2,b], Lina Liu[1,c]

[1]College of Resources and Environment, Hebei Polytechnic University, Hebei Tangshan, China

[2] Mining Development and Safety Technology Key Lab of Hebei Province, Hebei Tangshan, China

[a]hanxl1965@126.com, [b]licc@heut.edu.cn, [c]llngaoxing@126.com

Keywords: Xuanhua iron ore, Processing Mineralogy, Oolitic Structure, Hematite.

Abstract. The chemical component, mineral composition and dissemination characteristics of Xuanhua iron ore are researched systematically. The result shows that: the iron ore mainly is oolitic structure and colloform, xenomorphic granular texture, the mineral composition is complex, the primary metallic minerals is hematite, and the rocky minerals mainly is quartz, followed by carbonate, epidote, chlorite, and amphibole. The diffraction size of hematite and rocky minerals is fine. It is difficult to liberate between hematite and rocky minerals and easy to be mud. The iron ore is very hard to separate, and it can be used in the process of stage grinding and concentration.

Introduction

With the rapid development of steel industry and mineral resources decreasing in recent years, the disparity between supply and demand of iron ore intensive, so more and more attention have been paid to explorate and utilize the low grade and intractable iron[1-5]. Hebei Iron and Stell Group have abundant of hematite ore, about 1.5 hundred million ton, the grade of mining ore is about 40%, because the iron ore mainly is oolitic structure, the component of the mineral is complex, and the dissemination grain size is fine, so the ore belong to one of the beneficiation of unwieldy iron ore in China. For take full advantage of the iron ore, the processing mineralogy of the iron ore was analyzed systemicly, the analyzation is aimed at providing basic data and theoretical basis for mineral processing.

Chemical analysis of iron ore

The chemical element and iron phase of the representative samples of iron ore crashed, grinded and screened are analyzed. The result is shown in the table below.

Table1. Chemical composition of iron ore %

element	TFe	FeO	SiO_2	MgO	Al_2O_3	CaO	S	P	burning loss
content	47.66	9.81	15.08	3.12	4.85	2.78	0.22	0.24	4.5

From the table 1: The grade of iron is 47.66%, the content of FeO only is 9.81%. The impurity of iron ore mainly is SiO2, followed by Al2O3, MgO and CaO, the content of S and P is 0.22%, 0.24%, CaO and MgO by SiO2 and Al2O3 equal 0.296. So the iron ore mainly is acidic hematite ore.

From the table 2: The main mineral of the iron ore is orehematite-limonite, which the distributive law is 91.17%, and than carbonated iron. the distributive law of magnetite, iron sulfide and iron silicate of the iron ore is low, which are 1.22%, 1.03% and 0.35%. Seen from iron occurrence mode, the iron ore is mainly orehematite-limonite ore.

Table2. Iron phase analysis of iron ore %

iron phase	magnetite	hematite-limonite	carbonated iron	iron sulfide	iron silicate	total iron
content	0.58	43.45	2.97	0.49	0.17	47.66
distributive law	1.22	91.17	6.23	1.03	0.35	100

Structure and texture of iron ore

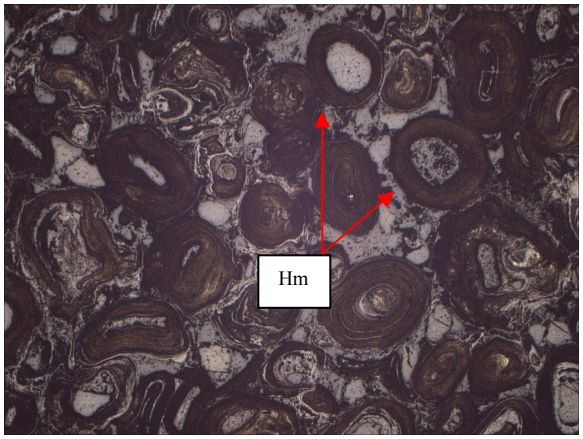
The representative iron ore are made of light sheets, and then tested by micropolariscope for research —Axioskop 40 pol of zeiss.

Structure of iron ore

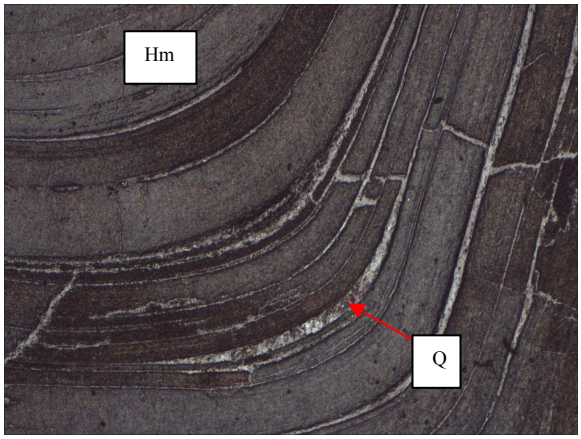
The type of structure of iron ore is simple, mainly is oolitic structure, followed by reniform and impregnation structure. (1) Oolitic structure: Colloidal hematites and rocky minerals are interbedded and form into cambium ring, the center of oolite mainly are quartz, carbonate, next is hematite.（see fig1a~c）. (2) Reniform structure: Clloidal hematite and rocky mineral are interbedded and form into big cambium ring. (3)Impregnation structure: the star dot hematite distribute in quartz and the other rocky mineral[6]（see fig1d）.

Texture of iron ore

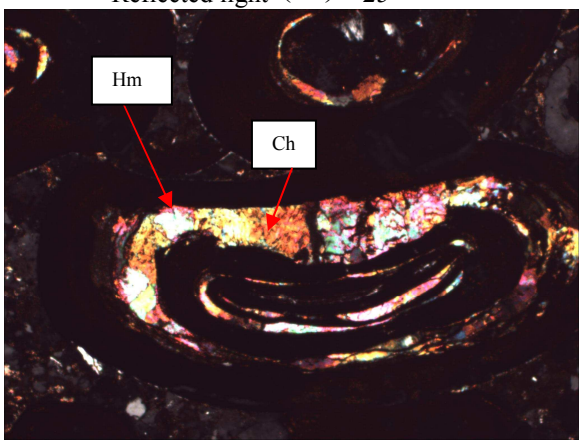
Xuanhua iron ore is colloidal textrue primarily, next is xenomorphic-granular, inclusion and broken texture. (1) Colloidal texture: The shape of the hematite is colloidal aggregate. Because of the grain size of the hematite is fine, and some of quartz, iron pyrites and the other rocky minerals are wapped in them, so it is difficult to liberate between hematite and rocky minerals and easy to be mud.（2）Xenomorphic-granular texture: The hematite and rocky minerals distribute with xenomorphic granular shaped. (3) Inclusion texture: Colloidal aggregate of the hematite and rocky mineral wapped each other and the xenomorphic granular of the hematite are wapped in the rocky minerals. (4)Broken texture: Crack is formed in iron pyrites because of exogenous process, and some rocky minerals filled in it.



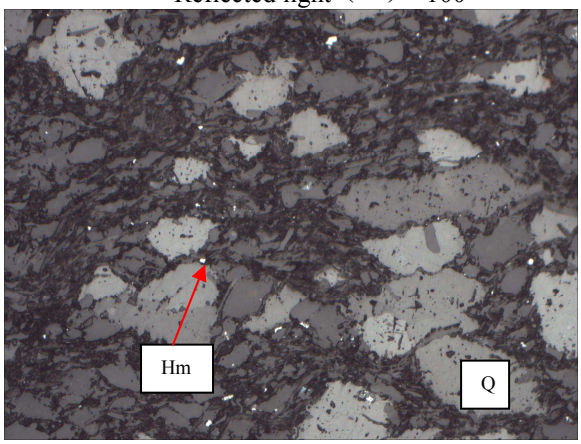
a Oolitic structure
Reflected light（—）×25



b Colloidal hematite and rocky mineral interbedded
Reflected light（—）×100



c Oolitic structure
Transmitted light（+）× 50



d Impregnation structure
Reflected light（—）×200

Hm—Hematite; Q—Quartz; Ch—Chlorite

Fig.1 The pictures of structure and texture of the iron ore

Mineral composition and embedded cloth of ore

Mineral composition of iron ore

Mineral composition and volume percentage of iron ore can be seen in table 3.

Table 3. Mineral composition and volume percentage of iron ore %

metallic minerals			rocky minerals					
hematite	limonite	Iron-pyrites	quartz	carbonate	epidote	chlorite	amphibole	others
60-65	2-3	1-2	15-20	5-7	3-5	2-3	1-2	1-2

According to Table 3, Mineral composition of iron ore is complex, metallic minerals is mainly hematite, the percentage is 60-65%, and then limonite and iron-pyrites; the main rocky minerals is quartz, followed by carbonate, epidote, amphibole and chlorite.

Embedded cloth of the main minerals

Metallic minerals

⑴ Hematite: Hematite is the main iron mineral of the iron ore, the percentage of the hematite is 60-65%. Most of the hematite are colloidal aggregate shaped, covering each other with quartz, carbonate, some other rocky minerals, and form the oolitic hematite, a little of hematite are fine xenomorphic granular and acerose shaped, which wapped in the quartz and other rocky minerals.The grain size of hematite is generally 0.005mm-0.3mm. The oolitic hematite mainly appears as circular and elliptical shaped, most the centers of oolites are small, and are consisted of quartz, carbonate, and other rocky minerals and hematite. Some tiny iron-pyrites and quartz can be seen in the colloidal aggregate hematite. The aperture gaps between oolites are filled up with quartz, carbonate, epidote and other rocky minerals, the contanct relation of the minerals in the gaps and the oolitic hematite is mainly syntopy. (see fig.2 a~f).

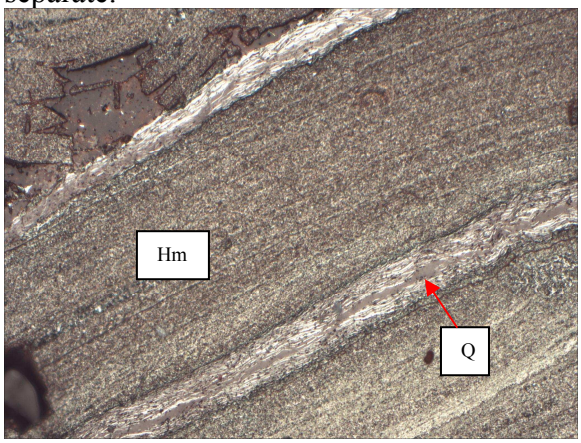
⑵ Iron-pyrites: Iron-pyrites is the main sulfur-bearing mineral and harmful mineral. The grain size of iron-pyrites varies, generally is 0.002mm-0.15mm. the iron-pyrites are wapped in colloidal aggregate heamtite, quartz and the other minerals with a shape of drops and stars, some of the rest appears as euhedral crystal, subhedron, and interbeded or inlay with the hematite. (see fig.2 e~f).

Rocky minerals

Quartz：Quartz is the main rocky mineral of the iron ore, the percentage is 15-20%. Most of the quartz is granular shaped, grain size varies from 0.005mm to 1.78 mm, the coarse grains fill in the center of oolitic hematite or fill between the oolites, and the tenuous grains are interbeded or mingled with colloform hematite. (see fig.2 a~d).

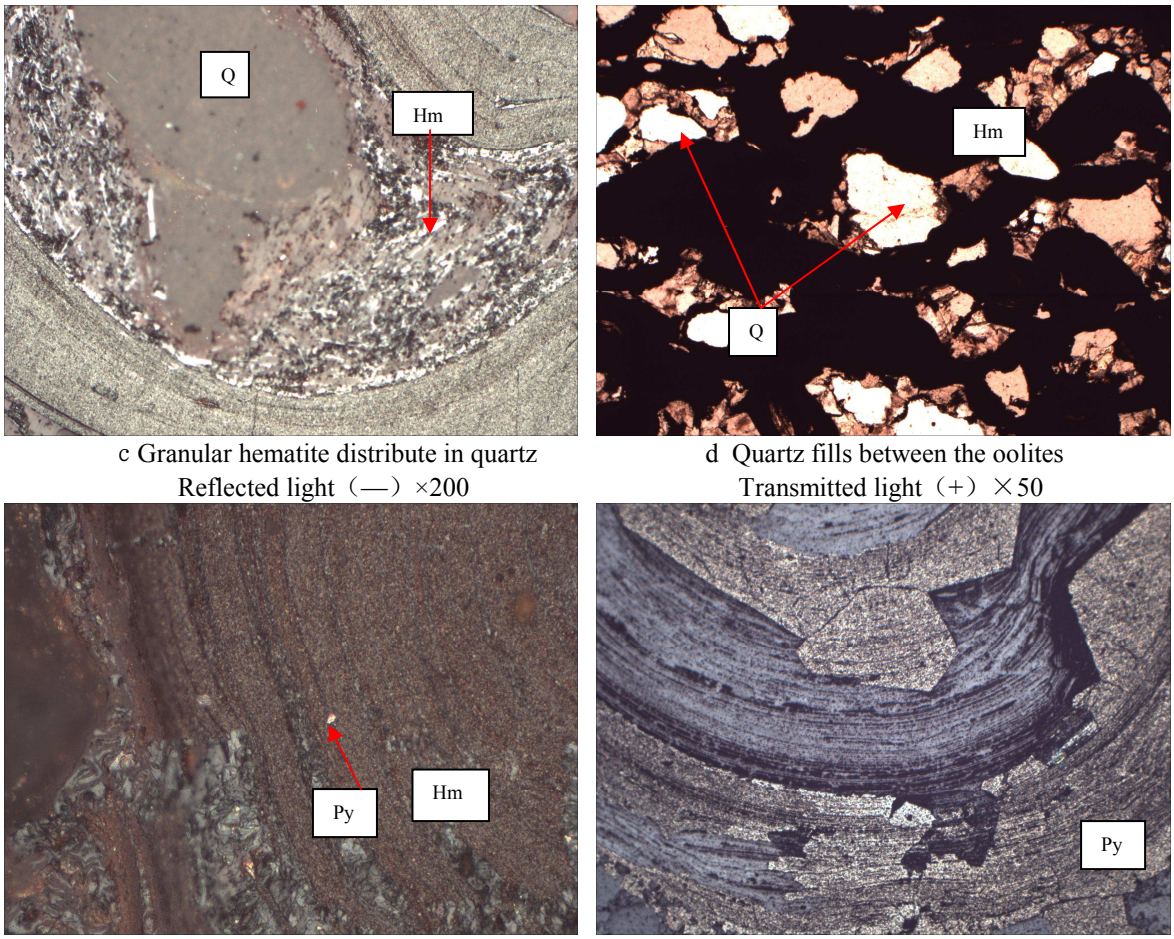
Carbonate most appears in the shape of granule, grain size is 0.04mm-0.17mm. Most of the carbonate inlay with the oolitic hematite, and some fill in the the center of oolites. Some carbonate braid can be seen in the print of ore.

The analysis result of the feature of Xuanhua iron ore shows that: the main iron mineral is hematite, the main rocky mineral is quartz, so the aim of concentration is made hematite and quartz to separate.



a hematite and quartz interbeded
Reflected light（—）×200

b Quartz fills in the center of oolite
Reflected light（—）×200



c Granular hematite distribute in quartz
Reflected light（—）×200

d Quartz fills between the oolites
Transmitted light（+）×50

e Particle iron-pyrites in colloidal hematite
Reflected light（—）×500

f Iron-pyrites and hematite interbeded
Reflected light（—）×50

Hm—Hematite; Q—Quartz; Py—Iron-pyrites

Fig.2 The Pictures of Inbuilt features of main minerals

Grain size analysis of hematite and gangue mineral in the mineral

In order to characterize the distribution of the hematite, the craft grain size of the hematite and rocky mineral was curved by using line measurement under the micropolariscope.

Table 4 The testing result of the dissemination size of hematite and rocky mineral %

size fraction [mm]	hematite distibution rate			rocky mineral distribution rate		
	content	positive accumulation	negative accumulation	content	positive accumulation	negative accumulation
+0.666	0.4	0.4	100	15.2	15.2	100
-0.666+0.333	8.4	8.8	99.6	23.4	38.6	84.8
-0.333+0.166	18.1	26.9	91.2	12.9	51.5	61.4
-0.166+0.083	21.2	48.1	73.1	10.5	62	48.5
-0.083+0.042	17.3	65.4	51.9	12.1	74.1	38
-0.042+0.021	17.3	82.7	34.6	13.3	87.4	25.9
-0.021+0.01	11.7	94.4	17.3	7.4	94.8	12.6
-0.01+0.005	5.1	99.5	5.6	4.5	99.3	5.2
-0.005	0.5	100	0.5	0.7	100	0.7

From the table 4: The size distribution of hematite and rocky mineral in size fraction of +0.083mm is up to 48.1%, 62%, in size fraction of -0.042mm is 519%, 38%, in size fraction of -0.01mm is 17.3%, 12.6%. This shows than the dissemination size of the hematite and rocky mineral is very fine, it is difficult to liberate between hematite and rocky minerals and easy to be mud, so that

the iron ore is very hard to separate, and it can be used in the process of stage grinding and concentration[7].

Conclution

(1) In Xuanhua iron ore, the content of the hematite-ironstone is 43.45%, the distributive law is 91.17%. The ore belongs to acidic hematite ore.

(2) The ore is oolitic structure, colloidal and xenomorphic granuler texture. The mineral composition is complex, the metallic minerals mainly is hematite, and the rocky minerals mainly is quartz, the next is carbonate, epidote, chlorite, amphibole.

(3) In the iron ore, the shape of the hematite mainly is colloidal aggregate; among the colloidal hematite include the iron-pyrites and quartz. Quartz and the other rocky mineral are located in the center of oolitic hematite or fill between the oolites.

(4) In the iron ore, the distribution size of the hematite and the rocky mineral is fine, it is difficult to liberate between hematite and rocky minerals and easy to be mud.

(5) The iron ore is very hard to separate, and it can be used in the process of grind stage and concentration.

References

[1] Xianjun Lv, Jun Qiu: Investigation on the Process Mineralogy of a Low Grade Iron Ore from Xinjian. Metal Mines, Forum Vol.6(2009), p.99-100.

[2] Jinrui Zhang, Like Hu: The research on refractory oolitic hematite and expectation. Mining and Technology of China, Forum Vol. 7-16(2007), p.74-76.

[3] Zhenfu Lv, Yuexin Han and Yanjun Li: Mineralogical Process Investigation on Liaoning Yezhugou Iron Ore. Metal Mines, Forum Vol. 2(2010), p. 80-83.

[4] Xiong Tong, Yingshu Li: Experiment of Refractory Oolitic Hematite with Ore Processing.Engineering Science of China, Forum Vol. 7(2005), p. 324-326.

[5] Weixin Li, Baoping Liu: Presenn Status and Development Orientation of China's Red Iron Ore Beneficiation Technology Research. Metal Mines, Forum Vol. 3(2005), p. 1-6.

[6] Leguang Zhou: Ore Foundation (Mitallurgical Industry Press, Beijing 1990)

[7] Leguang Zhou: Process Mineralogy (Mitallurgical Industry Press, Beijing 1990)

doi:10.4028/www.scientific.net/AMM.50-51.756

Application of combination evaluation for academic title evaluation

Baofeng Li[1, a], Jingguo Qu[2,b] and Puyu Hao[1,c]
[1] Department of Mathematics and Information Science, Tangshan Normal University, Tangshan, China
[2] College of Light industry, Hebei Polytechnic University, 063000 ,China
[a]lbf6713@sina.com, [b]qujingguo@163.com, [c]puyu_hao@163.com

Keywords: entropy weight method, Topsis method, double base points method, combination evaluation, academic title evaluation.

Abstract. In this paper, using the relevant data of 34 teaching staffs who participate in the academic title evaluation of associate professor in 2010, firstly it introduces the entropy weight method, Topsis method with subjective weight, Topsis method with objective weight and double base points method with subjective weight to evaluate and sort the performance of 34 teaching staffs. Secondly, two combination evaluation models are constructed to do the same work and the conclusions are more science and rational.

Introduction

Evaluation work of the professional and technical positions involves the performance evaluation about the teaching staffs. It not only helps improve the scientific and technological management level and efficiency, but also to mobilize the enthusiasm of teaching staffs. So the establishment of a scientific, objective, fair and quantifiable evaluation method of performance is the exploration direction for many research scientists. To the academic title evaluation in the university, the job is to evaluate the teaching staffs from the aspects of moral character, diligence and performance and select the excellent one. Therefore the process of the academic title evaluation belongs to multiple attribute decision making problems. At present, there are many methods to solve the multi-attribute decision making problem, such as simple additive weighting (SAW), approaching the ideal solution for sorting methods (TOPSIS)[1], Analytic Hierarchy Process (AHP)[2], principal component analysis (PCA)[3], etc. To a decision making problems, you can use a variety of decision-making methods (model) to solve it. But due to the difference among the methods in the theoretical basis and information, the results (scheme order) are generally different and sometimes even have a great difference. At the same time, different methods have their own characteristics, scope, advantages and disadvantages. Therefore, in order to get a more reliable result, we can use several methods to evaluate and compare, synthesize the results. The new result is more reasonable and convincing. In this paper, it uses 34 teaching staffs of Tangshan Normal University as the evaluation object to validate the mathematical model.

The basic steps of multi-attribute combination evaluation

To the multi-attribute combination evaluation, the decision-making approach in it is the synthetically sorting method. And the evaluation result is the ranking results of decision scheme. The basic steps of combination evaluation are as follow:

1. Choosing some representative evaluation methods. Generally we choose the evaluation methods in the two basic kinds evaluation methods (one kind is the method with subjective weight in which It determines weights by experts based on experiences; Another kind is the method with objective weight in which the weights are determine by certain rules and do not depend on subjective judgments.) and select at least one method in each one. If these two methods are combined, the defects can be complementary and the evaluations are more comprehensive.

2. Comparing the different evaluation results. When we compare the different sort obtained by different evaluation methods, since they address the same issue, the results should not be a greater difference. Otherwise, we should consider replacing part of the sorting method. For any two sorted results obtained by the two methods, their level of consistency can be described by Spearman rank correlation coefficient which is calculated as follows:

$$r = 1 - \frac{6}{m(m^2 - 1)} \sum_{i=1}^{m} d_i^{\,2} .$$

where m is the number of decision scheme, d_i is the difference between the two ranking results of decision scheme A_i. Observe that, the lager r is, the more consistent the two ranking results are. If the two ranking results obtained by two methods are exactly the same, then we have $r = 1$.

3. Composing the different evaluation results and obtaining the combination evaluation value. The common combination evaluation methods are as follow:

(1) The mean value method. Suppose that there are s kinds of evaluation methods, p_{ik} is the sorting number of decision scheme A_i in the kth evaluation methods ($i = 1,..., m$, $k = 1,..., s$). So we let

$$a_i = \frac{1}{s} \sum_{k=1}^{s} p_{ik}, i = 1,2 \cdots, m .$$

Based on the value of a_i, the decision schemes are sorted. Observe that, the smaller a_i is, the better A_i is. If $a_i = a_j$, then we compare their variances. Their variances can be computed by

$$\sigma_i^{\,2} = \frac{1}{s-1} \sum_{k=1}^{s} (p_{ik} - a_i)^2 \text{ and } \sigma_j^{\,2} = \frac{1}{s-1} \sum_{k=1}^{s} (p_{jk} - a_j)^2 .$$

The variance is, the decision scheme is.

(2) Borda method. In the results obtained by different methods, if the number of that decision scheme A_i is better than A_j is greater than the number of that A_j is better than A_i, then .it note that $A_i S A_j$; If the number is equal, then note that $A_i E A_j$. So we can structure the Borda matrix as follow:

$$B = (b_{ij})_{m \times m}, \text{where } b_{ij} = \begin{cases} 1, if\ A_i E A_j \\ 0, other\ cases \end{cases} .$$

Define the score of A_i is

$$b_i = \sum_{j=1, j \neq i}^{m} b_{ij} .$$

Then the b_i is the "excellent" (or "wins") number of A_i. Observe that, the greater b_i is, the better A_i is. If $b_i = b_j$, then we compare their variance. The variance is, the decision scheme is.

Case

Taking the information about 34 teaching staffs who participate in the academic title evaluation of associate professor in Tangshan Normal University in the 2010 as example, it uses combined evaluation to evaluate and sort their capability. The original data can be seen in Table 1(The name of the teaching staffs are substitute by their second name).

And the value of each index about the teaching staff is calculated by the same and fixed formula which is accepted by the teaching staffs in Tangshan Normal University.

Table 1 The date of the 34 teaching staffs

Name	Educational background	Teaching quantity	Scientific research workload	Job performance
Cheng	5	47	30.31	9
Mou	4	51	21.585	7.75
Yingzi	5	32.4	10.805	7
Yufen	4	58	9.67	7
Dong	3	46.12	8.735	7
Junxia	5	48.5	22.19	7
Liqiang	5	44.1	11.09	7
Jining	5	60.3	15.92	7
Shugeng	5	44.62	6.5	8.75
Zhixin	6	57	21.48	7
Xiaohui	5	60	8.19	5
Wei	5	49	9.59	7.3
Baofeng	5	63	24.935	8
Gang	3	46	16.53	8.25
Haicheng	3	59	11	7.75
Jian	5	44.5	14.65	6.25
Yuanyuan	3	67	14.5	9
Fafen	5	59.57	9.1975	7
Junhong	4	58	19.645	9
Yanzhi	5	61	29.665	8
Xiangguo	3	47	8	9
Yongqing	4	51	22.5	5
Lina	5	44.11	21.19	7.5
Shengcai	5	55	16.7	7.5
Zhenling	5	60.12	36.7	7
Yanyu	3	54.5	12.85	9
Jing	6	55	30.68	5.25
Liying	4	58.8	5.73	7.35
Yuanjing	4	52.5	24.4	7.5
Yongping	6	41.8	20.145	5
Xin	5	65.73	26.17	8.65
Lili	6	58	15.145	7.75
Yan	5	70.81	15.525	7
Yuhua	5	54.5	15.735	9

1. Using six evaluation methods to evaluate. The paper uses Topsis method with subjective weights (note 1), Topsis method with objective weights(note2), double base points method with subjective weights(note 3)[4], double base points method with objective weights(note 4), the entropy weight method[5] (note 5) and principal component analysis method(note 6) to evaluate and sort the date. The results can be seen in the Table 2. And the subjective weights used is

$$w_1 = (0.1, 0.45, 0.35, 0.1)$$

which is given by experts and objective weights is

$w_2 = (0.3342, 0.12, 0.328, 0.218)$

which is calculated by entropy method.

Table 2 The results of six methods

Name	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6
Cheng	5	2	11	2	3	6
Mou	11	12	16	18	13	12
Yingzi	34	24	34	24	34	34
Yufen	25	31	22	28	23	23
Dong	32	34	32	34	33	29
Junxia	10	9	21	12	12	20
Liqiang	30	23	31	22	30	30
Jining	16	18	10	15	18	16
Shugeng	33	28	33	19	32	26
Zhixin	8	7	8	6	11	13
Xiaohui	26	29	20	26	26	32
Wei	29	26	28	21	29	28
Baofeng	6	6	5	7	5	5
Gang	21	21	26	29	24	18
Haicheng	22	30	18	32	16	15
Jian	27	20	29	20	27	31
Yuanyuan	15	22	7	27	8	4
Fafen	24	27	17	23	21	24
Junhong	12	14	9	16	6	7
Yanzhi	2	4	3	4	4	2
Xiangguo	31	32	30	33	31	21
Yongqing	9	13	19	25	14	27
Lina	14	10	25	11	20	22
Shengcai	18	16	14	13	15	17
Zhenling	1	1	1	1	1	3
Yanyu	23	25	23	30	22	10
Jing	3	3	6	3	7	19
Liying	28	33	24	31	28	25
Yuanjing	7	8	12	17	10	11
Yongping	17	11	27	9	25	33
Xin	4	5	2	5	2	1
Lili	19	15	13	8	17	14
Yan	13	19	4	14	9	8
Yuhua	20	17	15	10	19	9

2. Comparing the different evaluation results. The correlation coefficients are as follow:

Table 3 The correlation coefficients

r_{12}	r_{13}	r_{14}	r_{15}	r_{16}	r_{23}	r_{24}	r_{25}	r_{26}	r_{34}	r_{35}	r_{36}	r_{45}	r_{46}	r_{56}
0.91	0.85	0.75	0.94	0.72	0.67	0.9	0.80	0.56	0.64	0.93	0.84	0.68	0.49	0.85

Through observation, the correlation coefficients between the double base points method with objective weight and other methods and principal component analysis method and other methods are poor.

So we remove the two evaluation methods and only use the other four methods to evaluate the date.

3. Computing the combination evaluation value (see Table 4). And by computed the correlation coefficient between the two combination evaluation methods is 0.994194.

Table 4 The results of the combination evaluation

Name	The result of the mean value method	The result of Borda method	Name	The result of the mean value method	The result of Borda method
Cheng	5	5	Fafen	22	23
Mou	11	13	Junhong	9	9
Yingzi	33	34	Yanzhi	2	2
Yufen	26	25	Xiangguo	31	31
Dong	34	32	Yongqing	14	14
Junxia	12	11	Lina	18	15
Liqiang	30	30	Shengcai	16	16
Jining	15	18	Zhenling	1	1
Shugeng	32	33	Yanyu	24	24
Zhixin	7	7	Jing	4	4
Xiaohui	25	26	Liying	29	28
Wei	28	29	Yuanjing	8	8
Baofeng	6	6	Yongping	20	19
Gang	23	22	Xin	3	3
Haicheng	21	21	Lili	17	17
Jian	27	27	Yan	10	10
Yuanyuan	13	12	Yuhua	19	20

Conclusion

In this paper, to the quantitative data of 34 teaching staff in Tangshan normal university, we propose two combined evaluation model. Firstly, four evaluation methods are used to sort. Then it uses two methods to combine their results and get the comprehensive ranking results. At he same time, the correlation coefficient between the two combined evaluation methods are computed which show that the credibility of the results show is high. In other word, the two combined evaluation model not only consider the differences between the single methods, but also consider the order factors. Therefore, the evaluation results are more comprehensive, scientific, rational, and more credible. In a word, this study provides new practical ideas and methods for the job of academic title evaluation.

References

[1] Yande Gong. Comprehensive Income Coordination Strategy for Three-echelon CLSC Based on TOPSIS Method. Journal of Systems & Management. Vol.3 (2010), p. 260

[2] Mingyu chen. Application of AHP in the undergraduate education evaluation. Journal of Xi'an Polytechnic University. Vol.13 (2010), p. 106

[3] Qingju lou.A Study on Constructing Investor Sentiment Index by Using Main Component Method. Journal of Yunnan University of Finance and Economics. Vol. 2(2010), p. 114

[4] Caimei Lu. Study on Dynamic Evaluation of Ecological Environment Quality of Shanxi Province. Soil and Water Conservation In China. Vol. 4(2004), p.16

[5] Guixiang Shen. Comprehensive evaluation on reliability of numerically-controlled machine tool based on entropy weight method. Journal of Jilin University. Vol. 5(2009), p. 1208

doi:10.4028/www.scientific.net/AMM.50-51.761

Existence of Positive Solutions for a Class of Higher-order Neutral Delay Difference Equations

Donghua Wang [1, a], Yuhuan Cui [2,b] and Puyu Hao [1,c]

[1] Department of Mathematics and Information Science, Tangshan Normal University, Tangshan, China

[2] College of Light industry, Hebei Polytechnic University, 063000 ,China

[a]me_diandian@yahoo.com.cn, [b]yuhuan,cui@yahoo.com.cn, [c]puyu_hao@163.com

Keywords: Higher-order neutral difference equations; Positive Solutions; Fixed-Point Theorem of Krasnoselskii.

Abstract. In this paper, a class of higher-order neutral delay difference equations is investigated. Some sufficient condition of the asymptotic behavior and existence of positive solutions for the equations are obtained. At last, we give their applications to some more general equations.

Introduction

Consider the following higher-order neutral delay difference equations.

$$\Delta^d (x(n) - p(n)x(n-k)) + q(n)\prod_{i=1}^{m} \left|x(n-\sigma_i)\right|^{\alpha_i} signx(n-\sigma_1) = 0 \quad n \geq n_0 , \tag{1}$$

where d, k, σ_i are nonnegative integers, $i = 1,2,...,m$; $\{p(n)\}$, $\{q(n)\}$ are sequences of real numbers; $\alpha_i > 0$ and $\sum_{i=1}^{m} \alpha_i = 1$; Let $\sigma = \sum_{i=1}^{m} \sigma_i$, $M = \max\{\sigma_i, i = 1,2,...,m\}$ and $k \geq M$; Δ is the forward difference operator defined by $\Delta x(n) = x(n+1) - x(n)$, $\Delta^d x(n) = \Delta(\Delta^{d-1} x(n))$. $\{x(n)\}$ is said be a solution of (1) if $x(n)$ has definition for $n \geq n_0 - k$ and satisfies the equation for $n \geq n_0$. The solution of (1) is said to be oscillatory if it is neither eventually positive nor eventually negative. Otherwise, the solution is said to be nonoscillatory.

In recent years, in the field of biomathematics, modern physics, socio-economic and other social sciences, the research about neutral delay differential equations becomes more and more important. So the existence of positive solution and oscillatory behavior of these equations get rapid development .There are many results about them, such as [1,2,3]. However, these studies are mostly about low order or linear equations. The research on existence of positive solution about higher order difference equations are little and most part are distinguish even and odd [4]. Therefore, using fixed-point theorem of Krasnoselskii [5], to a class of high order equations, this paper get some sufficient condition about its solution.

Main Results

LEMMA 1 Assume that $\{p(n)\}$ is a nonzero sequence of same sign, $|p(n)| > \delta$ (where δ is a positive constant) for $\forall n \geq n_0$ and $(-1)^d q(n) > 0$. If

$$\frac{1}{(d-1)!}\sum_{i=n_0}^{\infty} (i - n_0 + d - 1)^{(d-1)} q(i) e^{n_0+\sigma-i} \tag{2}$$

converges, then the following statements hold:

(I) If $\{p(n)\}$ is a sequence of positive real numbers, then (1) has an eventually positive solution $x=\{x(n)\}$ and it satisfies

$$\lim_{n\to\infty} x_n = 0\,; \tag{3}$$

(II) If $\{p(n)\}$ is a sequence of negative real numbers, then (1) has a solution which satisfies(3).

Proof Suppose that H is a set of bounded sequence. For any $y=\{y(n)\}_{n=n_0}^{\infty} \in H$, we define

$$\|y\| = \sup_{n\ge n_0} |\, y(n)\,|\,.$$

Observe that H is a Banach space.

(I) If $\{p(n)\}$ is a sequence of positive real numbers, then $p(n)>\delta,\ for\ n>n_0$. Let

$$P_1 = \sup_{n\ge n_0}\{\frac{1}{p(n)}\},$$

then there exists $r>0$ and integer $n_1 \ge n_0$ such that

$$0 < P_1 e^{-rk} < 1,\ n \ge n_1, \tag{4}$$

$$P_1 e^{-rk} + \frac{(-1)^d}{(d-1)!} P_1 \sum_{i=n+k}^{\infty} (i-n-k+d-1)^{(d-1)} q(i) e^{r(n+\sigma-i)} \le 1. \tag{5}$$

Let

$$Q(n) = P_1 e^{-rk} + \frac{(-1)^d}{(d-1)!} P_1 \sum_{i=n+k}^{\infty} (i-n-k+d-1)^{(d-1)} q(i) e^{r(n+\sigma-i)}, \tag{6}$$

$$G(n) = \frac{(-1)^d}{(d-1)!} P_1 \sum_{i=n+k}^{\infty} (i-n-k+d-1)^{(d-1)} q(i) e^{r(n+\sigma-i)}. \tag{7}$$

Form the known conditions, we obtain that $\{G(n)\}$ is monotone decreasing and

$$\lim_{n\to\infty} G(n) = 0\,. \tag{8}$$

Therefore, for any $\varepsilon>0$, there exists integer $n_2 \ge n_1$ such that

$$G(n) < \varepsilon\ ,\ n \ge n_2\,. \tag{9}$$

Let

$$A = \{x \in H : 0 \le x(n) \le G(n), n \ge n_2\}\,. \tag{10}$$

Observe that A is a bounded closed convex set in H. For any $y=\{y(n)\}, z=\{z(n)\} \in A, n \ge n_2$, we define the mapping as following:

$$T_1 y(n) = \frac{e^{-rk}}{p(n+k)} y(n+k)\,; \tag{11}$$

$$T_2 z(n) = \frac{(-1)^d}{p(n+k)(d-1)!} \sum_{j=n+k}^{\infty} (j-n-k+d-1)^{(d-1)} q(j) e^{r(n+\sigma-j)} \prod_{i=1}^{m} z(j-\sigma_i)^{\alpha_i}\,. \tag{12}$$

For $|T_1 y(n) - T_1 z(n)| \le P_1 e^{-rk} |y(n+k) - z(n+k)| \le P_1 e^{-rk} \|y - z\|$ and form (4), we have T_1 is a contraction mapping on A.

On the other hand, form (5), (6) and (9), we obtain

$$|T_2 y(n) - T_2 z(n)| \le Q(n) \mid \prod_{i=1}^{m} y(j-\sigma_i)^{\alpha_i} - \prod_{i=1}^{m} z(j-\sigma_i)^{\alpha_i} \mid < \varepsilon\,.$$

Thus, T_2 is a continuous mapping. And form (4), (10) and the monotonicity of $\{G(n)\}$, we get

$$0 \le T_1 y(n) \le e^{-rk} P_1 y(n+k) \le e^{-rk} P_1 G(n+k) \le e^{-rk} P_1 G(n)\,, \tag{13}$$

$$0 \le T_2 z(n) \le \frac{(-1)^d}{(d-1)!} P_1 \sum_{j=n+k}^{\infty} (j-n-k+d-1)^{(d-1)} q(j) e^{r(n+\sigma-j)} G(n)\,. \tag{14}$$

Then by (6), (13) and (14), we have $0 \le T_1 y + T_2 z \le Q(n)G(n) \le G(n)$. Hence, $T_1 y + T_2 z \in A$.

Let $B = \{y \in A : y(n) = 0, n \ge n_2\}$, we know B is a $\varepsilon -$net of A. Therefore A is a sequentially compact set. At the same time, form (14), we have $T_2 A \in A$. Then by Fixed-Point Theorem of Krasnoselskii, we obtain that there exists $\mu = \{\mu(n)\} \in A$ which satisfies $\mu = T_1 \mu + T_2 \mu$. Let $x(n) = \mu(n) e^{-rn}$, by the upper formula, we have

$$x(n) = \frac{1}{p(n+k)} x(n+k) + \frac{(-1)^d}{p(n+k)(d-1)!} \sum_{j=n+k}^{\infty} (j-n-k+d-1)^{(d-1)} q(j) \prod_{i=1}^{m} x(j-\sigma_i)^{\alpha_i}\,.$$

Observe that $x(n) > 0$ holds. So for all $n \ge n_2 + k$,

$$x(n) - p(n)x(n-k) = \frac{(-1)^{d+1}}{(d-1)!} \sum_{j=n+k}^{\infty} (j-n-k+d-1)^{(d-1)} q(j) \prod_{i=1}^{m} x(j-\sigma_i)^{\alpha_i}$$

holds. Thus, $\Delta^d (x(n) - p(n)x(n-k)) = -q(n) \prod_{i=1}^{m} |x(n-\sigma_i)|^{\alpha_i}$.

Therefore, x is an eventually positive solution of (1). And from (8), we have $\lim_{n\to\infty} x_n = 0$.

(II) Let $P_2 = \sup_{n \ge n_0} \{|\frac{1}{p(n)}|\}$, then there exists $r' > 0$ and integer $n_3 \ge n_0$ such that $0 < P_2 e^{-r'k} < 1$ and $P_2 e^{-r'k} + \frac{(-1)^d}{(d-1)!} P_2 \sum_{i=n+k}^{\infty} (i-n-k+d-1)^{(d-1)} q(i) e^{r'(n+\sigma-i)} \le 1, n \ge n_3$.

The following proof of II) is similar to that of I), and will be omitted here. The proof is complete.

LEMMA 2 Suppose that $\{p(n)\}$ is s a nonzero sequence of same sign, $0 <| p(n) |< 1 - \delta$ (where $0 < \delta < 1$ is a constant) for any $n \ge n_0$ and $(-1)^{d+1} q(n) > 0$ .If (2) holds, then (I) If $\{p(n)\}$ is a

sequence of positive real numbers, then (1) has an eventually positive solution $x=\{x(n)\}$ and it satisfies (3); (II) If $\{p(n)\}$ is a sequence of negative real numbers, then (1) has a solution which satisfies(3).

The proof of lemma 2 is similar to that of lemma1and will be omitted here.

Extension

The more general form of (1) will be study as follow:

$$\Delta^d(x(n)-p(n)x(n-k))+q(n)f(x(n-\sigma_1),...,x(n-\sigma_m))=0 \quad n\geq n_0, \tag{15}$$

where d,k,σ_i are nonnegative integers, $i=1,2,...,m$; $\{p(n)\}$ and $\{q(n)\}$ are sequences of real numbers; $\sigma=\max\{\sigma_i,i=1,2,...,m\}$ and $k\geq\sigma$; f satisfies the following statements:

1) $f\in C(R^m,R)$, $x_i\cdot f(x_1,...,x_m)\geq 0$, $x_i\in R$,and for any x_i , f is monotone non-decreasing, $i=1,2,...,m$;

2) $\lim\limits_{(x_1,..,x_m)\to(0,..,0)}\dfrac{|f(x_1,...,x_m)|}{\prod\limits_{i=1}^{m}|x_i|^{\alpha_i}}=M$, $M\geq 0$, $\alpha_i>0$ and $\sum\limits_{i=1}^{m}\alpha_i=1$;

3) $|f(x_1,...,x_m)-f(y_1,...,y_m)|\leq C\sup\limits_{1\leq i\leq m}|x_i-y_i|, C\geq 0, x_i,y_i\in R$.

LEMMA 3 Assume that $\{p(n)\}$ is a nonzero sequence of same sign, $|p(n)|>\delta$ (where $\delta>1$ is a constant) for $\forall n\geq n_0$ and $(-1)^d q(n)>0$. If $\sum\limits_{i=n_0}^{\infty}i^{d-1}|q(i)|<\infty$, then the following statements hold: (I) If $\{p(n)\}$ is a sequence of positive real numbers, then (15) has an eventually positive solution $x=\{x(n)\}$ and satisfies (3);(II) If $\{p(n)\}$ is a sequence of negative real numbers, then (15) has a solution which satisfies(3).

Proof Also suppose that H is a set of bounded sequence. For any $y=\{y(n)\}_{n=n_0}^{\infty}\in H$, we define $\|y\|=\sup_{n\geq n_0}|y(n)|$.

The proof of lemma 3 is similar to that of lemma1and will be omitted here.

LEMMA 4 Suppose that $\{p(n)\}$ is a nonzero sequence of same sign, $0<|p(n)|<1-\delta$ (where $0<\delta<1$ is a constant) for any $n\geq n_0$ and $(-1)^{d+1}q(n)>0$ .If (2) holds, then (I) If $\{p(n)\}$ is a sequence of positive real numbers, then (15) has an eventually positive solution $x=\{x(n)\}$ and it satisfies (3); (II) If $\{p(n)\}$ is a sequence of negative real numbers, then (15) has a solution which satisfies (3).

The proof of lemma 4 is similar to that of lemma1and will be omitted here.

Conclusion

This paper gets several sufficient conditions about a class of higher order neutral delay difference equations and the results are extended to the more general form. But when the condition (2) does not establish, the properties of the difference equations remain to be further studied.

References

[1] Zhang G , Cheng S S. Positive solution of a nonlinear neutral difference equation. *Nonlinear Anal.* Vol.28 (1997), p. 729

[2] Yong Zhou. Existence of Non-oscillatory Solutions of Higher-order Neutral Difference Equations with General coefficients. *Applied Mathematics Letters.* Vol.15 (2002), p. 785

[3] Xianyong Huang. The Existence of Nonoscillatory Solutions to the Second-Order Neutral Delay Difference Equation. *Journal of Guangdong Education Institute.* Vol. 27(2007), p. 36

[4] Wenlei Dong. Existence of Positive Solutions for High-Order Neutral Difference Equations. *Journal of Shijiazhuang Institute of Railway Technology*. Vol. 6(2007), p.1

[5] T.A.Burton. A Fixed-Point Theorem of Krasnoselski. *Applied Mathematics Letters*. Vol. 11(1998), p. 85

doi:10.4028/www.scientific.net/AMM.50-51.766

An evaluation method Based on grey incidence coefficient and double-point method

Puyu Hao[1,a], Baofeng Li[2,b], and Yuhuan Cui[2,c]

[1] Department of Mathematics and Information Science, Tangshan Normal University, Tangshan, China

[2]College of Light industry, Hebei Polytechnic University, 063000 ,China

[a]puyu_hao@163.com, [b]lbf6713@sina.com, [c]yuhuan.cui@yahoo.com.cn

Keywords: economic enterprises, correlation coefficient, double-point method, evaluation.

Abstract. To one company, it takes 8 major economic indicators as evaluation indexes. Firstly, it uses correlation method to determine the weight and uses the double-point method to sort the 8-years comprehensive economic enterprises. Then through the analysis of the correlation matrix, it identifies the main factors. Finally, the ranking and principal component analysis results are compared and analyzed.

Introduction

In the evaluation of economic efficiency of enterprises, how to establish scientific and reasonable evaluation index system, by means of scientific evaluation method, is very important [1,2]. So finding an accurate quantitative analysis, an integrated evaluation of business plans to develop a scientific, rational policy and long-term planning is very necessary. Generally, in a multi-indices of the issue of multi-use integrated assessment method, formula, overall balance law [3], we propose a correlation to determine the weight (objective weight) of the double base points, which can comprehensively evaluate the enterprise economic benefits.

The basic steps of evaluation method

1. Standardization of the indexes.

Suppose that the original target set is $X=(x_{ij})_{m\times n}$, so if x_{ij} is a benefit indexes, we let

$$y_{ij}=\frac{x_{ij}-x_j^{-}}{x_j^{+}-x_j^{-}},\quad i=1,2,\cdots,m,\quad j=1,2,\cdots,n;$$

If x_{ij} is a cost-based indexes, let

$$y_{ij}=\frac{x_j^{+}-x_{ij}}{x_j^{+}-x_j^{-}},\quad i=1,2,\cdots,m,\quad j=1,2,\cdots,n,$$

where $x_j^{-}=\min_i\{x_{ij}\}$, $x_j^{+}=\max_i\{x_{ij}\}$, $i=1,2,\cdots,m$, $j=1,2,\cdots,n$, $y_{ij}\in[0,1]$, $j=1,2,\cdots,n$, let $Y=(y_{ij})_{m\times n}$.

2. The computing of the correlation coefficient and correlation degree.

The ideal program notes y_0 and $y_0(j)=\{y_j^{+}\,|\,j=1,2,\cdots n\}$, where $y_j^{+}=\max_i\{y_{ij}\}$, $j=1,2,\cdots,n$.

Then to the jth index, the correlation coefficient $\xi_i(j)$ between the ideal program and the property value of the ith program $y_i(j)=\{y_{ij}\,|\,i=1,2,\cdots m, j=1,2.\cdots,n\}$ can be defined as

$$\xi_i(j)=\frac{\min\limits_i\min\limits_j|y_0(j)-y_i(j)|+\rho\max\limits_i\max\limits_j|y_0(j)-y_i(j)|}{|y_0(j)-y_i(j)|+\rho\max\limits_i\max\limits_j|y_0(j)-y_i(j)|}$$

where $\rho\in(0,1)$ is resolution coefficient and Its role is to reduce the impact of distortion caused by the large absolute value. Obviously, the larger the value of $\xi_i(j)$ is, the closer two programs are. So the correlation degree can be computed by $\varpi_i=\frac{1}{n}\sum_{k=1}^{n}\xi_i(k)$ .Observe that the lager ϖ_i is, the program is better.

3. The calculating of the closeness degree [4].

Case

We adopt the above method to evaluate the economic enterprises of one company (see [3]), the original date can be seen as follow:

Table 1 The original date

Year	Gross Industrial Output Value (Million yuan)	The cost of 100 yuan output value (yuan)	Total profit (Million yuan)	Total of profit and tax (Million yuan)	Sales Income (Million yuan)	Selling tax (Million yuan)	Turnover days of circulating capital (day)	Capital pay-off rate (%)）
1990	2658	63	282	1109	3247	829	75	41.4
1991	2734	60	242	1121	3464	879	79	41.7
1992	3133	60	327	1399	4227	1072	51	56
1993	4483	48	503	2223	6338	1720	67	52.7
1994	5702	45	701	2605	8343	1243	100	28.7
1995	5273	46	182	2060	9008	1062	99	16.7
1996	5689	54	180	1970	9884	1011	99	15
1997	7051	51	719	3023	11814	1178	72	23.6

In the eight economic indicators, except the cost of 100 yuan output value and turnover days of circulating capital are cost indicators, the other six are benefit indicators. So the correlation coefficient matrix can be calculated by the previous formulas. The results are shown as follows:

Table 2 The resolution coefficient of each index

Year	Gross Industrial Output Value (Million yuan)	The cost of 100 yuan output value (yuan)	Total profit (Million yuan)	Total of profit and tax (Million yuan)	Sales Income (Million yuan)	Selling tax (Million yuan)	Turnover days of circulating capital (day)	Capital pay-off rate (%)
1990	1.0000	1.0000	0.7254	1.0000	1.0000	1.0000	0.4949	0.4371
1991	0.9666	0.7500	0.8130	0.9876	0.9518	0.8991	0.5385	0.4343
1992	0.8222	0.7500	0.6471	0.7674	0.8138	0.6471	0.3333	0.3333
1993	0.5462	0.3750	0.4549	0.4621	0.5809	0.3333	0.4261	0.3522
1994	0.4191	0.3333	0.3409	0.3901	0.4567	0.5183	1.0000	0.5994
1995	0.4565	0.3462	0.9926	0.5016	0.4265	0.6566	0.9608	0.9234
1996	0.4202	0.5000	1.0000	0.5264	0.3922	0.7100	0.9608	1.0000
1997	0.3333	0.4286	0.3333	0.3333	0.3333	0.5607	0.4667	0.7045

Then we can get the weights which are shown as follow:

Table 3 The weights

ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8
0.9110	0.9004	0.8681	0.8083	0.8528	0.8726	0.8810	0.8622

According to the weights obtained, we standardize the data and get closeness degree by double base points. The results of the comprehensive economic benefits are as follows:

Table 4 The closeness degree and the ranking

Year	Closeness degree	Ranking
1990	0.0277	8
1991	0.0356	7
1992	0.0742	4
1993	0.1217	1
1994	0.0883	2
1995	0.0696	5
1996	0.0596	6
1997	0.0801	3

The results showed that: the company's economic efficiency has gradually increased, but in 1995 and 1996 decreased. This conclusion matches with the actual situation of the company.

In order to study the differences between the two results, we calculate Spearman rank correlation coefficient of the two results and get the coefficient is $w = 0.905$. So the results of two evaluation results can be considered consistent. Then to test the above conclusions, we can do as follow:

H_0: Evaluation results are consistent;

H_1: Evaluation results are inconsistent.

We construct statistics $\chi^2 = \sum_{i=1}^{8} \frac{(a_i - b_i)^2}{a_i}$, where a_i is the sorting in [3] and b_i is the sorting in this paper. For $P(\chi_{0.975}{}^2 < \chi^2 < \chi_{0.025}{}^2) = 0.05$, we look up table and get $\chi_{0.975}{}^2 = 1.68987$, $\chi_{0.025}{}^2 = 16.0128$. So $\chi^2 = 1.95$. Can not reject the null hypothesis, that is, through non-parametric hypothesis test the two results can be considered be consistent.

Summary

Using χ^2 test method, we make a comparison between the results of this paper and the results of literature [3], which show that there is no significant difference between the two results. And by this method used in this paper, it avoids the subjective factors in process of determining the weight.

References

[1] Jingai Liu. Estimating Synthetically Economic Benefits of Corporation by Principle Multivariate Analysis. Journal of Yanbian University. Vol.24(1998), p.5

[2] Yang Zhang. Study on Project Economic Benefits Prewarning based on Gray Disaster Prediction Model. Pioneering with Science & Technology Monthly. Vol.5 (2010), p. 91

Mingyu chen. Application of AHP in the undergraduate education evaluation. Journal of Xi'an Polytechnic University. Vol.13 (2010), p. 106

[3] Zhao De-yong. A Method of Ameliorative Mult.i-objective Synthetic Evaluation Based on Entropy Weight and Its Application. Journal of Ordnance Engineering College. Vol.13 (2001), p.47

[4] Yinxin Jie. The Application of TOPSIS method on Synthetical Evaluation of Industrial. Economic Benifitis. Vol.20(1994), p. 58

doi:10.4028/www.scientific.net/AMM.50-51.770

The Research on the Stimulation of Particle Flow System Based on the Event-driven

Fan Haiyan

Department of Mathematics and Information Science Tangshan Normal University, Tangshan, China

fhy_sea@163.com

Keywords: Particle system, event-driven, the stimulation of fountain

Abstract. The particle system has unique advantages on the stimulation of the natural phenomena of rain, snow, flowing water and dust, physical phenomena and space warp. The paper discussed using the event-driven particle flow system to stimulate the fountain on the stimulation of fountain on particle system, in which there are disadvantages of by using the non-event-driven particle flow system and lacking of the effects of external factors such as resistance, gravity and the speed of wind. The event-driven particle flow system is made up of a series of events. Every particle flow emitted the particles that will experienced a series of events. In every event we can use operator to the effect of all the changes in particles. Among the tests in the events, the events are connected together. So we can stimulate the water jetting, falling, spray, ripple and the effect of disappear with the wind of the fountain. All these can prove that the event-driven particle system is necessary in producing the realistic scene.

Introduction

The simulation of natural scenery has been the focus of computer graphics. Because the shapes of the natural scenery are changing with the time, the surfaces usually contain a lot of details. So it is hard to describe them by using the traditional surface analysis. When simulation the objects such as cloud, smoke and fire, the particle system is thought as the most successful way. It is used widely in the fields of computer games, video and advertisements.

Compared with mountains, plants and cloud, the simulation of the fountain is more difficult. The reason is that the movement of the water sprayed by the fountain is changing all the times, collision with the obstacles arising splashes phenomenon. All of this is hard to be described in traditional way. Since the particle system is proposed by Reeves[1] in 1983, there are many particle systems to simulate the work of natural scenery, including the simulation of water flow. Peachey[2] and Fournier[3] use particle system to simulate the wind blowing the wave and the spray formed by the waves. Goss[4] used the particle system to simulate the locus when the ship going. Wan hua gen[5] used particle system to simulate the wind blowing the wave and the splash formed by the wave[67]. The article analysis the water flow from the nozzle will be affected by the gravity to fall, then will meet the wind and the wind will the spray to blow more smaller sprays, then the spray will fall into the pool to form ripples. The speed and quantity of the water from the fountain will change irregularly. All the effects are necessary to form the realistic scene.

The basic principle of particle system

The simulation of natural scenery has been the focus of computer graphics. Because the shapes of the natural scenery are changing with the time, the surfaces usually contain a lot of details. So it is hard to describe them by using the traditional surface analysis. When simulation the objects such as cloud,

smoke and fire, the particle system is thought as the most successful way. It is used widely in the fields of computer games, video and advertisements.

The basic thought of particle system is: taking many small particles with simple shapes as basic element, using them to describe the irregular and fuzzy objects. The particles with their life cycle will go through the three steps of the "forming", "movement and growing" and "extinguish". The particle system is a system with "life", so it can not form the instantaneous static landscape pictures, it can form a series of pictures with movement. So it is possible to stimulate the dynamic natural scenery.

The basic steps of forming pictures by using the particle system:

Forming the new particle.

Give every new particle some property.

Delete the particle that not in the period of survival.

According to the dynamic property of the particle to move and change the particle.

Display the picture made up by the particles with lives.

The particle system is taking the random process to control the number of the particles and to determine the original random property, such as the original movement direction, original size, original color, and original transparency, original shape and the period of survival. And changing the properties during the process of the movement and growing up. The random of the particle system make the stimulation of irregular fuzzy objects vey simple.

The basic principle Event-driven Particle Flow

Non event-driven particle system is referred to the cartoon system that forming the particles with time changing. It is easy and fast to set and fit for making the simple particle cartoon such as fountain, snow and rain and dust. The particle flow is a particle system driven by events. It contains a special transmitter. Every particle system is made up by many different particle flow, and the particle flows own different transmitter. The transmitter can define the behavior of the particle, set the test conditions of life, collision and speed. It can react according to the test result, set the strong flexibility and controllability fitting for making complicated particle cartoon picture such as the fragments, flame and smog.

The working method of the event-driven particle flow system. We can think that the working method of the particle system as follows: Choose some actions, using the actions to define the special condition of some particle and make a group of movements to form the event and then using the test to send the particle from the event to the other. A series of events form the particle flow. The event-driven particle flow belongs to one kind of particle system, and the particle flow is made up by a series of event and every event is made up by one of more actions.

The life cycle of the particle . The particle system is like the program that can produce the lives. The program can affect the movement of the particle, the property of the particle, testing the interact of the particle and other objects in the scenery and can define the condition and behavior of the particle in every time. The life cycle of the particle can be described the change of the particle from begin to the end, and during the time the external forces and a variety of triggers can decide the next period condition of the particles.

The basic flow of the life cycle of the particle.

The particle is set up based on special settings, but the particles have to be guided and commanded by some events.

The movement is added on some position of the particle to make the particle accelerate to move to the direction of the object. The series of movements are controlled by force.

The particle will be in some condition until an event happens.

Event test can change the current condition of the particle, they are like the triggers. When an event happens, a decision have to be made, the particle then can go into a new condition.

A new condition can change the property of the particle, such as: speed, shape, size and rotation or make the old particle to produce the new particle.

This kind of force can be all kinds of force, such as the power of wind and gravity.

A particle can be tested to hit with other object or be forbidden to move on some object.

The life cycle of the particle is another property , it can be tested, can be used to change the condition of the particle or disappear after some frame.

The particle can be pointed as any kind of material.

The key technology of the stimulation of fountain

We make a fountain in this paper using common non-event-driven particle system. Although the whole effect is pretty good, but there is fatal flaw—lacking of details, and the particle flow just have these advantages.

The realistic effect of the fountain is when the water flow coming out of the nozzle, the water flow is affected by the gravity then will meet the wind, the wind will blow the spray into smaller sprays, then the spray will fall into the pool to produce the ripples. There will be irregular changes of the speed and the number of the water. All the effects is necessary details to make the realistic scenery.

The function of the particle is very powerful, but the flow chart is not intuition like the non-event-driven particle system. The stimulation of the paper will need to use the "particle view", the "particle view" is made up of a series of "events".

Related model

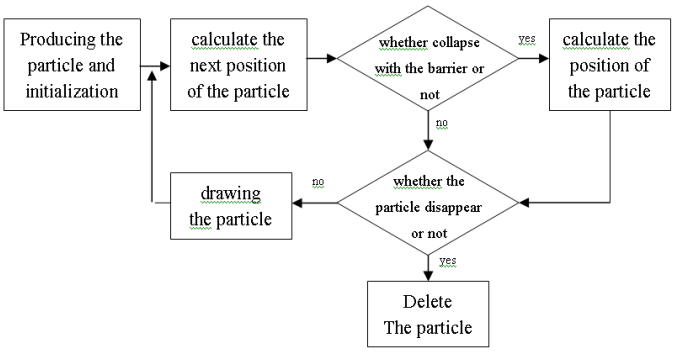


Fig. 1 Model diagram

The steps of stimulation

Set up the particle emitter. After adding a particle emitter in the scenery, in the particle view there will be an event named by particle flow source, the event is global event, the operator in the event will affect all the particles in the particle flow source. With the global event there will be an event -2. In this event there must be a Birth operator, then the particle system can launch particles normally.

Space warp. To produce the effects about the particle system, it is necessary to have the space warp. The stimulation of scenery of fountain will need three space warps, they are gravity, wind and POmniFlect.

In the non-event-driven particle system when using POmniFlect, using the tool of Biond To Space Warp to bind the particle system and POmniFlect, in the particle flow using the POmniFlect, in the particle view adding the Collision test or the Collision Spawn test for the event.

In the event of "particle emitting event" there will be particles continuously, these particles pass the test of collision test. All the particles that collapse with the space warp "water surface POmniFlect" will be passed to the event of "splash event", the event through the Spilt Amount operation to pass part of the particles to the event of ripples, in the event we need the shape mark 01 operator to stimulate the ripple effect of the water surface.

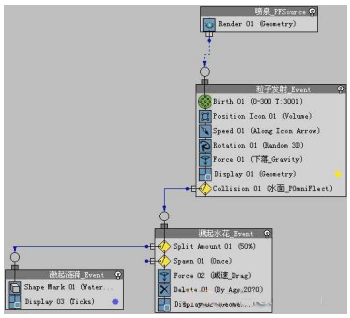


Fig. 2 the picture of related matters

In order to produce more realistic effects, the wind is necessary. The drops from the nozzle will be blown in more smaller drops, the speed of these drops will slow down and will dance with the wind. This kind of effect will be realized by using Spawn, and the falling speed of small drops will slow down and the dancing effects will use Force operator.

The last step of stimulation of fountain just as the ways used in the non-event-driven particle system, before rendering the particles we need to set the particles properly. For the non-event-driven particle system, we can set the type of the particles as Facing, but for the particle flow we still need the operator. We can add a material dynamic operator to the three events.

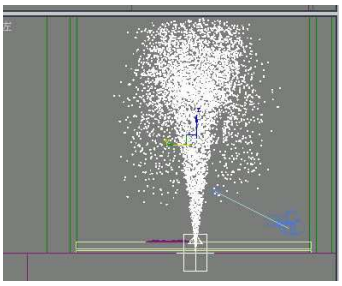


Fig. 3 The stimulation model diagram of non-event-driven particle system of fountain

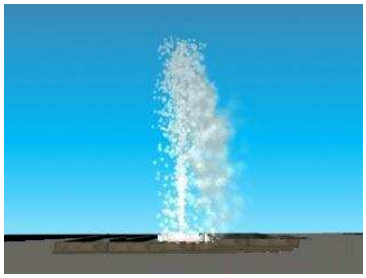


Fig. 4 The diagram of event-driven particle system of fountain

The stimulation of the fountain is almost finished, but there are a lot of places needed to be improved. Such as add Scale operator to the event of “particle emit event”, “disappear with the wind event” to make it become large during the process of movement, the strength of the gravity and wind can be adjusted more in details. We can add noise to the birth and speed of the “particle emit event” to make the speed and number of the particle be full of changes to make the fountain more realistic.

Calculation

The paper uses event-driven particle flow system to stimulate the fountain. Based on the characters of the event-driven particle flow system, we can use the test in the event or between the event to show the effects of the three factors of gravity, wind and collision to the stimulation of the fountain. It is more changeable than using the non-event-driven particle system. It supply the references for the users to stimulate the fountains or other natural scenery in different environments.

References

[1] Reeves,W.T. Particle Systems–A Technique for Modelling a Class of Fuzzy Objects. Computer Graphics,1983,17(3):359-376.

[2] Peachey D.Modeling waves and surf. ComputerGraphics,1986,20(4):65-74.

[3] Fournier A,Reeves W.T. A simple model of ocean waves. Computer Graphics,1986,20(4):75-84.

[4] Goss M E. A real time particle system for display of ship wakes. IEEE Computer Graphics and Applications,1990,10(3):30-35.

[5] Wan Hua-Gen, Jin Xiao-Gang, Peng Qun-Sheng. Physics based real time animation of fountain. Chinese Journal of Computers, 1998, 21.

[6] Peachey D．Modeling waves and surf．Computer Graphics，1986，20(4)：65—74.

[7] Fournier A，Reeves W T．A simple model of ocean waves．Computer Graphics，1986，20(4)：75—84.

[8] Ma jun, Zhu hengjun.the Stimulation of fountain based on the dynamic texture and particle system [J]. Beijing Jiaotong university academic journal, 2005, 29(1): 90-95.

[9] Zhao jingmi, Zhanghui, Zheng guoqin. The stimulation of the fountain based on the particle system [J]. The application research of the computer, 2006, (1): 244-246.

[10] Zhao jingmi, Zhanghui, Zheng guoqin. The stimulation of the fountain based on the particle system [J]. The application research of the computer, 2006, (1): 244-246.

doi:10.4028/www.scientific.net/AMM.50-51.775

Research on Key Factors in Shader Programming of Dynamic Water Surface

Yufen Feng

Department of Mathematics and Information Science Tangshan Normal University,Tangshan China

fengyufen_ts@163.com

Keywords: Dyamatic water surface,shader,texture mapping ,virtools.

Abstract. Research on real time dynamic water surface is a hotspot in the study of virtual reality with very broad application prospect. With the rapid development of calculating ability of GPU (Graphics Processing Unit) and the development of HLSL (High-Level Shader Language), the tasks such as vertex shader ,pixel shader, calculation of normal vector, dynamic texture mapping and calculation of water reflections and refractions can be done by GPU. In our research we studied the key factors affecting the effect of dynamic water surface and indicated that the writer presents some opinions which can generate more accurate reflections and refractions of water through adjusting light intensity and colorimchy by lighting and perturbing 2D texture mapping with normal mapping. A shader program in Virtools has been written and applied in virtual reality project. In the actual operation the effect is real and the rendering speed has greatly increased.

Introduction

In virtual reality system, to generate real water surface is very difficult because the color of water depends largely on the surface of the light reflection and refraction and the different observations have different reflection and refraction effects. In the large water surface, the effect of reflection will be weaker nearly and stronger in the distance, the effect of refractive is stronger in nearby surface than far into the water, the color of water is darker near the surface and lighter in the distance. Because of the fluctuation of water by force, in the near water surface the ripple is clear and obvious and in the distance it is more ambiguous. Shader programs can show physical characteristics of the water model and texture changing incisively and vividly according to various optical principles, using the surface processing capacity of GPU.

In recent years, many people study dynamic water surface with a variety of graphical image processing technology, designing and implement with different algorithms and languages[1-5].

Li Guangxin and Ding Zhenguo present a method simulating realism of water surface based on the perlin noise and B-spline surfaces, noise is applied the characteristic control points of B-spline surface to control height field indirectly, using CubeMap texture technology to simulate reflection, using the tracking refraction vector method to simulate the refraction and implemented with OpenGL graphics interface [6].In the dynamic simulation of large water surface ,Cheng Tiantian absorbs Johanson's grid projection method , combines the feature of GPU which supporting shader model of 3.0, uses the function of vertex texture fetching, visits the high field of lake in programmable vertex shader, solves the 2D equation of wave by using finite difference method and simulated raindrops falling into the water [7].Zhao Ping, Chen Dingfang research on water flow effects in the visual scene system and multiple resolution texture techniques based on shader technique of Cg, stimulating the wave, the reflection and refraction of water by using the techniques of bump texture and HDR(High Dynamic Range) [8]. Cheng YongXin improves the normal mapping and HDR technology by applying dynamic texture to normal mapping and using the powerful functions of GPU to generate shader in the real rendering of dynamic water[9].Obviously, a lot of research work of dynamic water has been done by the predecessors. However, in the actual application, when we design a new virtual reality project, we will find that many shader programs written before can't be used or may not be suitable for the environment. In order to express the

physical characteristics of water and dynamic surface texture fully in different environment and different scenes, the author wrote shader program based on HLSL in virtools combining with the optical effect of water through studying several key factors affecting dynamic water.

Development Technology

The development of shader. Shader drive the programmable graphic pipeline, and define the final surface properties of an object. It provides a series of instructions programmers which can finish 3d image rendering through executing the instructions on programmable pipeline of GPU[9]. Technology of fixed pipeline of 3D graphics card is used before, now programmable pipeline is a kind of mainstream technologies. Technology of fixed pipeline is that programmer just set the vertex position, normal line, UV coordinates and light into the graphics card by providing API and setting good parameters, then graphics card can render a picture. Shader is a kind of technology applying to 3D graphics card of programmable pipeline; there are two kinds of shaders . One is vertex shader, it is replaced fixed rendering pipeline in the transformation and lighting part, programmers can control vertex transformation and lighting. The other is pixel shader, it replaced the rasterization part of rendering in fixed pipeline, programmers can control pixel color and texture sampling etc[10-13].

shader in virtools. Virtools software is a rich interaction module of real-time 3D virtual scene editing software. It can produce many 3D products with different uses. Virtools software supports the senior rendering language of DirectX9 fully. It develops a shader reference for researcher, all kinds of usages and the parameters value of the state can be seen[14].Writing programs for each shader is that some parameters are filled in, these parameters can be filled in material properties page and can be filled by BB(Building Block) [15]. A shader program can contain one or more techniques, but there is only one technique applying to render materials. Each technique may contain one or more passes, each pass executes sequently according to their order defined. Each vertex and each pixel is execute in every pass in order to complete different purpose of render.

Main Factors and the Relevant Implementation Codes

Lighting. It is the indispensable content that we add lighting to virtual scene. Lighting model is that which way is adopted to calculate shade information of the vertex (or pixels) according to the direction of light, the vertex (or light pixels) position, the vertex normal information etc. The common Phong model is used in shader usually.

$$I = I_{Ambient} + I_{Diffuse} + I_{Specular} \tag{1}$$

Ambient light is that the light from every direction projects onto an object, it is also called the global light. This type of light arrives the surface of object by other surface reflection, illuminates the whole scene, and combines the light intensity and chromaticity.

$$I_{Ambient} = Ai * Ac \tag{2}$$

Diffuse light propagates along the specific direction . When it reaches a surface, it reflects along each direction and has no relations with the observer's position. Fig. 1 is diagram of diffuse light vector.

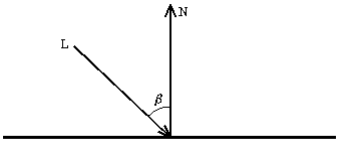
L is diffuse light vector, N is normal line, is the angle between L and N, therefore

$$I_{Diffuse} = \|N\| * \|L\| * \cos\beta * n \tag{3}$$

If we take N and L as unit vectors,

Then $I_{Diffuse} = \cos\theta$ (4)

Specular light reflectes along the particular direction, When the kind of light reaches a surface, it will reflect along another direction strictly, thus forms high brightness lluminatation which can be seen from a certain angle range. So specular light has relation with the angle between incidence light and the observer's position, It is usually used to simulate the highlights of objects. Fig.2 is the diagram of specular light vector.



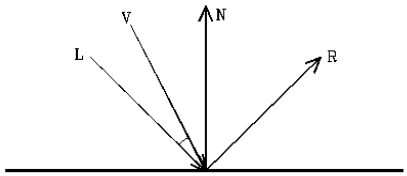


Fig. 1 diagram of diffuse light vector Fig. 2 diagram of specular light vector

L is a incident ray vector, R is a reflecting vector, V is the observer view vector, is the angel between R and V. The value of θ is smaller , the specular reflection is stronger. N is reflection coefficient of different materials, so the specular model is showed as follow.

$$I_{Specular} = \|R\| * \|V\| * \cos\theta * n \tag{5}$$

If we take R and V as unit vectors,

Then $I_{specular} = \cos\theta$ (6)

common lighting model

$$I = Ai * Ac + Di * Dc * \cos\beta + Si * Sc * \cos\theta * n \tag{7}$$

Reflection. Reflection is one of the most basic optical effects of water. From the lighting model we studied in the previous chapter, we have a conclusion that the water surface reflection is divided into two parts; they are global reflection and local reflection. Object reflected of global reflection can be considered infinity, because the reflective surfaces is far away, so the effects of reflection and reflection associate with the angle between the reflection vector and the surface. For example, the sky reflected is the global reflection, it can be seen that it doesn't move with the viewpoint. Object reflected of local reflection is the near object from the surface. It is not only related to the incident angle, but also related to the reflection coordinate of surface. If using light tracing algorithm to calculate local reflection, the rendering speed will be reduce greatly. In our research , we adopt 2D reflection texture to simplify the local reflection, the reflection texture regards the surface as a smooth mirror to calculate. First render the scenery reflected on the surface on a texture map upside down, then affix the map to the polygon of the surface for the approximate effect of local reflection. Because the water is fluctuations, so the reflection is distorted through different directions of normal vector . We use normal mapping, adjust the texture coordinates with pixel shader and perturb the map, we can see the finally effect of the water. At last we use the method which combine the texture coordinates transformed from projection coordinate with reflection texture and write shader code through analyzing the basic principle of refraction.

The main codes are showed as fo11ows:

```
// The (1,1) projection coordinate is transformed into (0,1) texture coordinate
float bumpFactor;//factor of water-waving
texture texRef; // reflection texture
sampler2D samplerRef=sampler_state
{       texture=<texRef>;
        MipFilter=LINEAR;
        MinFilter=LINEAR;
        MagFilter=LINEAR;};
```

```
texture texNorm; //normal texture for generating  waves
sampler2D samplerNorm=sampler_state
{       texture=<texNorm>;
        MipFilter=LINEAR;
        MinFilter=LINEAR;
        MagFilter=LINEAR;
};
```

Refraction . When we observe the underwater objects, the light which can be seen will deviates from an angle, this kind of phenomenon is refraction. The phenomenon generation is caused by the light reaches into one's eyes after it reaches the underwater objects and then water reflects to the surface. The water is shallower, the light to underwater objects is stronger, the refraction effect is more obvious. At the same time, the effect of light refraction and the viewpoint position also have relations. The viewpoint position from surface is closer; the refraction effect is more obvious.

The refraction principle of incident light into the water is as shown in Fig. 3, The corresponding formula was brought forward in the 19th century by Fresnel.

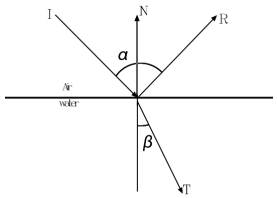


Fig. 3 figure of refraction principle

Where is refractive coefficient of the medium of incident light, is refractive coefficient of the medium of refraction light, (Light in air refractive coefficient is 1.0003, light in water refractive coefficient is 1.3333) is incident angle, is refractive angle. According to the principle of light refraction, we know how to deal with 2D texture to generate refractive effect for underwater objects, and rendering realistic effects finally. Reflection and refraction are the same; the paper adopts texture mapping perturbation technology to simulate refraction. First render a texture mapping for the whole scene, mark the reflected objects, then use normal mapping disturb the marked texture, according to the color information of normal mapping, change the texture coordinates and add displacement for the texture coordinates.

We have written a shader program of dynamic water surface in virtools. Fig. 4(a) and (b) are effect figures of reflection and refraction in the virtual community roaming system. In the shimmering waters there are reflections of the buildings, the trees and the stones, the refractive effect of stone in water is clearly visible.

Fig. 4 (a) effects of reflection and refraction (b) effects of reflection and refraction

Conclusion

In our paper, we studied the key factors that had influence on dynamic water surface. According to the lighting model we adjust the light intensity and chromaticity to realize overwater reflection and refraction by normal mapping perturbing 2D texture mapping. It is of the realistic water surface and high real-time rendering speed when you apply it in the virtual community roaming system by

programming the shader in virtools. We need to research the real traces when the shop is across the water for the future. And we need also to make ripples effects when the rain drops. It is mysterious for us that one can obtain the intuitive visual feedback of code by shader programming at once. You can also realize a tree sway with real-time dynamic swaying effect of leafs on the lake by shader programming. But it needs to consider the necessity using shader in order to utilize the computer system resources rationally.

References

[1] Wang Shengzheng,in:A Novel Approach to Dynamic Ocean Surfaces Simulation for Realtime 3D.edited by Computer Appliation(2007), p. 1147–1149.

[2] Yang Chun etc., in:Research on Experimentation Data Analysis of Virtual Experiment Based on 3D and Virtools Technologies,edited by Computer Engineering and Design,(2007),p.2589-2591.

[3] Chen Yongxin,in:Research and Realization at Real-time Render of Large-scale Water Surface and Scene Magement,edit by Master Dissertation,University of Electronic Science and Technology of China(2009).

[4] Cem Yukel,Donald H House,John C Keyser,in:Wave Particle.ACM Transactions on Graphics,Article 99,Jul 2007,26(3)

[5] Enright D,Marschner 95,Fedkiw R,in:Amination and Rendering of Complex Water Surface.edited by ACM Transactions on Graphics,Vol.21,No.3(2007), p.736-744.

[6] Li Guangxin ,Ding Zhenguo, in:A Modeling Alorithm for Realistic Water Surface Simulation in Vistual Environment ,edited by Journal of Computer Research and Development,Vol 41,No 9(2004) ,p.1580-1585.

[7] Cheng Tiantian,in:Research on Dynamic Water Scene Modeling Method for Tai Lake Water Area, edited by Master Dissertation ,Suzhou University(2008)

[8] Zhao Ping ,Chen Dingfang etc. in:the Application of Shader in the Development of Visual Scence Based on Cg, edited by Journal of Hubei University of Technology,Vol 22,No 3(2007) ,p.65-67.

[9] Cheng Yongxin,in: The Research and Implementation of the Large Water Surface Realtime Render and Scene Management In 3D Engine,edited by Master Dissertation,University of Electronic Science and Technology of China(2009)

[10] Tan Xiaohui etc. , in:Real Time Simulation of Large Scale Water Surface Based on GPU. edited by Computer Simulation,Vol 26,No 9(2009) p.211–214.

[11] Frank D.Luna Introduction to 3D Game Programming with DirectX 9.0 ,Copyright 2003 by Wordware Publishing.

[12] Wang Daochen,Wan Wanggen etc.in:Real-time Rendering Alogrithm for Water Surface Based on GPU,edited by Computer Engneering, Vol 34,No 20(2008), p.233–237.

[13] Zhang Long etc., in: Real-Time Simulation of Dynamic Wetland Scence, edited by Journal of Computer-Aided Design&Computer Graphics, Vol 20,No 8(2008),p. 1007–1010 .

[14] Fu Zhiyong ,Gao Ming, Three Dimensional Game Design by Virtools Dev 4.0,Beijing:Tsinghua University Press,(2008).

[15] Wang Dandong,Xu Yingxin etc.in: Treasury of Knowledge of Designer on 3D Game,edited by Ceap Electronic & Audiovisual Press（2009）

doi:10.4028/www.scientific.net/AMM.50-51.780

Study on the modification of polishing putty by inorganic filler

Bi Chunbo

College of Light Industry, Hebei Polytechnic University, Tangshan 063000, Hebei, China, bicbo2005@126.com

Keywords: modification; couple; putty; inorganic filler

Abstract. In order to decrease the cost of polishing putty, the interrelations between talcum powder as the filler and KH550, KH560, or silane compound coupler were investigated by atomic force microscopy (AFM) and Fourier transform infrared spectrum (FTIR). It was proved that the effect of Modification of KH560 coupler is the best.

Introduction

Polishing putty with titanium dioxide as the inorganic filler has excellent properties., Although the putty matching with curtain wall is favorite in the market, it has quite high cost compared with other putty with inorganic filler. In order to decrease the cost of putty, the polishing putty was modified by talcum powder as inorganic filler.

Experimental

According to the literatures, silane coupling agents are most popular, the consumption is also the largest, and they are mainly used for the surface treatment of granular filler with silicon [1].
The inorganic filler was modified via wet processing method [2-5]. Silane coupling agent was firstly dissolved in ethanol, then reacted with inorganic filler, ethanol was removed by filtering and drying after full reaction, and then the treated filler was obtained. The advantages of the processing method are the best coupling effects, the inorganic filler can be covered by the inorganic filler, and consequently, the error was reduced..

The processing methods of coupling agent KH560 and KH550

Ethanol is used for the dilution of KH550 KH560. 25% solution of silane coupling agent was prepared by mixing ethanol and silane coupling agent, the solution was sprayed on the stirring filler, and then the treated filler was baked with 1000w an incandescent lamp for 15min.

Processing methods of silane coupling agent composite

The filler was heated to 100℃ in high-speed mixer, the water content was less than 0.3% by stirring and drying, and then, coupling agents were cut into pieces. According to the dosage, coupling agents were put into the treated filler slowly in batches, and were activated for 6-8 min. In this experiment, the coupling agent were KH550, KH560, silane coupling agent composite.

3 Silane coupling agent modification mechanism analysis of talcum powder

Modified micro distribution analysis of talcum powder

In the wet modification after drying process, and the particle size and distribution are characterized and assessment of the effect of surface modification, is one of important index of surface modification process and the formula of important macroeconomic-control means. This research is characterized by atomic force microscope to the modification effect. Talcum powder On the modified

or unmodified talcum powder and different modifier modification effect, and choose better modifier. Pure talcum powder to join in accordance with the above method is modified agents after treatment, by atomic force microscope to sample were tested. This study adopts Shanghai love built nano science and technology development Co., LTD Ⅲ - AJ scanning probe microscopy SPM type on the control box were characterized. Among them, the coupling agent for 1% of talcum powder. Characterization is shown in figure 1 and figure 2 results Shown.

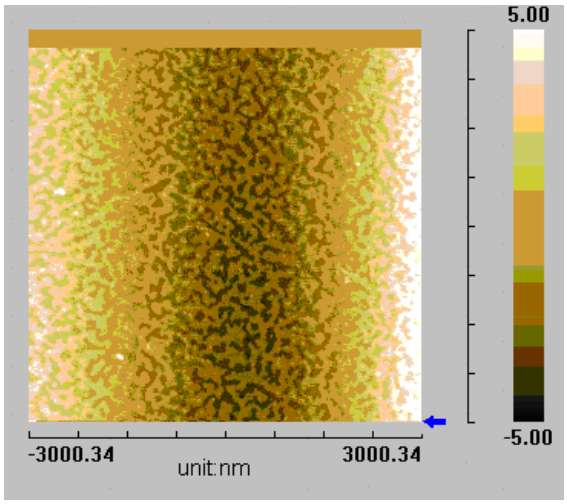
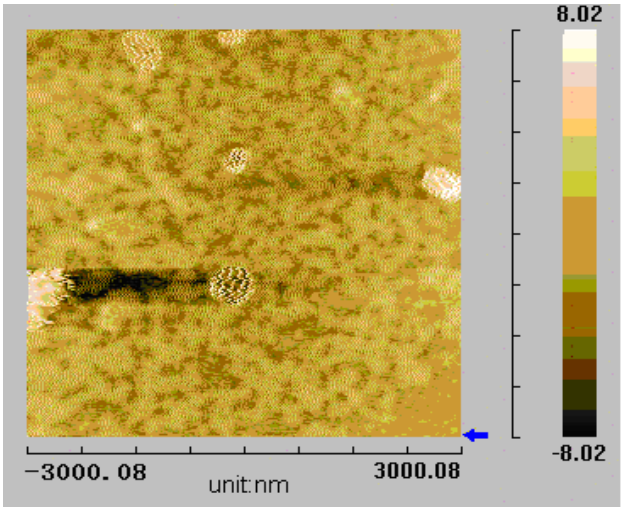
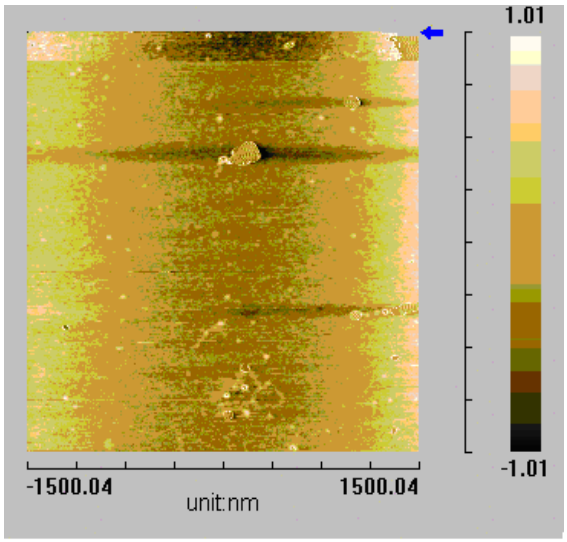


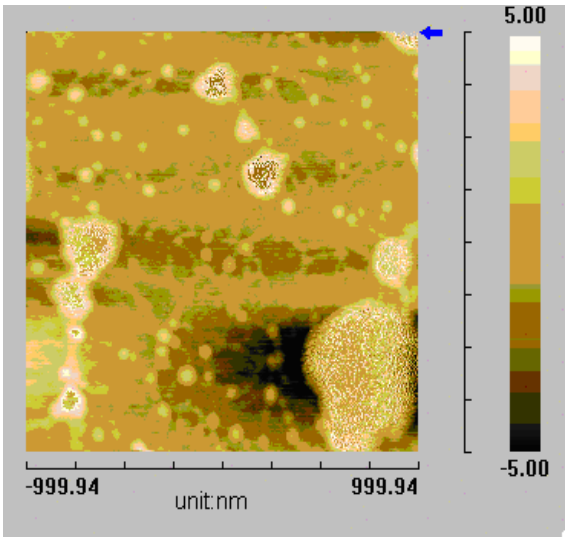
Fig.1 Photograph of the uncoated talcum powder



(a) Talc silane coupling agent added composite photos



(b) KH550 talc added photos



(c) Photos added KH560 talc

Fig.2 Microstructure of the coated talcum powder

From figure 1 and figure 2 can see, modifier to change the shape, but talcum powder silane coupling agent and KH550 compound modification effect of talcum powder, and the modification of talcum powder KH560 effect more apparent. By coupling agent mechanism analysis of silane coupling agent, which formed silicon alcohol and hydrolysis and inorganic powder particles on the surface of the hydroxyl, hydrogen and condensation into -- the SiO - M or inorganic powder particles (M said surface) [6]. While all the silane molecule silicon alcohol associating mutual mesh structure formation in the film, covering the surface of talcum powder, organic and inorganic powder talcum powder surface. After the modification of talcum powder structure, distribution, it is modified resin dispersion of talcum powder in. Unmodified talcum powder particles due to exist between the van der Waals force and producing reunite phenomenon, resulting in polymer particle, the dispersion viscosity.

The infrared spectrum analysis of samples of talcum powder

Infrared spectroscopy (IR) is the test sample composition and structure is one of the most important methods, it has been widely applied to identify the additive, polymer materials, packing, plasticizer monomer. Fourier transform infrared spectroscopy (FTIR) instrument has adequate sensitivity and high selectivity, it has simple operation, convenient system, which can be used to study the characteristics of the material surface structure and chemical reaction. According to the sample appearance, high-translucent characteristics and the required information, by Fourier transform infrared spectroscopy analysis coupling agent of inorganic packing surface structure has reported [7-9]. Using IR360 ft-series type infrared spectrometer Infrared spectra of the sample.。

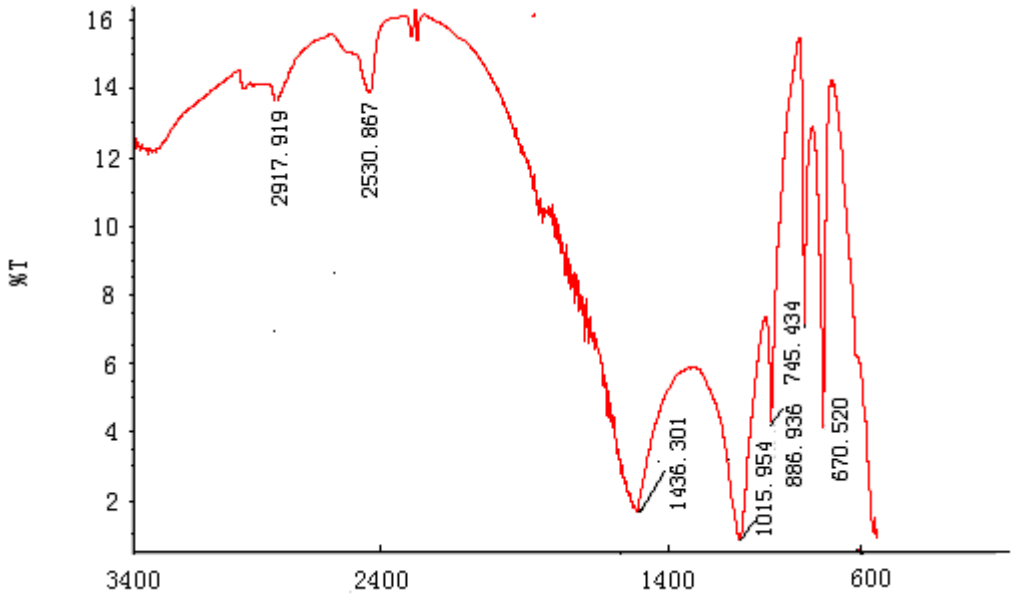
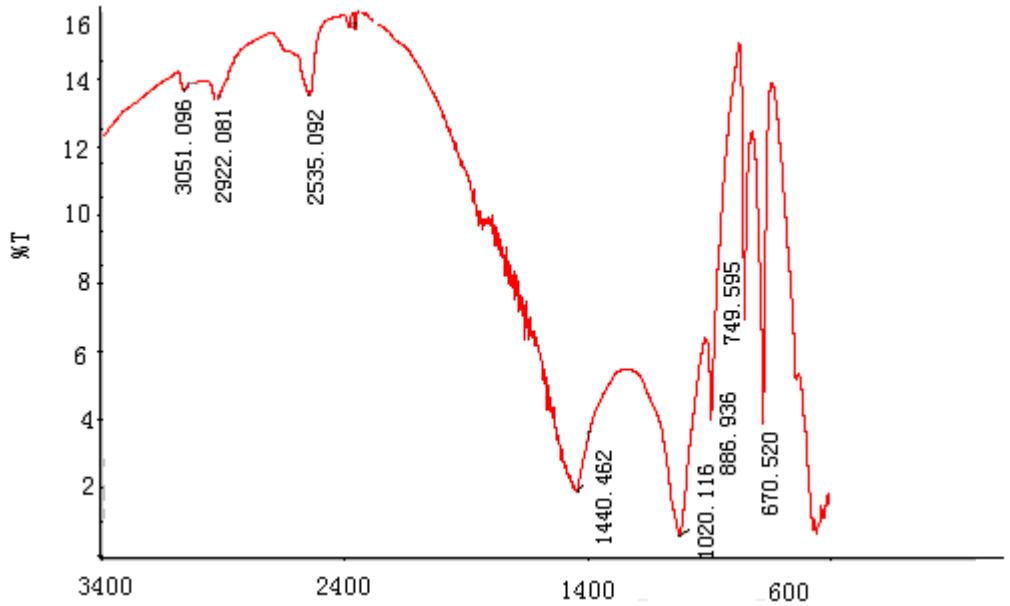
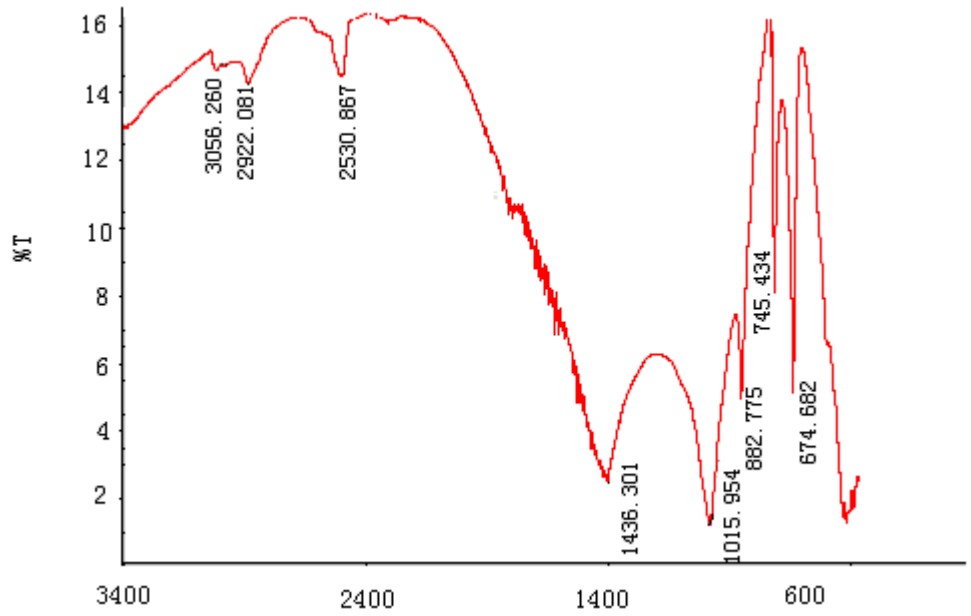


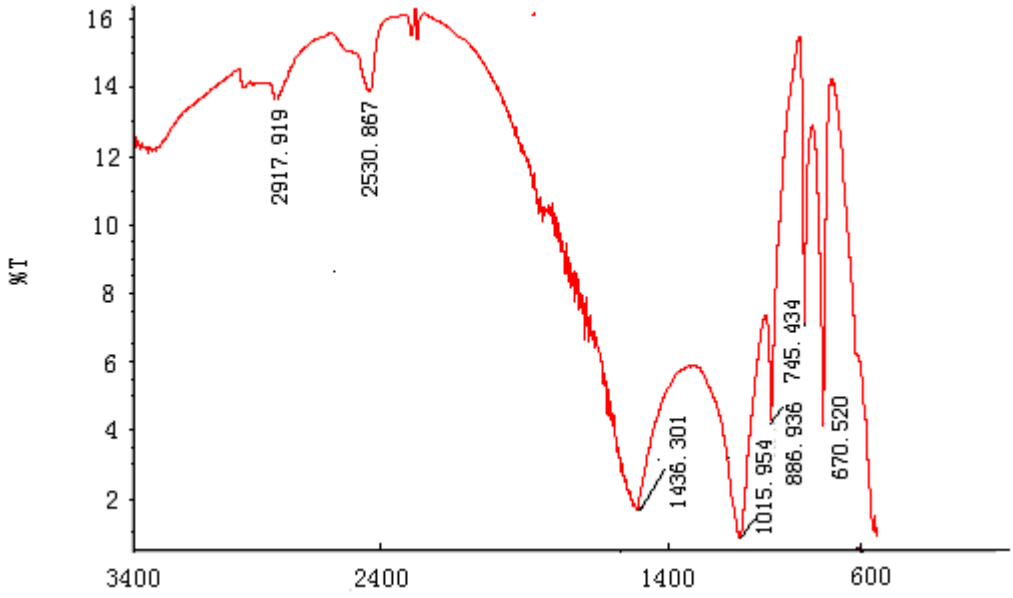
Fig.3 FTIR of the uncoated talcum powder



(a) adding KH550 infrared spectra of talc



(b) adding KH560 infrared spectra of talc



(c) the silane coupling agent added composite infrared spectra of talc

Fig.4 FTIR of the coated talcum powder

Figure 3 and 4 are before the coupling agent talcum powder and processing of infrared spectra of talcum powder. Figure 3 can be seen by at 2917.919 cm, wave number 1 - nearby existing O - H, while in figure 4 (a) and (b) wave number into 2922.081 cm - 1, this shows that after modification O - H keys have energy changes that occurred in KH560 adsorption, the talcum powder surface.

By 3 and 4 (c), silane composite with coupling agent modification of spectra and talcum powder unmodified chart no changes, also confirmed that figure 2 (a) no change in the structure of talcum powder, also showed the silane coupling agent fails to reach the compound modification effect. Figure 4 (a) and (b) was not characteristic peak value. The KH550 KH560, except with the processing of talcum powder itself characteristic peak in the infrared, 3000cm - 1, a regional KH560 c-h keys on the vibration absorption peaks, the telescopic KH560 that has to physical or chemical way adhere to the surface of talcum powder. From the above phenomenon and analysis in the KH550 KH560 knowable, talcum powder surface formed a physical or chemical combination. According to the microscopic photographs can judge KH560 atoms KH560 was formed in talcum powder surface chemical combination layer, happened chemical adsorption or chemical reactions.

For KH550, although use KH550 modified ir talcum powder has changed, but according to the chart 2 (b), and it is not up to the requirements of the research work, and through the analysis of the end of may not thoroughly, because bake the existence of impurities in infrared spectra of change.

Conclusion

The titanium oxide powder can be replaced by modified talcum powder. the atomic force microscope and Fourier transforms the infrared spectrum (FTIR) are used to study the interaction between the talcum powder and KH550, KH560, the silicon hydride compound coupling; Modified effect of the KH560 coupling is confirmed to be best.

References:

[1] WangShuJiang NiYuDe coating process, the census. [M].*beijing: chemical technology press,*1997:228.
[2]M. R, Turner, E.Duguet and C.Labrugere. Charaterization of silane modified Zro_2 powder surfaces. *Surface Analysis*[J],1997,25:917~923.
[3]Zoltan Demjen, Bela Pukanszk, Eniko Foldes etc. Interaction of Silane Couping Agents with calcium carbonate.*Journal of Colloid and Interface Science*[J],1997,190:427~436.
[4]Li Beixing,Zhang wensheng. Chenmical Reaction Between Ployvinyl Alcoho and Titanate Couping Agent with X-Ray Photoelctron Spectroscopy.*Journal of Wuhan University of Technology*[J],2003,18(2):71~74.
[5]R.Doufnical,F, Chebira,N.Haddaoui.*Effect of Titanate Couping Agent on the Mechanical Properties of Materials*[J],2003,52(11):967~984.
[6]ZhengShuiLin. Powder surface modification [M], Beijing: *China building materials industry press*,2003,8.
[7]LiuYuJuan by FTIR and HRPGC method. The study of silane coupling agent/inorganic packing interface [D].*electronics. Beijing: tsinghua university,*2001.
[8]JingXuYing etc. Ir guide [J].*journal of tianjin: tianjin science and technology publishing house*,1992.
[9]WangXueFei. Titanate coupling agent with the preparation and the mechanism of inorganic packing study [D].*electronics. Wuhan: wuhan university of technology of master degree theses.*

doi:10.4028/www.scientific.net/AMM.50-51.785

Numerical Studies on Resistive Wall Instabilities in Cylindrical Plasma Confined by Surface Current★

Shaoyan Cui[1,a] and Peng Xie[1,b]

[1]School of Mathematics and Information, Ludong University, Yantai 264025, China

[a]shycui@ldu.edu.cn, [b]ytxiepeng@yahoo.com.cn

Keywords: Resistive wall mode, Cylindrical model, Plasma flow, Surface current

Abstract. The stability of resistive wall mode is studied in cylindrical plasma confined by surface current, which is Dirac δ-function distribution. For Dirac δ-function distribution case, it is shown that the perturbations oscillate and even decline wherever the initial perturbation seed is placed. The whole system is stable and the plasma flow has little effect on it.

Introduction

The stability of large-scale magnetohydrodynamic (MHD) modes is crucial to the magnetic confinement of toroidal plasmas[1]. In tokamaks, the maximum achievable plasma β (the pressure ratio of the plasma to the magnetic field) is mainly limited by low mode number external kink modes[2] which can however be stabilized by a perfectly conducting wall of the plasma confinement device. Nevertheless the finite resistivity of the real conducting wall can convert the external kink mode into the slowly growing resistive wall mode (RWM) [3-7]. The RWM can then partially brings the instability back and sets a limit for plasma β value. Thus study of the resistive wall mode is important for the advanced tokamaks. In order to control the resistive wall mode, relative rotation between the plasma and the conducting wall or control coil feedback can be introduced[8-10]. On the other hand however, the supersonic rotation of the plasma may generate a new type of flow-driven resistive wall instability[11-13] introduced first in a slab model. It has been shown that the resistive wall instability can occur if the flow is super-Alfvénic in incompressible plasmas[11,12], or supersonic in compressible plasmas[12,13].

More realistically, the cylindrical geometry, instead of the slab, should be applied to discuss the resistive wall instabilities. In such geometry, for the initial equilibrium, a uniform axial plasma flow and a surface current on the plasma surface were considered with an ideal MHD model [14]. Nevertheless, in the instability analysis, the incompressible condition, i.e., $c_s \to \infty$, was used. Thus, it is necessary to reconsider the resistive wall instability of compressible plasmas in the cylindrical geometry with a surface current distribution. In this paper, we consider a δ-function surface current.

Cylindrical Geometry and Initial Conditions

We consider a cylindrical geometry shown in Fig.1. The plasma is in the cylinder with a radius a and assumed in a stationary equilibrium with a uniform flow velocity V_0 parallel to the axis. A current sheet is generated at the plasma surface $r = a$ to confine the plasma. And, the system is surrounded by a resistive wall with radius b and the thin wall thickness $\delta << a, b$.

★Project supported by Science Fund of Ludong University (No.LZ20082704, No.LY20082704) and Promotive research fund for excellent young and middle-aged scientisits of Shandong Province(No. BS2010NJ003).

[1] Corresponding Author: shycui@ldu.edu.cn

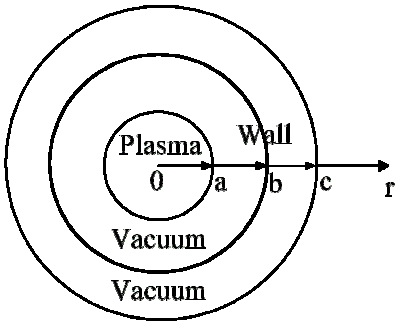


Fig. 1. The cross-section of the cylindrical model

The distribution of the surface current $\mathbf{J_0}$ parallel to the axis of the cylinder can either be a Dirac δ-function form or with a small not finite thickness. Thus, for the first case,

$$\mathbf{J_0}(r) = \hat{z} J_0 a \delta(r-a), \tag{1}$$

the magnetic field produced by this surface current is then

$$\mathbf{B_0}(r) = \begin{cases} 0 & 0 \le r < a \\ \dfrac{\mu_0 a^2}{r} J_0 & r \ge a \end{cases}. \tag{2}$$

A constant axial magnetic field B_{0z} can also be applied. Without the poloidal surface current, it is assumed that $B_{0z} = B_0 = const.$. Then the pressure has to be balanced by azimuthal magnetic field (2), and derived from $(\nabla P)_r = -J_z B_\theta$ to get

$$P(r) = \begin{cases} P_0 & 0 \le r < a \\ 0 & r > a \end{cases}. \tag{3}$$

with $P_0 = \dfrac{\mu_0 J_0^2 a^2}{2} = \dfrac{B_{0\theta}^2(a)}{2\mu_0}$.

Physical Models

The linear magnetohydrodynamics (MHD) equations of compressible plasmas are as follows,

$$\frac{\partial \rho_1}{\partial t} + \rho_0(r)\nabla \cdot \mathbf{V}_1 + \mathbf{V_0} \cdot \nabla \rho_1 + \mathbf{V}_1 \cdot \nabla \rho_0(r) = 0 \tag{4}$$

$$\rho_0(r)\frac{\partial \mathbf{V_1}}{\partial t} + \rho_0(r)(\mathbf{V_0} \cdot \nabla)\mathbf{V}_1 = -\nabla P_1 + \frac{1}{\mu_0}\left((\nabla \times \mathbf{B}_1) \times \mathbf{B_0} + (\nabla \times \mathbf{B}_0) \times \mathbf{B}_1\right) + \nu \nabla^2 \mathbf{V}_1 \tag{5}$$

$$\frac{\partial \mathbf{B_1}}{\partial t} = \nabla \times (\mathbf{V}_1 \times \mathbf{B_0}) + \nabla \times (\mathbf{V}_0 \times \mathbf{B}_1) + \eta \nabla^2 \mathbf{B}_1 \tag{6}$$

$$P_1 = c_s^2 \rho_1 \tag{7}$$

with the subscripts "0" for the initial equilibrium quantities and "1" for the perturbed quantities, and c_s the sound speed $(P_0/\rho_0)^{1/2}$. The Fourier form of the perturbations is denoted as $f_1(r,\theta,z,t) = \tilde{f}_1(r,t)\exp(i(kz+m\theta))$.

In the vacuum regions, the perturbed magnetic field satisfies $\nabla \times \mathbf{B}_1 = 0$ and $\nabla \cdot \mathbf{B}_1 = 0$ and then is represented by the magnetic flux ψ, i.e.,

$$\mathbf{B}_1^V = \nabla \psi, \tag{8}$$

with

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \psi}{\partial r}\right) - \left(k^2 + \frac{m^2}{r^2}\right)\psi = 0. \tag{9}$$

In the numerical calculation, the variables are normalized by the plasma radius a, the uniform density ρ_0, and the axial magnetic field $B_{0z} = B_0$. Therefore, the plasma flow is scaled by the Alfvén velocity $V_A = B_0 / \sqrt{\mu_0 \rho_0}$, the time by $\tau_A = a / V_A$, and the magnetic flux by $B_0 a$.

Boundary Conditions

The boundary conditions are needed to connect the internal plasma region and external vacuum regions and the wall. The first boundary condition on the plasma-vacuum interface is,

$$\left[p_1 + B_{0\theta}^p B_{1\theta}^p / \mu_0 + B_{0z}^p B_{1z}^p / \mu_0 \right]_{r=a-0} - \left[B_{0\theta}^V B_{1\theta}^V / \mu_0 + B_{0z}^V B_{1\theta}^V / \mu_0 \right]_{r=a+0}$$
$$= \frac{\xi_r}{2\mu_0} \left(\left[\frac{\partial (B_{0\theta}^V)^2}{\partial r} \right]_{r=a+0} - \left[\frac{\partial (B_{0\theta}^p)^2}{\partial r} \right]_{r=a-0} \right), \tag{10}$$

where ξ_r is the radial displacement. The second boundary condition on this interface is the continuity of the tangential electric field,

$$B_{1r}^V = i(B_{0\theta}^V m / r + k B_{0z}^V) \xi_r \tag{11}$$

Substituting Eq. (11) into Eq. (10) we can obtain the combined boundary condition on the interface,

$$\left[p_1 + B_{0\theta}^p B_{1\theta}^p / \mu_0 + B_{0z}^p B_{1z}^p / \mu_0 \right]_{r=a-0} - \left[B_{0\theta}^V B_{1\theta}^V / \mu_0 + B_{0z}^V B_{1\theta}^V / \mu_0 \right]_{r=a+0}$$
$$= \frac{i B_{1r}}{2\mu_0 (m B_{0\theta}^V / r + k B_{0z}^V)} \left(\left[\frac{\partial (B_{0\theta}^p)^2}{\partial r} \right]_{r=a-0} - \left[\frac{\partial (B_{0\theta}^V)^2}{\partial r} \right]_{r=a+0} \right) . \tag{12}$$

On the resistive wall, using the thin wall approximation, we have

$$\left. \frac{d\psi}{dr} \right|_{b+\delta} - \left. \frac{d\psi}{dr} \right|_{b} = \frac{1}{c_w} \frac{\partial \psi}{\partial t} , \tag{13}$$

where $c_w \equiv (\sigma \mu_0 \delta)^{-1}$ is the wall penetration speed with δ the wall thickness and σ the wall conductivity. On $r = 0$, we use the symmetric boundary condition, and on the outer boundary $r = c$, the zero boundary condition is applied.

Numerical Results and Discussions

The set of equations under the appropriate boundary conditions is then solved as the initial value problem numerically. In the simulation, we choose the time step $\Delta t = 10^{-4}$, and the radial space step $\Delta r = 10^{-3}$. An initial "seed" perturbation $B_{1z} = \zeta \exp(-(r - \tilde{r})^2 / \lambda^2)$ is applied to trigger the instability, and a small magnitude of $\zeta = 10^{-4}$, with $\tilde{r}$ the radial location and $\lambda << 1$ the "size" of the "seed". To investigate the resistive wall instability, an initial axial uniform plasma flow $V_0 \hat{z}$ is applied.

The simulation parameters are introduced as follows. The lengths in Fig. 1 are scaled by the plasma radius, as $a = 1.0$, $b = 1.2$, and $c = 2.0$ respectively. The periodic boundary condition in the axial direction is applied as $\tilde{f}(z) = \tilde{f}(z + 2\pi R_0)$ with $\tilde{f}$ an arbitrary perturbation to the initial equilibrium and $R_0 = 4.0$. It then leads to the "toroidal" mode number $n = k R_0 = 4k$. The other parameters are $\alpha \equiv \beta_0^{1/2} = \mu_0 J_0 a / B_0 = 0.2$, then $\beta_0 \equiv 2\mu_0 P_0 / B_{0z}^2 = 0.04$, and the sound speed $c_s = \sqrt{\beta_0 / 2} \approx 0.14$, as well as the boundary safety factor at the plasma-vacuum interface

$$q_a = a B_{0z} / R_0 B_{0\theta}(a) = q_{a0} / (1 - 2\varepsilon / a)$$

with $q_{a0} = 1.25$, where the δ-function current corresponds to $\varepsilon \to 0$.

Since in the plasma region, the safety factor $q \to \infty$, the internal mode should be stable.

The typical safety factor $q(r)$ outside of the surface current is plotted in Fig. 2. Obviously the $q(r)$ outside of the plasma monotonously increases from the boundary value q_a. We consider the effect of the internal mode, i.e., the initial perturbation is placed in the plasma region, for example, choosing $\tilde{r} = 0.5$, the radial magnetic field perturbations evolving with time with different initial plasma flow velocities. As expected the modes oscillate only, with no evidence of growth.

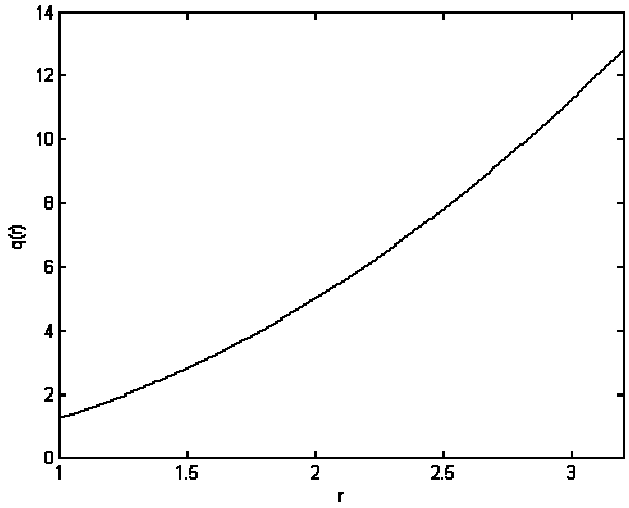


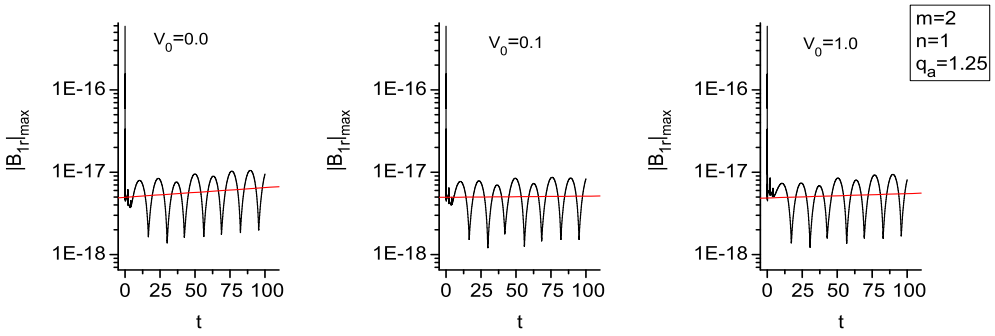
Fig. 2. The safety factor $q(r)$ distribution outside of the surface current

In general, the most unstable external modes in the static equilibrium are $m = 2$ modes. Thus, we first test the $m/n = 2/1$ mode, with $m = 2$, $k = n/4 = 0.25$ and the initial "seed" perturbation on the $q(r) = 2$ surface in the vacuum region. As seen in Fig. 3(a), the radial field perturbations are also oscillating in a very low magnitude of less than 10^{-17}. To the static equilibrium, a slow growth tendency is shown, though not clearly. With the flow however, the mode is more stable than the no flow case. Therefore the flow has stabilizing effect on the mode, as predicted in other models previously[8-10].

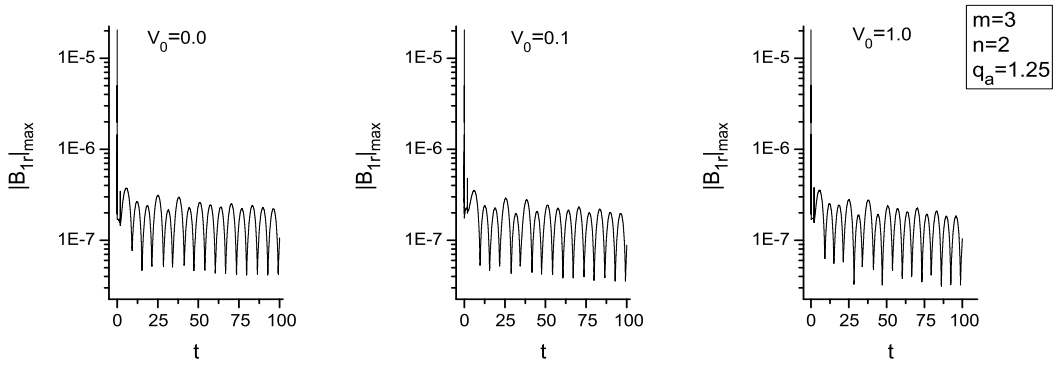
The $m/n = 3/2$ external mode is also tested and it is found that the mode is slowly damped, see Fig.3(b). Again, the flow has a stabilizing effect on the mode, even for a very fast Alfvénic flow.

Summary

In conclusion, we have studied the stability of resistive wall modes in cylindrical plasma confined by surface current. It is found that due to the strong stabilizing effect of the surface current distribution, both the internal and external MHD modes are stable. For the δ-function current distribution, the equilibrium flow has an additional stabilizing effect on the modes. And instead of the instability discussed in the incompressible plasma case[14], the flow has little effect on the modes.



(a)



(b)

Fig. 3. The evolution of the external modes with the different initial flow velocities of (a) $m/n=2/1$; (b) $m/n=3/2$

References

[1] R. Aymar, P. Barabaschi, Y. Shimomura: The ITER design. Plasma Phys. Control. Fusion. Vol. 44(5), p. 519-565 (2002)

[2] F. Troyon, R. Gruber, H. Saurenmann, S. Semenzato, S. Succi: MHD-Limits to Plasma Confinement. Plasma Phys. Control. Fusion. Vol. 26(1A), p. 209-215 (1984)

[3] D. Pfirsch, H. Tasso: A theorem on MHD-instability of plasmas with resistive walls. Nucl. Fusion. Vol. 11(3), p. 259-260 (1971)

[4] T. H. Jensen, M. S. Chu: A Linear Model for the Tearing Mode of a Tokamak Plasma with Flow and a Resistive Wall Boundary Condition. J. Plasma Phys. Vol. 30, p. 57-63 (1983)

[5] C. G. Gimblett: On Free Boundary Instabilities Induced By A Resistive Wall. Nucl. Fusion. Vol. 26, p. 617-625 (1986)

[6] J. P. Freidberg, *Ideal Magnetohydrodynamic,* New York: Plenum Press, (1987)

[7] L. E. Zakharov, S. V. Putvinskii: Effect of plasma rotation in a tokamak on the stabilizing effect of a conducting wall. Sov. J. Plasma Phys. Vol. 13, p. 118-119 (1987)

[8] A. Bondeson, D.J. Ward: Stabilization of External Modes in Tokamaks by Resistive Walls and Plasma Rotation. Phys. Rev. Lett. Vol. 72(17), p. 2709-2712 (1994)

[9] D. J. Ward, A. Bondeson: Stabilization of ideal modes by resistive walls in tokamaks with plasma rotation and its effect on the beta limit. Phys. Plasmas Vol. 2(1), p. 1570-1580 (1995)

[10] R. Betti, J. P. Freidberg: Stability Analysis of Resistive Wall Kink Modes in Rotating Plasma. Phys. Rev. Lett. Vol. 74(15), p. 2949-2952 (1995)

[11] J. A. Wesson: Magnetohydrodynamic flow instability in the presence of resistive wall. Phys. Plasmas Vol. 5(11), p. 3816-3819 (1998)

[12] B. M. Veeresha, S. N. Bhattacharyya, K. Avinash: Flow driven resistive wall instability. Phys. Plasmas Vol. 6(12), p. 4479-4486 (1999)

[13] S. Y. Cui, X. G. Wang, Y. Liu, and B. Yu: Numerical studies for the linear growth of resistive wall modes generated by plasma flows in a slab model. Phys. Plasmas. Vol. 13, p. 094506-094510 (2006)

[14] C. N. Lashmore-Davies: The resistive wall instability and critical flow velocity. Phys. Plasmas. Vol. 8(1), p. 151-157 (2001)

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.790

Use of Information Discrepancy Measure to Register Medical Images★

Shaoyan Sun[1,a], Lei Chen[2,b]

[1] School of Mathematics and Information, Ludong Universtiy, Yantai 264025, China

[2] Department of Information Engineering and Art and Design, Shandong Business Institute, Yantai 264025, China

[a]renasun@gmail.com, [b]softbunny@gmail.com

Keywords: Image registration, Kullback-Leilber divergence, Mutual information, Optimization

Abstract. Function of Degree of Disagreement (FDOD), a new measure of information discrepancy, quantifies the discrepancy of multiple sequences. This function has some peculiar mathematical properties, such as symmetry, boundedness and monotonicity. In this contribution, we first introduce the FDOD function to solve the three-dimensional (3-D) medical image registration problem. Numerical experiments illustrate that the new registration method based on the FDOD function can obtain subvoxel registration accuracy, and it is a competitive method with the mutual information based method.

Introduction

The geometric alignment or registration of medical images is a fundamental task in numerous applications in three-dimensional (3-D) medical image processing [1]. Recently there has been active research into the use of voxel-similarity-based (VSB) measures of multimodality medical image registration [2-6]. The main advantage of VSB methods is that feature calculation is straightforward or even absent when only grey-values are used, such that the accuracy of these methods is not limited by segmentation errors as in surface based methods. Among these methods, the mutual information (MI) based method is one of the most accurate methods, which measures the statistical dependence between images by measuring the discrepancy between the joint probability distribution and the product of the marginal distributions associated to the case of complete independence by means of the Kullback-Leibler divergence (KLD) measure, so we also denote this method as the KLD based method in the present paper.

This paper applies Function of Degree of Disagreement (FDOD), a new measure of information discrepancy, to three dimensional (3-D) medical image registration problem. The FDOD function has a close connection with the KLD measure, and in contrast to the KLD measure, it has some peculiar properties, such as symmetry, boundedness, monotonicity, effectiveness in singular case, convexity [7]. All of these facts motivate us to investigate whether this function can be successfully used in the field of medical image registration.

In the next section, we give a description of the FDOD function. And then introduce the use of this function to solve 3-D medical image registration problem in Section 3. Numerical experiments are demonstrated in Section 4. Finally, in Section 5, we present some concluding remarks.

FDOD Function

In recent years, a bounded FDOD function has been introduced in the study of information discrepancy among some multiple information sources [7]. In the case of measuring the discrepancy of a group of distributions, it is reduced to

★Project supported by the School Foundation of Ludong University (No.LY20082701).

[1] Corresponding Author:renasun@gmail.com

$$FDOD(U_1,U_2,\cdots,U_s)=\sum_{k=1}^{s}\sum_{i=1}^{t}p_{ik}\log\frac{p_{ik}}{\sum_{k=1}^{s}p_{ik}/s},U_k\in\Gamma_t \tag{1}$$

where $\Gamma_t=\{(p_1,p_2,\cdots,p_t)^T\mid\sum_{i=1}^{t}p_i=1,p_i\geq 0,i=1,2,\cdots,t\}$, $P=(p_1,p_2,\cdots,p_t)^T\in\Gamma_t$ denotes a probability distribution, and $0\cdot\log 0=0$, $0\cdot\log 0/0=0$ are defined.

To make it relevant to image registration context, in the following we only consider two probability distributions $P=(p_1,p_2,\cdots,p_t)^T$ and $Q=(q_1,q_2,\cdots,q_t)^T$. The discrepancy of distributions P and Q can be measured by the FDOD function as follows:

$$FDOD(P,Q)=\sum_{i=1}^{t}(p_i\log\frac{p_i}{(p_i+q_i)/2}+q_i\log\frac{q_i}{(p_i+q_i)/2}) \tag{2}$$

It can also be measured by the KLD measure:

$$KLD(P,Q)=\sum_{i}p_i\log\frac{p_i}{q_i} \tag{3}$$

Obviously, the FDOD function is symmetric about the two distributions P and Q, while the KLD measure is not. So the FDOD function can be viewed as a symmetric version of the KLD measure.

Medical Image Registration Based on FDOD Function

Considering image intensity values, a and b, of a pair of corresponding voxels in the two images that are to be registered to be random variables A and B, respectively, estimations for the joint distribution $p_{AB}(a,b)$ can be approximated by histogramming [2]. The joint distribution $p_{AB}(a,b)$ denotes that a voxel in image A has intensity a while the corresponding voxel in image B has intensity b. Their marginal distributions $p_A(a)$ and $p_B(b)$ can be respectively obtained as follows:

$$p_A(a)=\sum_{b}p_{AB}(a,b),\quad p_B(b)=\sum_{a}p_{AB}(a,b) \tag{4}$$

Variables A and B are statistically independent if $p_{AB}(a,b)=p_A(a)\cdot p_B(b)$, while they are maximally dependent if they are related by a one-to-one transformation T:

$$p_A(a)=p_B(T(a))=p_{AB}(a,T(a)). \tag{5}$$

If the discrepancy between the joint distribution $p_{AB}(a,b)$ and the distribution associated to the case of complete independence $p_A(a)\cdot p_B(b)$ are maximal, we say that two images of the same subject are geometrically aligned.

This discrepancy can be measured by the FDOD function and the KLD measure respectively, the latter case is often referred to as mutual information (MI):

$$FDOD(A,B)=\sum_{a,b}(p_{AB}(a,b)\log\frac{p_{AB}(a,b)}{(p_{AB}(a,b)+p_A(a)\cdot p_B(b))/2}+p_A(a)\cdot p_B(b)\log\frac{p_A(a)\cdot p_B(b)}{(p_{AB}(a,b)+p_A(a)\cdot p_B(b))/2}) \tag{6}$$

$$KLD(A,B)=I(A,B)=\sum_{a,b}p_{AB}(a,b)\log\frac{p_{AB}(a,b)}{p_A(a)\cdot p_B(b)} \tag{7}$$

In the present paper we have restricted the geometric transformation to rigid body transformation T_α (α consists of 3 translations and 3 rotations). The trilinear partial volume (PV) interpolation distribution was used to update the joint histogram. In all experiments, the joint

histogram size was 256 ×256. Similar to the KLD based algorithm, the registration method based on the FDOD function states that the images are geometrically aligned by the transformation T_{α^*} for which $FDOD(A,B)$ is maximal.

Powell's method was then used to maximize similarity functions. This method is a reasonable compromise between robustness and speed. Powell's method involves a series of one-dimensional minimizations for each dimensional. These minimizations are carried out by Brent's one-dimensional optimization. Having found an optimum in one direction, the minimization is continued in the next direction, starting from the current position. Once all six parameters (3 translations and 3 rotations) have been optimized, the loop is repeated until convergence is reached. The fractional precision convergence parameters for the Brent and Powell optimization algorithm were set to 10^{-3} and 10^{-5} respectively [8, 9].

Experimental Results

We implemented the FDOD function based method and the KLD measure based method on clinical data furnished by Vanderbilt University as the practice data set. For an in depth description of this database, please refer [10, 11]. In each registration task, CT or PET image is taken as floating image while MR image is always the reference one. The CT image was 2-voxel subsampled in x and y directions to accelerate the registration process and the initial registration parameter are set to zero vector $[0,0,0,0,0,0]^T$ for all the registration tasks. We compare our results with the stereotactic registration solutions provided by Vanderbilt University. The difference between the reference and each of the registration solutions computed by the FDOD function and the KLD measure was evaluated at eight points near the brain surface. The mean error of our results was computed as follows:

$$meanerror = \frac{1}{8}\sum_{i=1}^{8}\left\| q_{i,requence} - q_i \right\|_2,\ i = 1,2,\cdots,8 \tag{8}$$

Where $q_{i,requence}$ represents the i_{th} point's coordinate provided by Vanderbilt University, and q_i is the coordinate we computed. We define the diagonal distance in a voxel to indicate the size of a voxel. For CT to MR registration, the size of a voxel is the diagonal distance in a MR image's voxel, shown at below:

$$\sqrt{1.25^2 + 1.25^2 + 4.0^2} \approx 4.373(mm) \tag{9}$$

For PET to MR, the size of a voxel is the diagonal distance in a voxel belonging to PET image, depicted as follows:

$$\sqrt{2.590723^2 + 2.590723^2 + 8.0^2} \approx 8.799(mm) \tag{10}$$

If the mean error of a registration method is smaller than the size of a voxel, then we can say this method succeed and reached the subvoxel accuracy.

Although the sub-sampling of the floating images and not adopting the multi-resolution technique lead to lower success rate and lower registration accuracy in the task, we aim to investigate the feasibility of the FDOD function and the performance compared with that of the KLD measure under the same conditions.

Table 1. Comparison of the mean error (in mm)/the number of registration function evaluations computed by the FDOD function and the KLD measure respectively

Method	CT-T1	CT-T2	CT-PD	PET-T1	PET-T2	PET-PD
FDOD	2.6483/514	3.8646/732	2.5566/497	5.3667/914	8.6069/429	7.7652/664
KLD	1.2886/591	3.8446/644	2.4058/500	5.5356/833	9.3733/409	7.8821/655

Table 1 shows our computational results. From Table 1, we can see that the registration errors based on the FDOD function are all less than the size of a voxel while the KLD measure failed once (PET-T2), this verifies that the FDOD function possesses higher success rate than the KLD measure.

Conclusions

The FDOD function is a new measure of information discrepancy, and in contrast to the KLD measure it has more appealing mathematical characteristics. Motivated by these facts, we investigate the performance of this function used in medical image registration. Our experiments on multi-modality brain images show that the new method based on the FDOD function is a competitive registration method and possesses a little superiority to the KLD measure. However, many good characteristics of this function have not emerged when used in medical image registration problem. Our future work is to further explore these good characteristics to improve the performance of the FDOD function used in medical image registration.

References

[1] D.L.G. Hill, P.G. Batchelor, M. Holden and D.J. Hawkes: Medical image registration. Physics in Medicine and Biology. Vol. 46(3) (2001), p. 1-45

[2] F. Maes, A. Collignon , D. Vandermeulen, G. Marchal and P. Suetens, Multimodality image registration by maximization of mutual information. IEEE Transactions on Medical Imaging. Vol. 16(2) (1997), p. 187-198

[3] C. Studholme, D.L.G. Hill and D.J. Hawkes: An overlap invariant entropy measures of 3D medical image alignment. Pattern Recognition. Vol. 32(1) (1999), p. 71-86.

[4] X.S. Lu, S. Zhang, H. Su and Y.Z. Chen: Mutual information-based multimodal image registration using a novel joint histogram estimation. Computerized Medical Imaging and Graphics. Vol. 32 (3) (2008), p. 202-209

[5] H.X. Luan, F.H. Qi, Z. Xue, L.Y. Chen and D.G. Shen: Multimodality image registration by maximization of quantitative–qualitative measure of mutual information. Pattern Recognition. Vol. 41(1) (2008), p. 285-298

[6] D. Zikic, B. Glocker, O. Kutter, M. Groher, N. Komodakis, A. Kamen, N. Paragios and N. Navab: Linear intensity-based image registration by Markov random fields and discrete optimization. Medical Image Analysis. Vol. 14(4) (2010), p. 550-562

[7] W.W. Fang, F.S. Roberts and Z.R. Ma: A measure of discrepancy of multiple sequences. Information Sciences. Vol. 137(1-4), (2001)p. 75-102

[8] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vettering: Numerical Recipes in C. Cambridge: Cambridge University Press (1992)

[9] J.P.W. Pluim, J.B.A. Maintz, M.A. Viergever: Mutual information matching in multiresolution contexts. Image and Vision Computing. Vol. 19(1-2) (2001), 45-52

[10]Vanderbilt Univ., Nashville, TN. Retrospective Image Registration Evaluation (RIRE). Information on http://www.vuse.vanderbilt.edu/~image/registration/

[11] J. West, J.M. Fitzpatrick, Y.W. Matthew et al: Comparison and evaluation of retrospective intermodality brain image registration techniques. Journal of Computer Assisted Tomography. Vol. 21(4) (1997), p. 554-566

doi:10.4028/www.scientific.net/AMM.50-51.794

Filling Holes in Triangular Meshes with Feature Preserved

Ping Hu[1,2,a], Changsheng Wang[1,2,b], Baojun Li[1,2], Mingzeng Liu[1,2]

1. School of Automotive Engineering, Faculty of Vehicle Engineering and Mechanics,
2. State Key Laboratory of Structural Analysis for Industrial Equipment,

Dalian University of Technology, Dalian 116024, P.R. China

[a]pinghu@dlut.edu.cn, [b]wangcshe@gmail.com

Keywords: hole-filling, triangulation, refinement, bilateral filter

Abstract. In this paper, a novel hole-filling algorithm for triangular meshes is proposed. Firstly, the hole is triangulated into a set of new triangles using modified principle of minimum angle. Then the initial patching mesh is refined according to the density of vertices on boundary edges. Finally, the patching mesh is optimized via the bilateral filter to recover missed features. Experimental results demonstrate that the proposed algorithm fills complex holes robustly, and preserves geometric features to a certain extent as well.

1 Introduction

As 3D optical scanners become widespread, triangle meshes can be easier created and widely applied in the fields of CAD and reverse engineering. Even with high-fidelity scanners, the data obtained is often incomplete. The existence of holes makes it difficult for mesh operations. In the process of One-Step inverse forming finite element analysis, the existence of holes may affect the robustness of the analysis and generate uncertain results. An illustration is shown in Fig. 1. Fig. 1(a) shows the initial mesh of beetle, Fig. 1(b) the result of unfolding the beetle directly via KMAS/One-Step and Fig. 1(c) the result of unfolding after inner holes filled.

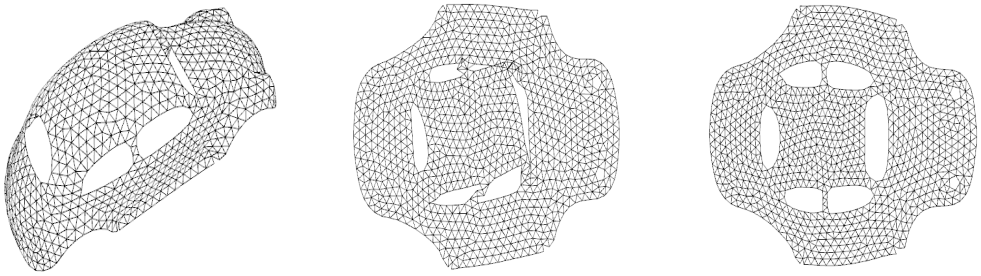


Fig. 1 (a) Initial mesh. (b) Unfold directly via KMAS/One-Step. (c) Unfold after inner holes filled

Surface feature refers to crest lines which are formed by points where the largest in absolute value principal curvature takes a positive maximum (negative minimum) along its corresponding curvature line [1]. It is a powerful shape descriptor and important for the appearance and expression of geometric models. Feature preserved hole-filling is significant in applications, which can effectively restore the initial data and improve the precision of finite element analysis. A feature preserved hole-filling algorithm is discussed in this paper.

Various mesh hole-filling approaches have been proposed in recent years. These approaches can be mainly classified into two categories [2]: volume-based and triangle-based. The volume-based approaches indirectly repair the model using an intermediate volumetric grid, while the triangle-based approaches identify and fill the holes directly on the model.

Volume-based approaches: The input model is first converted into an intermediate volumetric grid, where each grid point is associated with a positive or negative sign indicating it is inside or outside the model. Next, a polygonal surface is reconstructed that separates the grid points of different signs. Davis et al. [3] apply a volumetric diffusion process to extend a signed distance function through this volumetric representation until its zero set bridges whatever holes may be present. Nooruddin et

al. [4] develop a similar approach for the simplification and the repairing of polygonal meshes. Ju [5] traces and patches hole boundaries on a dual surface of the primal grid, which yields an inside/outside partitioning on the primal mesh. Volume-based methods excel in their robustness in resolving complex holes, however, they are time consuming and the generated mesh may be incorrect in some cases.

Mesh-based approaches: The holes are explicitly searched and filled directly on triangle mesh. Liepa [6] first fills the hole with a minimum area triangulation of its 3D contour [7], then the triangulation is refined so that the triangle density agrees with the density of the surrounding mesh triangles [8], at last a fairing technique based on an umbrella operator [9] is used to smooth the filled hole. Brunton et al. [10] first unfold hole boundary onto a plane and triangulate it, then embed the triangular mesh as a minimum energy surface in R^3. Zhao et al. [11] optimize the vertex position by solving Poisson equation after getting the initial patch mesh by advancing front method. Chen et al. [12] first get an initial repaired model by interpolation-based process then a sharpness-dependent filtering algorithm is applied to the initial triangulation. A piecewise hole-filling algorithm is proposed by Jun [13], it splits the complex hole into several simple holes and then fills them with planar triangulation method. Li et al. [14] adopt the same strategy based the concept of edge expansion. Li et al. [15] have applied the polynomial blending in hole-filling, which can preserve the detailed features of the original mesh to a certain extent.

In this paper, a new approach of hole-filling is proposed which can preserve geometric features to a certain extent. We first triangulate the hole using modified principle of minimum angle, then the triangulation is refined according to the density of vertices on boundary edges, finally the refined mesh is optimized via the bilateral filter [16, 17] to recover the missed features.

The rest of the paper is organized as follows. In section 2, the detail of algorithm is presented. Experimentations and results are discussed in section 3. The paper is concluded in section 4.

2 Hole-filling algorithm

A triangular mesh is defined as a set of vertices and a set of oriented triangles that join these vertices. If two triangles share a common edge, they are adjacent triangles. An edge usually links two triangles. A boundary edge is an edge that only connects one triangle and a boundary vertex is a vertex on boundary edge. A hole is a closed loop of boundary edges.

All mesh models in this paper are assumed to be oriented, manifold and connected, and a given hole is assumed to have no islands. In the hole-filling process, we refer to the original mesh as the surrounding mesh and the mesh that fills the holes as the patching mesh. The main steps of the method are hole triangulation, mesh refinement and mesh optimization.

Triangulation

After hole identification, the first step in filling it is to triangulate it with modified principle of minimum angle. Denote LV as the collection of ordered vertices on the hole boundary, NT the collection of new triangles to be added to the hole and sum(LV) the number of vertices in LV. Given an angle threshold σ, main steps of triangulation are as follows:

Step 1: If sum(LV) > 3, go to Step 2, otherwise go to Step 3.

Step 2: Search for the vertex $\boldsymbol{v}$ with minimum angle in LV, denote α as the angle constituting of two normals of its adjacent vertices in LV. If $\alpha > \sigma$, search for the vertex with minimum angle except $\boldsymbol{v}$ and calculate the angle α until $\alpha \le \sigma$. Denote the vertex as $\boldsymbol{v}$ and its two adjacent vertices in LV as $\boldsymbol{v}_1$,$\boldsymbol{v}_2$. Then add the oriented triangle ($\boldsymbol{v}_1$, $\boldsymbol{v}$, $\boldsymbol{v}_2$) to NT and delete vertex $\boldsymbol{v}$ from LV.

Step 3: Add the oriented triangle to NT.

The threshold σ is associated with the maximum angle of two normals constituting of two arbitrary vertices on the hole boundary.

Refinement

In order to fill a hole with a mesh that approximates the density of the surrounding mesh, we refine the patching mesh by the algorithm given by Pfeifle and Seidel [8]. Given a density control factor α and the patching mesh, main steps of refinement are as follows:

Step 1: For each vertex $\boldsymbol{v}_i$ on the hole boundary, compute the scale attribute $\sigma(\boldsymbol{v}_i)$ as the average length of the edges that are adjacent to $\boldsymbol{v}_i$ in the surrounding mesh. Initialize the patching mesh as the given hole triangulation.

Step 2: For each triangle $(\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k)$ in the patching mesh, compute the centroid $\boldsymbol{v}_c$ and the corresponding scale attribute $\sigma(\boldsymbol{v}_c) = (\sigma(\boldsymbol{v}_i) + \sigma(\boldsymbol{v}_j) + \sigma(\boldsymbol{v}_k)) / 3$. For $m=i, j, k$, if $\alpha\|\boldsymbol{v}_c - \boldsymbol{v}_m\| > \sigma(\boldsymbol{v}_c)$, $\alpha\|\boldsymbol{v}_c - \boldsymbol{v}_m\| > \sigma(\boldsymbol{v}_m)$ and the maximum inner angle of triangle $(\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k)$ is less than $5\pi/6$, then replace the triangle $(\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k)$ with three triangles: $(\boldsymbol{v}_c, \boldsymbol{v}_j, \boldsymbol{v}_k)$, $(\boldsymbol{v}_i, \boldsymbol{v}_c, \boldsymbol{v}_k)$ and $(\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_c)$ in the patching mesh, and then relax the edges $(\boldsymbol{v}_i, \boldsymbol{v}_j)$, $(\boldsymbol{v}_i, \boldsymbol{v}_k)$ and $(\boldsymbol{v}_j, \boldsymbol{v}_k)$.

Step 3: If no new triangles were created in Step 2, the patching mesh is complete.

Step 4: Relax all interior edges of the patching mesh.

Step 5: If no edges were swapped in Step 4, go to Step 2, otherwise go to Step 4.

Generally, density control factor $\alpha=2^{1/2}$ is a good choice. To relax an edge means, for the two triangles adjacent to the edge, check whether each of the two non-mutual vertices of these triangles lies outside of the circum-sphere of the opposing triangle (see Fig. 2). If this test fails, the edge is swapped.

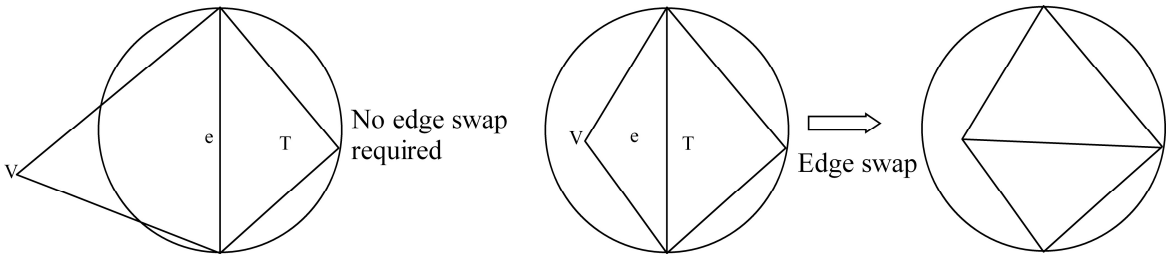


Fig. 2 Edge relaxation

Optimization

As noted in the introduction, the mesh optimization process for the refined patching mesh can be regarded as bilateral filter [16, 17]. The bilateral filter shifting mesh vertices along their normal direction can preserve sharp features. The pseudo-code for applying a bilateral filter to a single vertex $\boldsymbol{v}$ is given in the following:

$\{\boldsymbol{q}_i\}$=neighborhood($\boldsymbol{v}$), K=$|\{\boldsymbol{q}_i\}|$, sum=0, normalize=0.
for i:=1,⋯, K
  t=$\|\boldsymbol{v}-\boldsymbol{q}_i\|$, h=<$\boldsymbol{n}$, $\boldsymbol{v}-\boldsymbol{q}_i$>, $w_c=\exp(-t^2/(2\sigma_c^2))$, $w_s=\exp(-h^2/(2\sigma_s^2))$
  sum += $(w_c \cdot w_s)\cdot h$
  normalize += $w_c \cdot w_s$
end
return Vertex $\boldsymbol{v}' = \boldsymbol{v}+\boldsymbol{n}\cdot$(sum/normalize).

Where $\{\boldsymbol{q}_i\}$ is the neighborhood of $\boldsymbol{v}$, $\boldsymbol{n}$ the normal of $\boldsymbol{v}$ and K the number of $\{\boldsymbol{q}_i\}$.

3 Implementations and results

This section illustrates how our algorithm is implemented on computer and demonstrates the experiment results. The algorithm is implemented in VC++6.0 and OpenGL. In the program, let σ_c and σ_s be 0.1 and 0.5 of the current vertex density respectively, $\{\boldsymbol{q}_i\}$ 1-ring neighbors of $\boldsymbol{v}$ and $\sigma = 45°$ when the maximum angle is less than 55°, else σ be 85% of the maximum angle. The number of iterations of bilateral filter is three.

The first example concerns the mesh model of component in vehicle notes as model a (Fig. 3(a)). To illustrate the power of feature preservation, a hole has been created in a blending area (Fig. 3(b)). The comparison between the method of Liepa [6] (Fig. 3(c)) and our method illustrates our algorithm (Fig. 3(d)) can preserves geometric features to a certain extent.

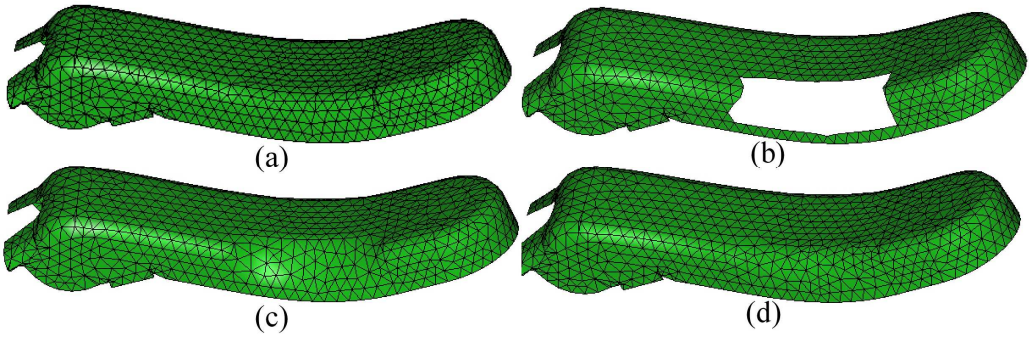


Fig. 3 Patch results of model a. (a) The initial mesh model. (b) A hole is created in a blending area. (c) The result of Liepa [6]. (d) The result of our method

Another two examples are given below. Fig. 4 shows the patch results of model b and c. Quality statistics of patching meshes of the models are given in Fig. 5. The horizontal axis denotes the quality of triangles, while the vertical axis denotes the number of triangles. We adopt the method of Field [18] to assess the quality of triangle by the ratio of twice the radius of inscribed circle to the radius of circum circle of a triangle. Generally, the triangle is of good quality if the ratio is not less than 0.7, and the closer to 1, the better of the triangle. Fig. 5 illustrates the patching meshes are of good quality, which can well approximate the density of surrounding mesh and preserve geometric features as well.

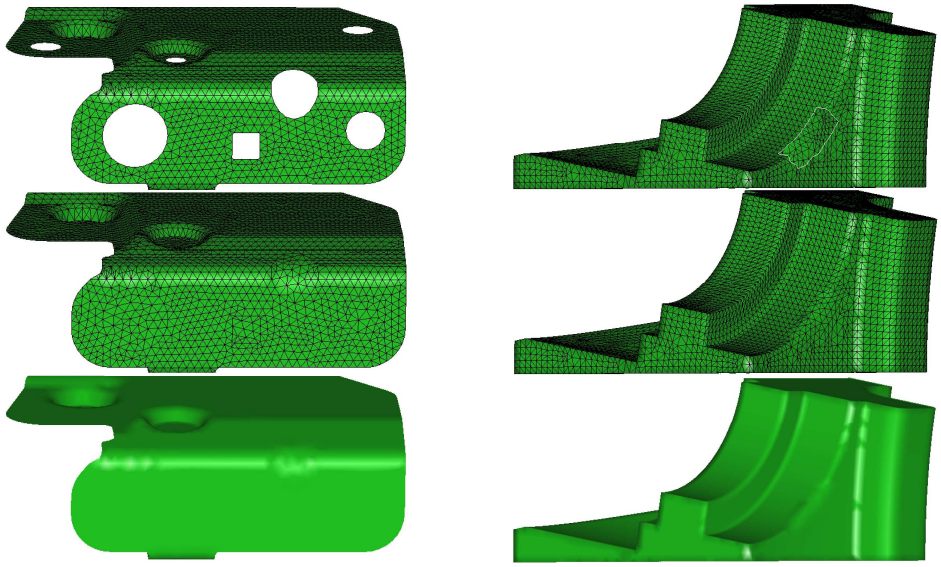


Fig. 4 Patch results of model b and c

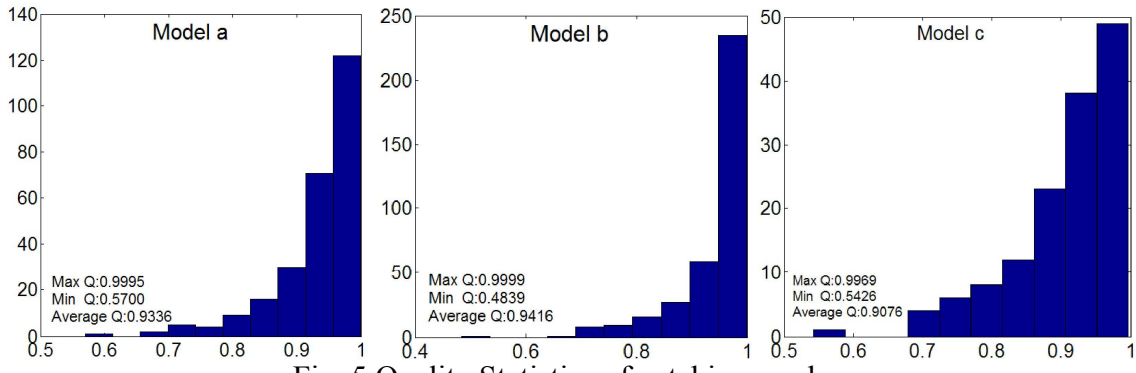


Fig. 5 Quality Statistics of patching meshes

4 Conclusions

In this paper, a novel hole-filling algorithm is proposed. The method first triangulates the hole into a set of new triangles, then refines the patching mesh according to the density of vertices and finally

optimizes the patching mesh via the bilateral filter. The proposed method is simple, stable and robust. Modified principal of minimum angle and bilateral filter is crucial for feature preservation. Experiments illustrate the method can preserve features to a certain extent.

Acknowledgments

We would like to thank Hui Wang, Yongxia Hao, Xiaochao Wang and Yan Zhao for their help. This work was funded by the Key Project of the NSFC (No. 10932003, u0935004), NSFC（No. 60873181）, “863” Project of China (No. 2009AA04Z101), “973” National Basic Research Project of China (No. 2010CB832700) and the Fundamental Research Funds for the Central Universities.

References

[1] Y. Ohtake, A. Belyaev and H.-P. Seidel. Ridge-valley lines on meshes via implicit surface fitting. *ACM Transactions on Graphics*, 2004, 23(3): 609-612.

[2] T. Ju. Fixing geometric errors on polygonal models: a survey. *Journal of Computer Science and Technology*, 2009, 24(1): 19-29.

[3] J. Davis, S.R. Marschner, M. Garr and M. Levoy. Filling holes in complex surfaces using volumetric diffusion. In: *First International Symposium on 3D Data Processing Visualization and Transmission*, Padova, 2002, 428-438.

[4] F.S. Nooruddin and G. Turk. Simplification and repair of polygonal models using volumetric techniques. *IEEE Transactions on Visualization and Computer Graphics*, 2003, 9(2): 191-205.

[5] T. Ju. Robust repair of polygonal models. *ACM Transactions on Graphics*, 2004, 23(3): 888-895.

[6] P. Liepa. Filling holes in meshes. In: *SGP '03: Proceedings of the 2003 Eurographics/ACM SIGGRAPH symposium on Geometry processing*. Aire-la-Ville, Switzerland: Eurographics Association, 2003, 200-205.

[7] G. Barequet and M. Sharir. Filling gaps in the boundary of a polyhedron. *Computer Aided Geometric Design*, 1995, 12(2): 207-229.

[8] R. Pfeifle and H.-P. Seidel. Triangular B-splines for blending and filling of polygonal holes. In: *Proceedings of Graphics Interface*, 1996, 186-193.

[9] L. Kobbelt, S. Campagna, J. Vorsatz and H.-P. Seidel. Interactive multi-resolution modeling on arbitrary meshes. In: *Proceedings of SIGGRAPH’98*, 1998, 105-114.

[10] A. Brunton, S. Wuhrer, C. Shu, P. Bose and E.D. Demaine. Filling holes in triangular meshes by curve unfolding. In: *IEEE International Conference on Shape Modeling and Applications*, 2009, 66-72.

[11] W. Zhao, S.M. Gao and H.W. Lin. A robust hole-filling algorithm for triangular mesh. *The Visual Computer: International Journal of Computer Graphics*, 2007, 23(12): 987-997.

[12] C.Y. Chen and K.Y. Cheng. A sharpness-dependent filter for recovering sharp features in repaired 3D mesh models. *IEEE Transactions on visualization and computer graphics*, 2008, 14(1): 200-212.

[13] Y. Jun. A piecewise hole filling algorithm in reverse engineering. *Computer-Aided Design*, 2005, 37(2): 263-270.

[14] G. Li, X.Z. Ye and S.Y. Zhang. An algorithm for filling complex holes in reverse engineering. *Engineering with Computers*, 2008, 24(2): 119-125.

[15] Z. Li, D.S. Meek and D.J. Walton. Polynomial blending in a mesh hole-filling application. *Computer-Aided Design*, 2010, 42: 340-349.

[16] S. Fleishman, I. Drori and D. Cohen-Or. Bilateral mesh denoising. *ACM Transactions on Graphics*, 2003, 22(3): 950-953.

[17] T.R. Jones, F. Durand and M. Desbrun. Non-iterative, feature-preserving mesh smoothing. *ACM Transactions on Graphics*, 2003, 943-949.

[18] D.A. Field. Qualitative measures for initial meshes. *International Journal for Numerical Methods in Engineering*, 2000, 47: 887-906.

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.799

Research of Centroid Extraction for Feature Point Based on FPGA

Bing Li[1,a], Xiao-hong Liu[1,b], Wen-jing Wu[2,c]

[1]Science and Technology on Electro-optic Control Laboratory Luoyang China 471009

[2]Luoyang Institute of Electro-Optical Equipment of AVIC Luoyang China 471009

[a]jluleeice@foxmail.com, [b]moongirl1110@yahoo.cn, [c]goodlucktoice@163.com

Keywords: Centroid extraction, FPGA, Morphological filter, Component labeling

Abstract. This paper focuses on the centroid extraction algorithm of feature point. We present a recognition algorithm to identify the feature point and extract centroid. This algorithm can extract the centroid of the feature point from the complex background by scanning the original image only one time. We design a hardware architecture and implement it based on FPGA. Experimental results show that it can extract the centroid coordinates exactly from the complex background in real time with low-cost hardware resources.

I. Introduction

The applications of machine vision systems contain identify objects and report their position, read characters or measure statistics and so on [1]. Feature extraction algorithms are essential to the implement of a machine vision system. A method for finding the pose of an object from a image is by extracting feature points centroid of the object in motion.[4] Centroid extraction processes are typically divided onto four steps, Capturing the original gray level feature points image by the image sensor, pre-processing the original image by a specified filter, converting the pre-processed image to a binary image and separating the foreground from background, labeling the connected components and extracting the centroid coordinates of the feature points.

Many feature extraction algorithms have been proposed to identify objects with a specific tag. Benny.T et al [1] present an image component labeler and feature extraction module based on FPGA. It can extract complex connected components in real-time. Dong Kyun Kim et al [2] design a hardware architecture that combines the labeling and boundary tracing procedures. It can extract many kinds of shapes by scanning the gray image twice. This algorithm is suitable for the objects we haven't specified. Boon Hean Pui et al [12] design a hardware system to compute in real-time the centroid of a scanning light beam. Jie Jiang et al [5] present a rapid star tracking algorithm which can identify and extract the centroid by matching the star map. For the feature point extraction which the shape is specified and the background is complex, those algorithms above are not suitable. A new algorithm need to present that can extract feature point from the complex background and separating from all connected components accurately.

To solve this problem we present in this paper a hardware image processing architecture for feature point centroid extraction which with the function of anti-jamming by a recognition algorithm in real-time. This hardware architecture can pre-process the original image and filter most of complex background interference, then label the connected components and recognize which connected components are belong to the feature points and extracting centroid.

This paper is organized as follows: in Section II we introduce an image pre-processing method. In Section III we provide a detail description of the feature point centroid extraction algorithm. The hardware architecture design is given in Section IV. In Section V we show experimental results and present the conclusion in Section VI.

II. Method for Image Pre-Processing

The Widely used image pre-processing approaches fall into two broad categories: spatial domain methods and frequency domain methods. The term spatial domain refers to the image plane it self, and approaches in this category are based on direct manipulation of pixels in an image.

Morphological filter is a kind of spatial domain methods which has been widely applied in the field of object identifying. Mathematical morphology includes erosion, dilation, opening operation and closing operation. In order to enhance the feature points in the original image, we first adopt the opening operation to the original image, then subtract the image which has been performing opening operation from the original image The model is defined as follows[8]:

$$\text{Filter}=f-(f\circ se) \tag{1}$$

Where

f	is the original image.
se	is the structuring element in morphology.
Filter	is the filtered image.

We set se as a flat, linear structuring element and the angle is 90° .The opening operation consists of two operations, eroding operation and dilating operation, Eroding operation is used to find the minimum value replace the current value and dilating operation is to find the maximum value replace the current value in the specified structuring element. The small targets are filtered and the background is retained after opening operation. Subtracting the open image form the original image can obtain the feature points without complex background interference. In fact these operates above are called Opening Top-hat (OTH).

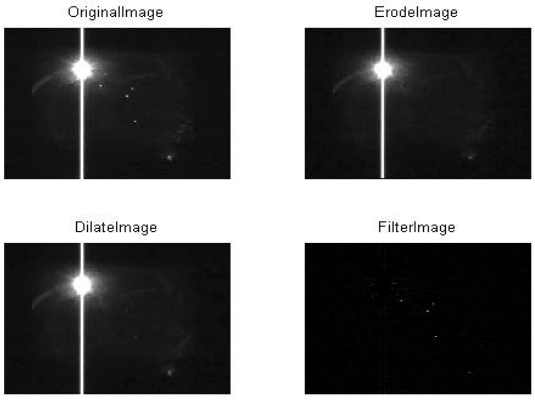


Figure.1 Result of applying the morphological filter given in Eq(1) .

As an illustration of morphological filter, Fig.1 shows the effect of image pre-processing by opening top-hat filter. The implement of morphology filter is suitable for pipeline structure based on FPGA. So we choose the morphological filter for image pre-processing in this paper.

III. Algorithms for recognizing and extracting

The pre-processed image can be converted to a binary image by a simply logic operation,

$$F(x,y)=\begin{cases} f(x,y) & f(x,y)>\text{Th} \\ 0 & f(x,y)\le \text{Th} \end{cases} \tag{2}$$

Where

$f(x,y)$	is the pre-processed gray image.
Th	is the threshold of global background image.
F(x,y)	is the binary image.

The next process is to label the connect component in binary image and calculate centroid coordinates.

Two broad categories of centroid extraction algorithms are described in [9, 10], one is based on target gray level. the other is based on the target contour. Studies have shown that the algorithm based on target gray level is more suitable for extracting small target centroid with symmetrical shape [11].

We define the connect component of the small target is expressed in terms of x and y coordinates, the centroid extraction algorithm is expressed as Eq 3,

$$Xc = \frac{\sum_{x=1}^{m}\sum_{y=1}^{n} F(x,y)\bullet x}{\sum_{x=1}^{m}\sum_{y=1}^{n} F(x,y)} \quad Yc = \frac{\sum_{x=1}^{m}\sum_{y=1}^{n} F(x,y)\bullet y}{\sum_{x=1}^{m}\sum_{y=1}^{n} F(x,y)} \tag{3}$$

Where

x, y are the pixels coordinates in a connected component.

X_c, Y_c are the centroid coordinates of the connected component.

The shape of the feature point is simply a dots so it is easy to label the connect component by scanning the binary image left to right, top to bottom. The labeling algorithm is given as follows,

1. Scan the current line until find the first pixel in the foreground, record its coordinate (X_s, Y_s) as the first coordinate data in a new region, continue to scan until the next pixel is belong to the background, record the current coordinate (X_e, Y_e) .Calculate and save statistical data of connected component for each line, total line connected component pixels number $\sum P$,the sum of gray level value $\sum G$,the sum of the coordinates multiplied by the gray value $\sum_{i=s}^{e}(X_i \bullet G_i)$ and $\sum_{i=s}^{e}(Y_i \bullet G_i)$,because the pixels in one line have the same Y coordinate, so $\sum_{i=s}^{e}(Y_i \bullet G_i)$ can be written as $Y_{No.} \bullet \left(\sum_{i=s}^{e} G_i\right)$,where $Y_{No.}$ is the current line number. Save the statistical data above and continue to scan the next line.

2. If the current line number $Y_{No.}$ and previous line number Y_{preNo} are continuous, this current line may combine to the previous record lines. We define a parameter $\max_{sub}$, set first foreground pixel coordinate in previous line is X_{preS} , first foreground pixel coordinate in current line is X_{CNo} , then

If $|X_{CNo} - X_{preNo}| \le \max_{sub}$, $|X_{CNo} - X_{preNo}|$ consider the current line and the previous are all belong to the same connected component region.

Else generate a connected component with the last line number Y_{preNo} , X_{CNo} is belong to a new connected component.

Several connected components may found when the scanning operation is over. The further recognition methods are provided to check which connect component need to calculate.

We propose two methods to recognize the connect component in binary image. One method is based on the area of the connected component. The other is based on the shape of the connected component

Set the maximum value of area is $Pixel_{MaxArea}$, the minimum value of area is $Pixel_{MinArea}$, $Pixel_{region}$ is the area value we calculated, the method is described as follows:

If $Pixel_{MinArea} \le Pixel_{region} \le Pixel_{MaxArea}$, consider this connected component meet standards, save and continue for the further recognition.

Else, consider this connected component is too large or small to meet standards, discard it.

We define two parameters aspect ratio and compact ratio to determine whether the connected component shape meet the standard. The definitions of two parameters are shown in Fig.2.

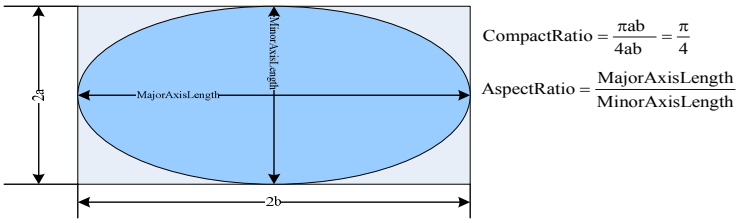


Figure.2 Definitions of the compact ratio and aspect ratio

The ideal aspect ratio is 1 and the ideal compact ratio is $\frac{\pi}{4}$ of a circle, Set the maximum aspect ratio is $\text{Aspect}_{\text{DotMax}}$, the minimum aspect ratio is $\text{Aspect}_{\text{DotMin}}$, the maximum compact is $\text{Compact}_{\text{DotMax}}$, the minimum compact ratio is $\text{Compact}_{\text{DotMin}}$. The calculated results of the aspect ratio and the compact ratio are $\text{Aspect}_{\text{region}}$ and $\text{Compact}_{\text{region}}$, the recognition methods are shown as follows,

If $\text{Aspect}_{\text{DotMin}} \le \text{Aspect}_{\text{region}} \le \text{Aspect}_{\text{DotMax}}$, consider this region meet standards, save and for the further recognition.

Else, consider this connected component does not belong to the feature point, discard it.

If $\text{Compact}_{\text{DotMin}} \le \text{Compact}_{\text{region}} \le \text{Compact}_{\text{DotMax}}$, consider this region meet standards, save and calculate the centroid.

Else, consider this connected component does not belong to the feature point, discard it.

IV. Hardware Architecture Design

The detail descriptions of extracting and recognition methods are presented in section V. We design a hardware architecture that can obtain successive original image data, pre-process the original image data and extract the centroid coordinates of feather point. Fig.3 shows the block diagram of this hardware architecture, this system comprises three main units. These units in order are as follows: Image pre-processing unit, image segmentation unit and centroid extraction unit.

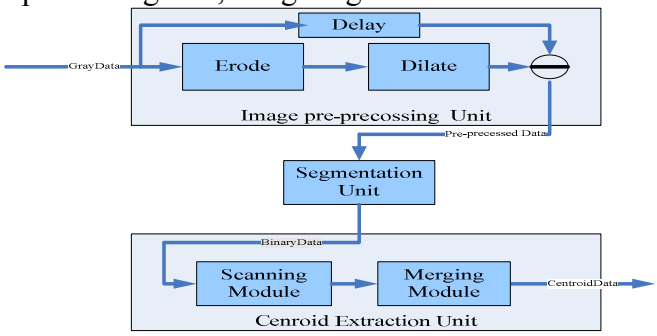


Figure.3 Hardware architecture of centroid extraction

A. Image Pre-processing Unit. The image pre-processing unit receives successive original gray image data from the image sensor. We design a 3-comparator for eroding operation. This comparator is capable of outputting minimum value of three consecutive data. The structure is shown in Fig.4.

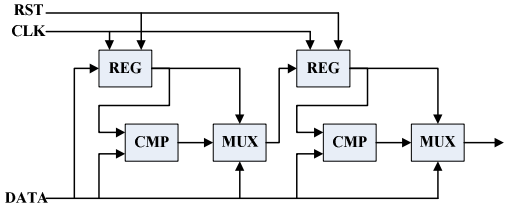


Figure.4 Structure of 3-comparator

Where

RST is the reset signal

CLK is the pixel clock

DATA is the gray image data

If the length of the selected se set to 1×5, for example, it is available by using two 3-comparators:

$$\min(a_1,a_2,a_3,a_4,a_5)=\min(\min(a_1,a_2,a_3),\min(a_2,a_3,a_4),\min(a_3,a_4,a_5))$$

The structure of dilating operation is similar to the eroding operation. The difference is that dilating operation for finding the max value in the structuring element

The relationship between the number of comparator and the length of the se is

$$CMP_{NO}=\frac{(Length_{se}-1)}{2} \tag{4}$$

Where

CMP_{NO} is the number of the comparators need to be generated.

$Length_{se}$ is the length of the se we specified.

It is convenient to automatically generate a specified length of comparators by modifying parameter $Length_{se}$.

B. Segmentation unit. The pre-processed image has been greatly enhanced so we chose a fixed value as the threshold to separate the foreground from background.

C. Centroid Extraction Unit. To implement the centroid extraction algorithm based on FPGA, we present the centroid extraction unit as follows: this unit is consisted of two modules. One is the scanning module. the other is the merging module.

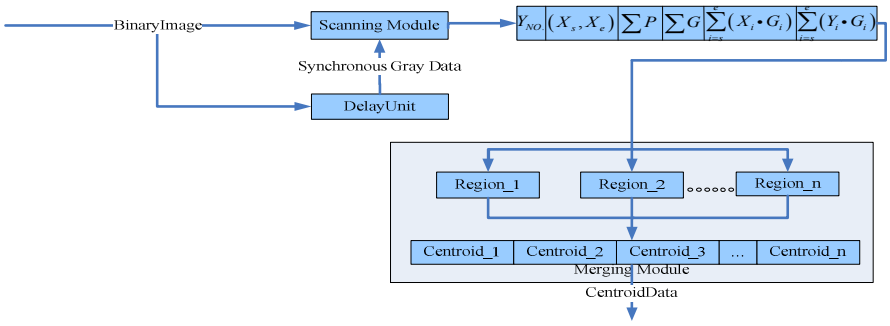


Figure.5 The structure of centroid extraction unit

The scanning module scan the connected component line by line, calculates and send $Y_{NO.}$, (X_s,X_e), $\sum P$, $\sum G$, $\sum_{i=s}^{e}(X_i\bullet G_i)$, $\sum_{i=s}^{e}(Y_i\bullet G_i)$ to the merging module.

For the merging module we set a data structure. This data structure is consisted of 11 elements:

Where

ValidFlag is the flag that the line is available to be merged.

NextLineNo is the next line number, which used to determine whether this connected component is scanned over.

FinalLineXs, FinalLineXe are the starting coordinate and the end coordinate of the connected component coordinates in the final line, which used to determine whether the new connected component line belong to the current scanned connected component.

The definitions of UpMaxY, DownMaxY, LMaxX, RMaxX are shown in Fig 6. These parameters are used to determine whether the shape of connected component meet the standard.

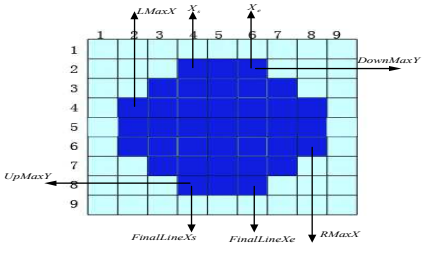


Figure.6 Parameters of the feature point

PixelCnt is the area of the connected component, which used to determine whether the calculated area meets the standard.

SumGray , SumXGray , SumYGray are the values used to calculate the centroid coordinates of feature point

Where

SumGray is the sum of gray of the connected component.

SumXGray is the sum of multiplying x and its corresponding gray value.

SumYGray is the sum of multiplying y and its corresponding gray value.

V. Experimental Results

An experiment based on the proposed hardware architecture is conducted. Fig.7 is an 800×600 experimental image which contains three feature points and two interference targets. The parameters are set as follows:

$se = 29$, $\max_{sub} \le 5$, $12 \le Pixel_{region} \le 200$, $0.5 \le Aspect_{region} \le 2$, $Compact_{region} \le 2$,

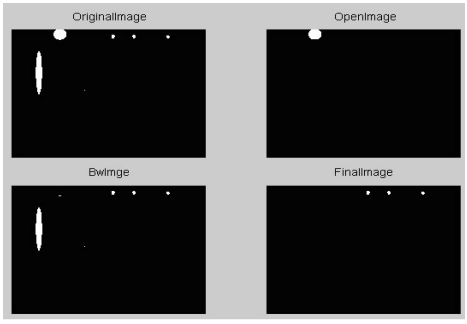


Figure.7 Effect of feature extraction

The results of centroid extracted by Matlab and Modelsim is shown in Fig.8

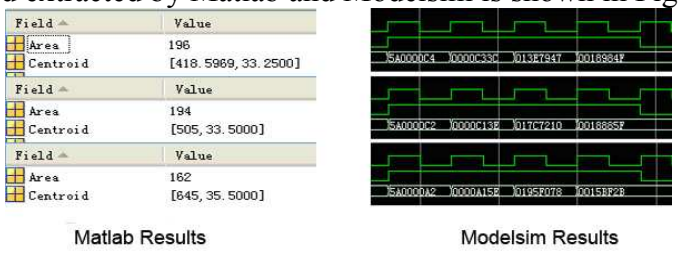


Figure.8 Results of the centroid extraction

Modelsim results are shown in Table 1.

Table 1. Modesim Results

Region_No.	ΣP	ΣG	Σ(XG)	Σ(YG)	X	Y
1	C4=196	C33C=49980	13E7947=20871495	18984F=1611855	417.5969	32.25
2	C2=194	C13E=49470	17C7210=24932880	18885F=1607775	504	32.5
3	A2=162	A15E=41310	195Ff078=26603640	15BF2B=1425195	644	34.5

Notice that the centroid extract between Modelsim and Matlab have a difference of 1. That because Matlab calculates the coordinates start at 1 and Modelsim start at 0, take x coordinate as an example:

$$X_M - X_F = \frac{\sum_{x=i+1}^{m+1} xf(x,y) - \sum_{x=i}^{m} xf(x,y)}{\sum_{x=i}^{m} f(x,y)} = \frac{\sum_{x=i}^{m} f(x,y)}{\sum_{x=i}^{m} f(x,y)} = 1 \quad (5)$$

Where

X_M is the Matlab result.

X_F is the result based on FPGA.

We can add 1 to correct the calculated centroid in FPGA. The hardware architecture is

implemented in a virtex-5 LX20TFF323 FPGA in the experiment Table 2 shows the design summary of the implementation architecture.

Table 2. Device Utilization for 5VLX20FF323

	Used	Available	Utilization
IOS	44	172	25.58%
Global Buffers	3	32	9.38%
Function Generators	5432	12480	43.53%
CLB Slices	1358	3120	43.53%
Dffs or Latches	5188	12480	41.57%
Block RAMs	0	26	0.00%
DSP48Es	14	24	58.33%

VI. Conclusion

We present a hardware architecture for feature point centroid extraction. This architecture comprises three main components: the image pre-processing unit, the image segmenting unit and the centroid extracting unit. The proposed algorithm has been designed by using Verilog and implemented in an FPGA. Experimental result shows that this hardware architecture can extract centroid with anti-jamming ability as a low-cost and high-speed application.

References

[1] Benny Thörnberg and Najeem Lawal. Real-time Component Labeling and Feature Extraction on FPGA, IEEE, 2009

[2] Dong Kyun Kim, Dae Ro Lee and Thien Cong Pham. Real-time Component Labeling and Boundary Tracing System Based on FPGA. Proceedings of the IEEE International Conference on Robotics and Biomimetics, 2007

[3] Holger Flatt, Steffen Blume and Sebastian Hesselbarth. A Parallel Hardware Architecture for Connected Component Labeling Based on Fast Lable Merging, IEEE, 2008, 144-149

[4] Daniel F. DeMenthon and Larry S.Davis. Model-Based Object Pose in 25 Lines of Code

[5] Jie Jiang, Guangjun Zhang and Xinguo Wei. Rapid Star Tracking Algorithm for Star Sensor, IEEE A&E Systems Magazine, 2009, 23-33

[6] P.J. Pietraski, ***Z.*** Zojceski, D.P. Siddons, G.C. Smith and B. Yu.Digital. Centroid-Finding Electronics for High-Rate Detectors, IEEE Transactions on Nuclear Science, 1999 Vol. 46, No.4

[7] Keqi Zheng, Shuching Chen. A Progressive Morphological Filter for Removing Nongrund Measurements From Airborne LIDAR Data. IEEE Transactions on Geoscience and Remote Sensing, 2003, 872-882.

[8] YE Bin, PENG Jia-xiong. Small Target Detection Method Based on Morphology Top-Hat Operator. Joural of Image and Graphics, 2007, 15-19

[9] P.J. Pietraski, Z. Zojceski, D.P. Siddons, G.C. Smith and B. Yu. Digital Centroid-Finding Electronics for High-Rate Detectors. IEEE Transactions on Nuclear Science, 1999,Vol.46, 810-816

[10] Takashi Komuro, Idaku Ishii, Masatoshi Ishikawa and Astrshi Yoshida. A Digital Vision Chip Specialized for High-Speed Target Tracking. IEEE Transactions on Electron Devices, Vol.50, 2003, 191-199

[11] Sharon S.Welch. Effects of Window Size and Shape on Accuracy of Subpixel Centroid Estimation of Target Images. NASA Technical Paper 3331, 1993

[12] Boon Hean Pui, Barrie Aayes-gill, Matt Clark and Mike Somekh The Design and Characterisation of an Optical VLSI Processor for Real Time Centroid Detection. Analog Integrated Circuits and Signal Processing, 2002, 32, 67-75

doi:10.4028/www.scientific.net/AMM.50-51.806

Design of Free-Form Surface Optic in Wide Field of View Off-axis Optical System

You Long[1, a], Ming-wu Kang[1,b],Xiao-xia Li [2,c]

[1]Science and Technology on Electro-optic Control Laboratory

[2]Luoyang Institute of Electro-Optical Equipment of AVIC Luoyang China 471009

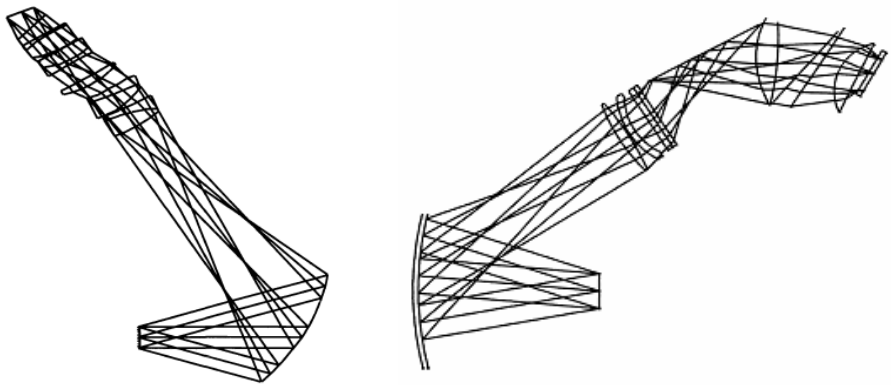
[a]longyou1122@gmail.com, [b]kang_m_w@163.com, [c]xlx9966@yahoo.com.cn

Keywords: Optical design, Free-form surface, Off-axis, Optical system

Abstract. Design of off-axis system is a hot and difficult topic in optical design area. The wide field of view(WFOV) off-axis system this paper involved always have the drawbacks of complicated large size and great weight. This paper introduces an application of free-form surface(FFS) in WFOV off-axis optical system design, analyzes FFS optics' material and fabrication methods. The system uses one FFS lens and one doublet, image quality meets the design specification, and characterizes by simple structure and lightweight.

1 Introduction

Off-axis optical system points that all optical units have no symmetry axis, these elements have tilts and decenters. Off-axis optical system design is more complex than symmetry ones. Fig.1 depicts the typical off-axis optical systems, they both have a tilted reflect surface and produce off-axis aberrations, so the relay optical system comprise lots of optical elements have to tilte and decenter to correct the aberrations, and the weight increase rapidly. The WFOV off-axis optical system's most important application is visor-type Helmet-Mounted Display(HMD)[1]. FFS optic is one of the advanced technologies in the field of optical engineering. In imaging system, it has irreplaceable advantages, such as correcting aberration, improving image quality, reducing the number of units, decreasing the weight of the system and provides great design freedom.



（a） （b）

Figure.1 Typical off-axis optical system

This paper first designs a WFOV off-axis system use one FFS lens and a normal doublet to correct the off-axis aberration brought by the reflect sphere surface. Then analyzes the fabrication of the FFS optics under the condition in existence. At last, the results of optical design are given and present the conclusion.

2 Designs of off-axis optical system

The Specifications of optical design are as blow:

Table 1 Design specifications

The size of image / inch		1	
Field of view/degree		40	
Wavelength range/nm	560	546.07	530
Exit pupil diameter / mm		10	
Weight/g		<50	
MTF@30lp/mm		>0.1	
Distortion		<10%	

Fig.2 indicates that the principal ray is not reflecting on-axis at the spherical mirror, there is an offboresight angle and bring off-axis aberration to the system. The system meridianal plot is asymmetrical, so the main aberrations are binodal astigmatism, linear astigmatism and axial coma. In traditional optical design, the strategies to correct these aberrations are[2]:

- Using a generalized reflect surface shape to reduce the bias aberration such as binodal astigmatism and axial coma.
- Tilted and decentered relay lens groups to reduce linear and binodal astigmatism.
- Using wedge to correct axial coma, linear astigmatism and lateral chromatic aberration

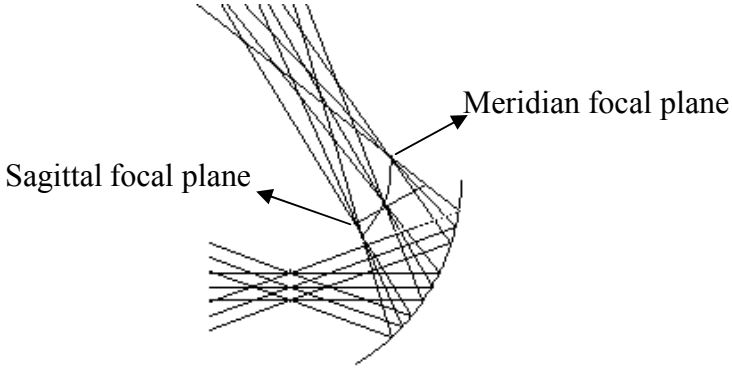


Figure.2 Two focal planes in off-axis system

Design an off-axis optical system using normal optical elements, shown in Fig.3[3]. The relay system contains 5 units: symmetrical even aspheric singlet 1, doublet 2&4, cylinder 3 and wedge 5. They all tilted and decentered from each other.

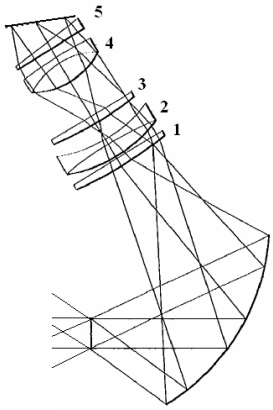


Figure.3 Off-axis system using normal optical elements

On the base of the upper system, we use a FFS lens in the system in stead of cylinder 3,doublet 4 and wedge 5. Optimize the system in optical design software CODE V[4].The final layout of system is shown in Fig.4.

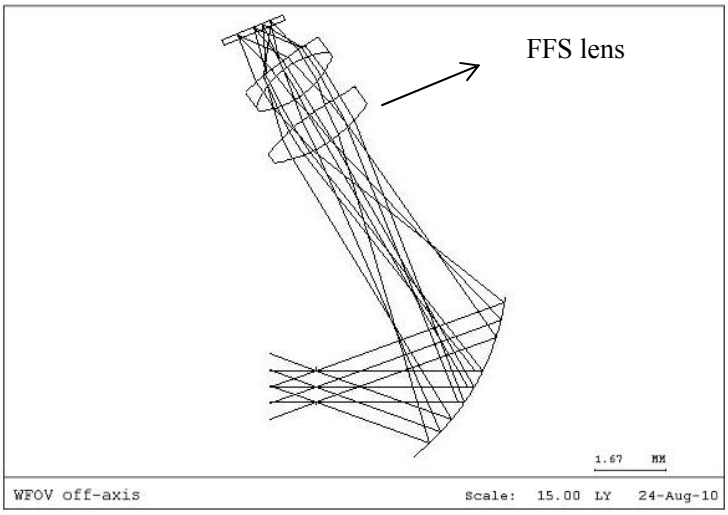


Figure.4 Layout of system

3 FFS optics

FFS optic is one of the advanced technologies in the field of imaging system, it provides great design freedom to correct off-axis aberration[5]. It can not be described in a math equation but a series of point data. Actually, any FFS can not be ray traced, so form a simple mathematic model is necessary, z is a function in term of x and y:

$$z=\left(cr^2\right)\left[1+\sqrt{\left\{1-(1+k)c^2r^2\right\}}\right]^{-1}+\sum_{ij}^{R}C_{ij}x^iy^j$$

$$\sum_{ij}^{R}C_{ij}x^iy^j=C_{00}+C_{01}y+C_{10}x+C_{02}y^2+C_{11}xy+C_{20}x^2+C_{03}y^3+C_{12}xy^2$$
$$+C_{21}x^2y+C_{30}x^3+C_{04}y^4+C_{13}xy^3+C_{22}x^2y^2+C_{31}x^3y+C_{40}x^4+\cdots \quad (1)$$

Where:

z is the amount of deviation from a plane tangent to the origin of the surface configuration.

c is curvature in polar coordinate direction.

r is radius.

x, y are the local coordinate of surface, and $r^2=x^2+y^2$.

k is quadratic term coefficient.

C_{ij} is normazied coefficient.

The designed FFS is 7-polynomial system, and discussed as above in section 2, the system meridianal surface is asymmetrical but the sagittal surface is symmetrical. So that:

$$C_{10}=C_{11}=C_{12}=C_{30}=C_{13}=\cdots=C_{70}=0 \quad (2)$$

The designed FFS equation is:

$$z=\left(cr^2\right)\left[1+\sqrt{\left\{1-(1+k)c^2r^2\right\}}\right]^{-1}+C_{00}+C_{01}y+C_{02}y^2+C_{20}x^2+C_{03}y^3+C_{21}x^2y$$
$$+C_{04}y^4+C_{22}x^2y^2+C_{40}x^4+C_{05}xy^5+C_{23}x^2y^3+C_{41}x^4y+C_{06}y^6+C_{24}x^2y^4$$
$$+C_{42}x^4y^2+C_{60}x^6+C_{07}y^7+C_{25}x^2y^5+C_{43}x^4y^3+C_{61}x^6y \quad (3)$$

Asphere is a special type of FFS, the used asphere equation in designed system is :

$$z = \left(cr^2\right)\left[1+\sqrt{\left\{1-(1+k)c^2r^2\right\}}\right]^{-1} + a_4r^4 + a_6r^6 \tag{4}$$

Where:

z,c,r,k have same meaning as in Eq.1.

a_4、a_6 are coefficientes.

Compare Eq.3 with Eq.4, we found that the first part of both equation are the same, just the last part have differences, but with the relation of $r^2=x^2+y^2$, we can get that:

$$\begin{aligned} C_{02} &= C_{20} = 0 \\ C_{04} &= C_{40} = \frac{1}{2}C_{22} = a_4 \\ C_{06} &= C_{60} = \frac{1}{3}C_{24} = \frac{1}{3}C_{42} = a_6 \end{aligned} \tag{5}$$

So we get the original coefficients of FFS, and substitute FFS for Asphere to optimize the system.

4 Fabrication of FFS optics

4.1 Selection of material

In traditional optical systems, optical glasses are common material. Glasses have a number of desirable features, low-impact strength and high density. Optical plastics like polymethyl methacrylate(PMMA) and polystyrene(PS) characterized by a high-impact strength and low density. Optical plastics have good performance in impact resistance and density compared with glasses. While the optical performance of plastics constantly upgrade, the refractive index range can meet the requirements. The designed WFOV off-axis optical system's biggest usage is HMD, these two characters is the most important in HMD optical design[6]. So the designed FFS lens choose PMMA as the optical material.

4.2 Fabrication

There are various processing methods of optical plastics, like optical precision processing, injection molding and pressing formation[7]. Optical processing points that process plastics directely on the machine as glasses. Injection molding points means that precisely-machine a metal mold, then inject the liquid plastic into the mold[8]. Both methods are feasible and effective. With the technology of precise production development, the fabrication of PMMA lens is no longer a question.

5 Optical performance

The image quality evaluation for off-axis system use lateral aberration as basic meaning. The weight of FFS system compare to the original system is shown in table 2.

Table 2 Weight of two systems

FFS system	Original system
34.2g	107.7g

Fig. 5 shows the diffraction MTFs for the tangential and sagittal ray fan plot across the field angles of ±20° . The lowest point at spatial resolution of 30lp/mm is 0.352.

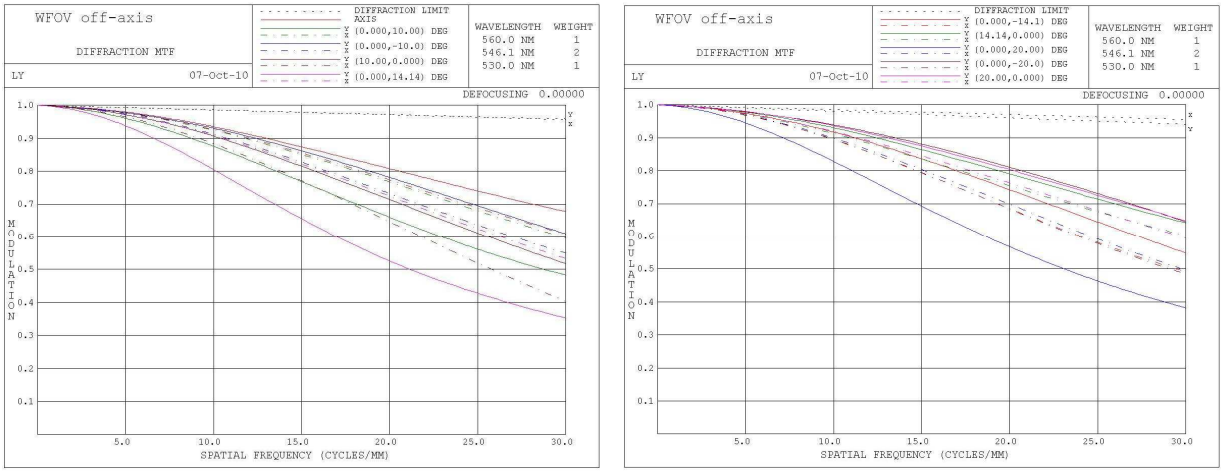


Figure.5 Plots of diffraction MTFs

The distortion of the FFS WFOV off-axis system is shown in Fig.6 below. The biggest distortion at edge field is 8.91%, less than 10%, meets the requirement as in Table 1.

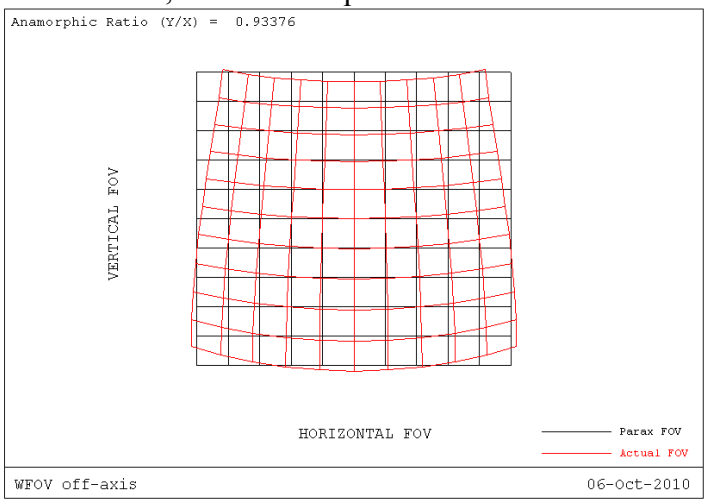


Figure.6 Plot of Distortion

Fig.7 shows the system's lateral color aberration curves, the biggest lateral color aberration is 0.021mm.

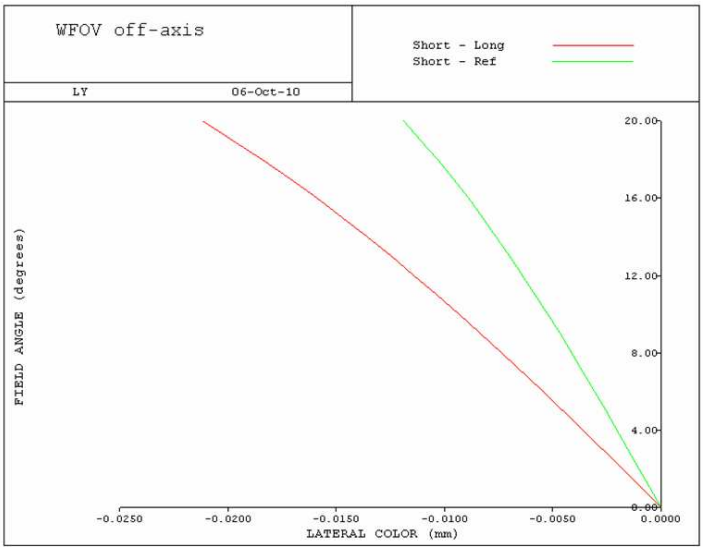


Figure.7 Plot of lateral color aberration

6 Conclusion

The present paper mainly research on the FFS optics' application design in WFOV off-axis system. The FFS system reduced 3 lenses compare to the original system which use normal surface figure and glass lenses. The volume and weight of system are obviously decreased, and the image quality meet the requirements. Exhibits FFS optics' great superiority and vast development space for structure simplification and the optical system strict with weight. The WFOV off-axis optical system design is an important reference to visor-type HMD.

References

[1] Jannick P. Rolland: Wide-angle, off-axis, see through head mounted display. Opt.Eng Vol.39(2000), p.1760-1767

[2] C.Bill Chen: Wide field of view, wide spectral band off-axis helmet-mounted display optical design. Proceedings of SPIE Vol. 4832 (2002),p.61-66

[3] Chungte W.Chen: Helmet Visor Display Employing Reflective, Refractive And Difftactive Optical Elements: US,Patent5,526,183.(1996)

[4] HE Yu-lan,LIU Jun, JIAO Ming-yin, et al：Design of free form surface optical system in CODE V． Journal of Applied Optics Vol.27(2006), p.120-123．

[5] M.M.Talha, Yongtian Wang, Dewen Cheng,et al. Design of a Compact Wide Field of View HMD Optical System using Freeform Surfaces ． Proc. of SPIE Vol. 6624, 662416 (2008), p. 662416-1-14

[6] Kazutaka Inoguchi, Motomi Matsunaga, Shoichi Yamazaki. The development of a high-resolution HMD with wide FOV using the shuttle optical system. Proceedings of SPIE Vol.6955(2008)

[7] Garrad, K. P., A. Sohn, R. G. Ohl, R. Mink and V. J. Chambers: Off-axis biconic mirror fabrication.Proceedings of the Third International Meeting of the European Society for Precision Engineering andNanotechnology (EUSPEN) (2002), p. 277 – 280,.

[8] DU Xue, WANG Er-qi, LI Rong-bin, et al: Investigation on error factor of injection molding plastic for freeform optics. Journal of Applied Optics Vol28(2007), p.684-688.

doi:10.4028/www.scientific.net/AMM.50-51.812

Orthodontic EMR Cloud Based on 2-Tier Cloud Architecture

Wenjun Zhang[1,a], Aimin Yang[2,b] and Yi Liu[3,c]

[1]Institute of Applied Computer Technology, China Women's University, Beijing, 100101, China

[2]College of Science, Hebei Polytechnic University, Shijiazhuang, Hebei, 063009, China

[3]Department of Orthodontics, Peking University School of Stomatology, Beijing, 100081, China

[a]voicefromzhwj@yahoo.com.cn, [b]aimin@heut.edu.cn, [c]liuyi@bj163.com

Keywords: EMR, Orthodontics, Cloud computing, 2TCAR, RIA, REST, SimpleDB

Abstract. Now for orthodontists there are no commercial orthodontic EMR systems suitable for their clinical needs in China, so we study orthodontist's daily workflow, analyze the requirements, and finally develop the orthodontic EMR Cloud for Peking University School of Stomatology. We propose and adopt 2-Tier Cloud-ARchitecture (2TCAR), which contains rich client tier based on Rich Internet Application (RIA) and server-side Cloud tier based on SimpleDB, to develop the orthodontist's EMR Cloud for orthodontics according to orthodontist's workflow. In the 2TCAR the rich client tier is maximized to implement almost all functionalities of user interfaces and transaction logic in the EMR. Functionalities in server-side Cloud tier are simplified only to implement data storage and query. Communication between the two Cloud tiers is also simplified via REST. In the article we research corresponding technologies such as Cloud computing, REST, Flex in RIA and SimpleDB Cloud. Further, in the orthodontic EMR Cloud we use Flex to implement UI presentation & interaction, transaction logic, and REST requests & responses in rich client tier; then design SimpleDB Cloud in server-side Cloud tier, and communication between two tiers via REST. And the EMR Cloud is integrated with existing Resister, Ward and Drugstore information systems. The practice shows that the orthodontic EMR Cloud based on the 2TCAR can fit seamlessly into orthodontist's daily workflow and effectively replace current paper medical records.

Introduction

Electronic Medical Record (EMR) has recently become critical to global medical care improvements [1]. Unlike paper medical records, EMR has the capability to integrate heterogeneous information into medical treatment information, and is able to satisfy the requirements of clinical medicine, medical management, research, education, and other legal medical affairs. EMR can create, store and apply a patient's medical records, such as examination reports, diagnosis, treatment records, nursing records, and medical images through electronic files.

There have been many EMR systems applied in many different subspecialties. However, few published articles have focused on the EMR systems for orthodontics. Now in the field of orthodontics no commercial EMR is suitable for orthodontist's needs.

Orthodontics is the branch of dentistry concerned with the study of growth of the craniofacial complex. The detection and correction of malocclusion and other dental abnormalities is one of the most important and critical phases of orthodontic diagnosis [2]. So we develop the EMR system for orthodontics to solve the above problems. The EMR is intended to replace current paper medical records used in orthodontics.

To be truly useful, of course, the EMR must fit seamlessly into orthodontist's daily workflow. The orthodontist's interface is the key problem facing the EMR and this design problem is the most important factor influencing future success. If the orthodontist is excited and challenged by the system he/she sees and it produces income and expedites the care of his/her patients, then we will see instant as well as long-term success. Therefore, we use RIA such as Flex to develop the user interface.

In addition Cloud computing [3] is an emerging technology which promises to revolutionize software and business life-cycles. By making data available in the cloud, it can be more easily and ubiquitously accessed, often at much lower cost, increasing its value by enabling opportunities for enhanced collaboration, integration, and analysis on a shared common platform.

Along with the advancements of the Cloud technology, new possibilities for Internet-based applications development are emerging. These new application models are mainly based on HTML on the client side, complex Web Services based on SOAP, and so on in Figure 3. And almost all components such as Web services, application logic, transaction logic, and data storage, are deployed and executed in the back-end Cloud. This traditional Cloud application architecture results in many problems [4]: 1) Web services API in Cloud is too complex because all services need to be defined; 2) Running cost is too high because all transaction logic is executed in Cloud and consumes much CPU computation; 3) Interaction between clients and Cloud servers is frequent and low efficient because in the client-side HTML Web pages do not contain the script, can not interact with users, only show contents; 4) The client-side requires data and contents only through request-response session because all datum and states host on the server-side. The content from server contains not only data but also a lot of redundant display formats. 5) Unlike desktop applications, Internet browsers (IE, Firefox) are not equipped with multifunction controls such as DataGrid, Tree, and PieChart.

The above problems stem from no using Rich Internet Application (RIA) and inappropriate functionality segmentation between client-side and server-side. That is, Web services, application logic and transaction logic are overly concentrated in the back-end Cloud. In this article we propose 2-Tier Cloud ARchitecture (2TCAR) [6], which by maximizing the client-side functionality via RIA, on the client-side RIA implements UI presentation, Web services, application logic and transaction logic; on the server-side functionalities are simplified and minimized into only storing and querying data via Amazon's SimpleDB Cloud; the API interfaces between the RIA client-side and the server-side are simplified via minimized Representational State Transfer (REST).

In this article, we adopt 2TCAR to develop the orthodontist's EMR Cloud for orthodontics according to orthodontist's workflow.

Workflow of the EMR Cloud on Orthodontics

Requirement Analysis and Features of the EMR Cloud. Firstly focusing on the contents of orthodontics and orthodontist' clinical practices, we analyze requirements and features of the EMR Cloud.

Preparation. Before an orthodontist begins to evaluate a patient he or she must have access to a base of knowledge that summarizes the current status of the patient, called his problem list. The EMR must offer a one page summary screen that is graphical and highly interactive that allows the total content of medical knowledge about a patient to be displayed. The design of this page is critical to the ability of the EMR to expedite care. The EMR Cloud should allow the orthodontist to instantly navigate a whole range of issues in seconds. This is similar to a heads-up display in the cockpit. This needs to have intelligence built into the display so that the practice pattern of the orthodontist can be incorporated into the presentation sequence.

Patient–Orthodontist Encounter. The patient–orthodontist encounter can be designed in several modalities so that the data is collected after the dialogue or during the process. The problem with current other systems is that they are essentially blank forms that the orthodontist has to complete. Most EMR systems ask the physicians to write their own templates and that is a job they are not prepared to do with their limited time and resources. These templates are the key to several doors if they are properly done and executed. These intelligent templates would include the diagnosis and treatment options allowable and also the range of test modalities approved for evaluation. It is totally irresponsible to expect a practicing physician to write and update his own templates. This is a complex and time consuming job and should be the basis of a whole orthodontics to support the new

EMR systems. So according to experts' suggestions on orthodontics, intelligent templates for orthodontists are developed and embedded in different sub-workflows in the EMR Cloud.

Decision-Support. The orthodontists in today's world are involved in a rapidly changing landscape of new knowledge and breakthroughs in medical research. To keep pace with this expanding universe of technology, the EMR Cloud must be made a part of this frontier. We need an intelligent support system that knows differential diagnosis and disease complications. It needs to reason for the orthodontist and be his intelligent assistant. This means that the template system that is created should be expanded with AI tools that apply this knowledge in problem solving tools and techniques. Using a highly structure relational database foundation for medical knowledge and the AI tools, we should be able to prepare some very useful tools that would interact with the orthodontist when needed. Therefore some intelligent support systems need to be integrated into the EMR. And the result data and images from the systems can be also embedded in the EMR.

Workflow of the Orthodontic EMR Cloud. The medical records of orthodontics are cumbersome, including three main examinations: general, special orthodontics, and temporomandibular-joint, each of which contains lots of items. But according to the above requirement analysis and orthodontists' daily workflow, it can be mainly divided into: patient's complaint, the relationship between front and back molars, degree of tooth-mouth-opening, and x-ray film of tooth positions, and so on.

Therefore, running processes of the EMR Cloud must be ensured to completely match the above orthodontist's daily workflow. And after finishing the three examinations, all related data and images, such as x-ray films and 3D figures provided by the decision-support system, will be integrated and embedded into the EMR Cloud. At the same time the patient's problem list will be generated automatically and intelligently so as to facilitate the orthodontist to understand the patient's all problems and determine the final treatment programs. Additionally, the EMR Cloud can also use a large number of patients' data for clinical research, analysis and statistics.

The workflow and software modules of the EMR Cloud mainly includes: Patient's Basic Information, Teeth Status, Initial Examination, Orthodontic Examination, Temporomandibular Joint, Problem List, Initial Diagnosis, Initial Disposition, Correct Programs, and Treatment as shown in Figure 1. To store data from different sub-workflows, we create the corresponding tables in the database and associations between the tables via the primary key case-id as shown in Figure 2.

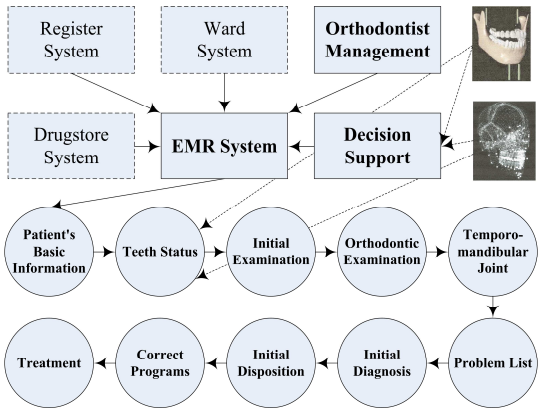


Fig. 1 Workflow of the EMR Cloud

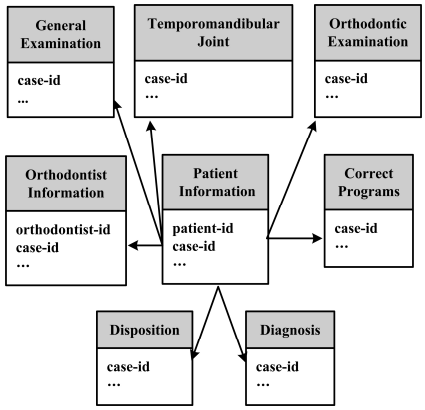


Fig. 2 Tables in the EMR's database

Web Architecture of the Orthodontic EMR Cloud Based on 2-Tier Cloud

2-Tier Cloud Architecture. Traditional development methods of Cloud applications as shown in Figure 3 are very complex. The main reasons are: 1) the architecture of Cloud implement is inappropriate; 2) the RIA's power can not be used fully; 3) almost all functionalities of application logic and transaction logic are overly concentrated on the server side. That is, inappropriate functionality segmentation between client-side and server-side.

To address the above problems, we propose a novel Cloud implement architecture: 2-Tier Cloud Architecture, shown in Figure 4.

In the EMR the main idea of 2-Tier Cloud Architecture [4] is that by using RIA on the client side rich client tier is maximized, implements almost all functionalities such as UI presentation, human-machine interaction, and application logic and transaction logic except of data storage. Rich client tier implements a complete application; by using SimpleDB Cloud, the Cloud tier is greatly simplified, and implements to store, update and query data sets in the cloud; the API interfaces between the RIA client tier and Cloud tier on the server side are also simplified via minimized REST.

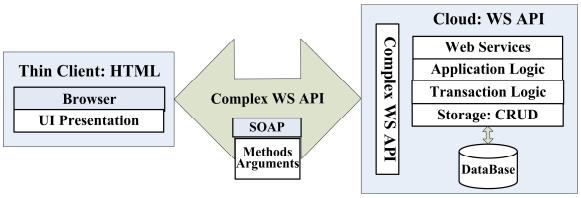


Fig.3 Traditional cloud architecture

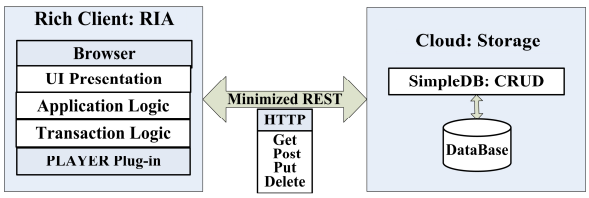


Fig.4 2-Tier cloud architecture

Rich Client Tier Based on Flex [5]**.** Almost all 2TCAR's functionalities in the EMR Cloud are implemented by MXML and ActionScript.

UI Presentation and Human-Machine Interaction. Using MXML, we can add components such as different controls and data charts into containers such as panel. Effects, transitions, data models, and data binding can is also introduced into UI design. Human-machine interaction is implemented by ActionScript and event-driven.

Application Logic and Transaction Logic. All logic functionalities are programmed by ActionScript's Classes, Interfaces and predifined Class library.

Communication between Rich Client Tier and Server-Tier in SimpleDB. SimpleDB REST calls are made using HTTP GET requests. The Action query parameter provides the called method and the URI specifies the target of the call. Additional call parameters are specified as HTTP query parameters. The response is an XML document that conforms to a schema. Using Flex's HTTPService component a request from rich client tier to server-tier in SimpleDB Cloud is constructed, but the format of the request has to comply with SimpleDB's REST API standard. And the response from SimpleDB is received by HTTPService's callback function, and is parsed to acquire the data and refresh user interface.

Server-Side Tier Based on SimpleDB Cloud. Amazon's SimpleDB [6] is a Web service for running queries on structured data. This service can work in close conjunction with RIA, collectively providing the ability to store, process and query data sets in the cloud. These services are designed to make Web-scale computing easier and more cost-effective for developers. SimpleDB is a database that is accessed via REST. SimpleDB easily uses and provides the core functionality of a database – real-time lookup and simple querying of structured data – without the operational complexity. SimpleDB requires no schema, automatically indexes our data and provides a simple API for storage and access.

When accessing SimpleDB [7] to store or query data into or from tables in the EMR's database via REST, we must provide the following items so that the request can be authenticated.

Authentication.

- AWSAccessKeyId—Our AWS account is identified by your Access Key ID, which AWS uses to look up our Secret Access Key.
- Signature—Each request must contain a valid HMAC-SHA signature, or the request is rejected. A request signature is calculated using your Secret Access Key, which is a shared secret known only to you and AWS.
- Date—Each request must contain the time stamp of the request.

Depending on the API we're using, we can provide an expiration date and time for the request instead of or in addition to the time stamp. For details of what is required and allowed for each API, see the authentication topic for the particular API.

Authentication Process. The following is the series of tasks required to authenticate requests to AWS using an HMAC-SHA request signature. It is assumed we have already created an AWS account and received an Access Key ID and Secret Access Key. We firstly perform the three tasks for creating authentication requests: 1) Create a request; 2) Create a HMAC-SHA signature; 3) Send the request and signature to AWS. Then AWS performs the next three tasks: 1) Retrieve your Secret Access Key; 2) Create a HMAC-SHA signature; 3) Compare the two signatures.

Authenticating REST Requests and REST Responses. We can send REST requests over either HTTP or HTTPS. Regardless of which protocol we use, we must include a signature in every REST request. If a REST request authenticated are confirmed SimpleDB will be responding a corresponding response.

Design and Implement of the Orthodontic EMR Cloud Based on 2-Tier Cloud

According to the above workflow and 2TCAR architecture of the EMR Cloud, we use Flex to implement almost all functionalities in the EMR Cloud exception of storing and querying data in the SimpleDB Cloud: 1) a special user interface (UI) for each sub-workflow detailed in Figure 1; 2) Application logic and transaction logic; 3) Communication between rich client tier and server-tier in SimpleDB. And different sub-workflow's UI can be specified and switched through clicking the corresponding button in the left-side main menu.

The data in different sub-workflow's UI is stored into and retrieved from the corresponding table in SimpleDB Cloud using Flex communication components such as HTTPService via REST. The resulting data and images from the Decision-Support systems can be also embedded in the EMR as shown in Figure 5.

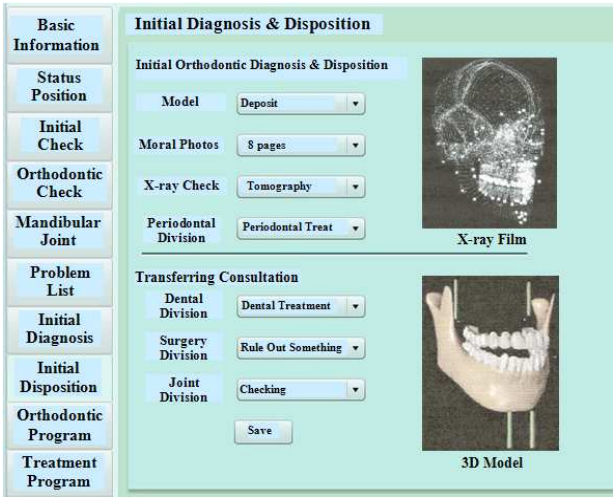


Fig. 5 User interface of initial disposition

The intelligent templates for orthodontists detailed in Figure 6 are also developed and embedded in different sub-workflows in the EMR Cloud. At the same time the patient's problem list can be generated automatically and intelligently from the data in the sub-workflows.

Additionally we also developed the system of query, analysis & statistics for research and education as shown in Figure 7. The EMR Cloud can also connect with existing Resister, Ward and Drugstore information system so as to share all data.

In short working processes of the EMR Cloud completely follows orthodontist's daily workflow to become orthodontist's intelligent assistant. In the EMR Cloud Flex greatly enhances the interactive user experience and implements the workflow on the rich client-tier in the Cloud; SimpleDB Cloud

also simply implements data storage and query on the server-tier in the Cloud; the API interface between the rich client-tier and server-tier in the Cloud is also simplified via minimized REST.

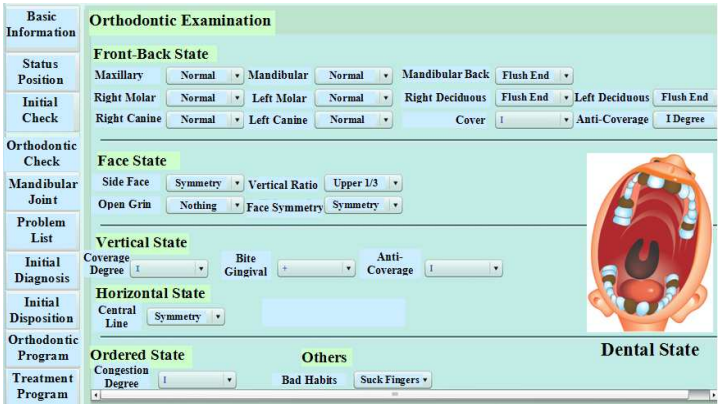


Fig. 6 User interface of orthodontic examination

Fig. 7 User interface of query, analysis & statistics

Conclusions

In the article the EMR Cloud for orthodontics is successfully developed according to orthodontist's daily workflow and 2-Tier Cloud architecture, and is put into operation in Peking University School of Stomatology. The practice shows that the EMR Cloud can fit seamlessly into orthodontist's daily workflow and effectively replace current paper medical records used in orthodontics.

Acknowledgement

This research project has been sponsored by Department of Orthodontics under Peking University School of Stomatology. We would like to express our heartfelt thanks to the sponsors.

References

[1] H.S. Fraser, F. Biondich and P. Moodley: Informat. Prim. Care Vol.13 (2005), p.83

[2] D. Zhang and M. Sonka: Springer ICMB 2010 Vol.1 (2010), p. 344

[3] S. Lohr: Google and IBM join in cloud computing research, New York Times (October 8, 2007)

[4] E. Jamil: What really is SOA? A comparison with Cloud Computing, Web 2.0, SaaS, WOA, Web Services, PaaS and others, White Paper, Soalib Incorporated (2009)

[5] W.J. Zhang: Journal of Software Engineering & Applications Vol. 3 (2010), p. 73

[6] W.J. Zhang: 2nd International Conference on Computer Engineering and Technology Vol. 6 (2010), p. 52

[7] Information on http: //simplerdb.rubyforge.org

©2010 Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.818

THE STRUCTURE OF A GROUP BIPRODUCT

Xiuyan Jiang, Fengzhi Sun, Guoqing Liu, Xing Qiao

Department of Mathematics, Daqing Normal University, Daqing 163712, China
Email:jiangxy1963@sohu.com

Keywords: Hopf group coalgebra; group byproduct; cosemisimplicity

Abstract. In this paper we mainly give the necessary and sufficient conditonsfor a group biproduct to be a Hopf group coalgebra and state some properties such as group integrals ,grouplike elements and cosemisimplicity.

Introduction and Preliminaries

It is well-known that a Smash product algebra and a Smash coproduct coalgebra arising from the theory of groups and affine group schemes respectively.([1])are very important structures in the theory of Hopf algebras([2]). Roughly speaking, a biproduct affords a Smash product algebra structure and a Smash coproduct coalgebra structure.

A Hopf group coalgebra was introduced by Turaev([3]) and the author gave its purely algebraic properties.([4],[5])In this paper we firstly introduce the notion of group biproducts and give the necessary and sufficient conditions for it to be a Hopf group coalgebra. Moreover we also reveal some properties such as group integrals, grouplike elements and cosemisimiplicity.

Definition 1 Let H be a Hopf group coalgebra and A a left H -module(i.e., for any $\alpha \in \pi$ A is a H_α -module) which is also an algebra. A is called a group H -module algebra if the following are satisfied:

（1） $h\bullet ab = \sum (h_{1\alpha}\bullet a)(h_{2\beta}\bullet b), h \in H_{\alpha\beta}, a, b \in A$;

（2） $h.1 = \varepsilon(h)1, h \in H_i$.

Definition 2 Let H be a Hopf group coalgebra and A a left H -comodule which is also an algebra. A is called a group H -comodule algebra if the following are satisfied:

（3） $\sum (ab)_{((-1),\alpha)} \otimes (ab)_{(0)} = \sum a_{((-1),\alpha)} b_{((-1),\alpha)} \otimes a_{(0)} b_{(0)}, a, b \in A;$

（4） $\sum 1_{((-1),\alpha)} \otimes 1_{(0)} = 1_\alpha \otimes 1.$

Definition 3 Let H be a Hopf group coalgebra and C a left H -comodule which is also an coalgebra. C is called a group H-comodule coalgebra if the following are satisfied:

（5） $\sum c_{1((-1),\alpha)} c_{2((-1),\alpha)} \otimes c_{1(0)} \otimes c_{2(0)} = \sum c_{((-1),\alpha)} \otimes c_{(0)1} \otimes c_{(0)2}, c \in C;$

Project supported by the Natural Science Foundation of Heilongjiang Province of China(No.A200906) and the Scientific Subject of Education Bureau of Heilongjiang(No.11553002).

（6） $\sum \varepsilon(c_{(0)})c_{((-1),\alpha)} = \varepsilon(c)1_\alpha, c \in C.$

Main Results

We can obtain the following lemmas from easy and straightforward computations.

Lemma 4 Let H be a Hopf group coalgebra and A a left H-module algebra. For any $\alpha \in \pi, A\Omega H_\alpha = A \otimes H_\alpha$ as a k-space. Then $A\Omega H_\alpha$ is an associative algebra with the unit $1\Omega 1_\alpha$ under the multiplication given by

$$(a\Omega h)(b\Omega g) = \sum a(h_{1i} \bullet b)\ \Omega h_{2\alpha} g,\ a, b \in A,\ h,\quad g \in H_\alpha.$$

Lemma 5 Let H be a Hopf group coalgebra and C a left group H-comodule algebra. Then $C\nabla H = \{C \otimes H_\alpha\}_{\alpha \in \pi}$ is an group coalgebra under the comultiplication given by

$$\Delta(c\nabla h) = \sum c_1 \nabla c_{2((-1),\alpha)} h_{1\alpha} \otimes c_{2(0)} \nabla h_{2\beta}, c \in C, h \in H_{\alpha\beta}.$$

Theorem 6 Let H be a Hopf group coalgebra and A a left group H-comodule coalgebra and a . H-module algebra. Then the following are equivalent:

（I） $A\Omega H$ is a semi-Hopf group coalgebra;

（II）A is a left H group comodule algebra and a H_i-module coalgebra, ε_A is an algebra map and

（7） $\Delta(1_A) = 1_A \otimes 1_A;$

（8） $\Delta(ab) = \sum a_1(a_{2((-1,i))} \bullet b_1) \otimes a_{2(0)} b_2, a, b \in A;$

（9） $\sum (h_{1i} \bullet a)_{((-1),\alpha)} h_{2\alpha} \otimes (h_{1i} \bullet a)_{(0)} = \sum h_{1\alpha} a_{((-1),\alpha)} \otimes h_{2i} \bullet a_{(0)}, a \in A, h \in H_\alpha.$

Proof. First we claim that $\varepsilon_{A\Omega H}$ is an algebra map if and only if ε_A is an algebra map and $\varepsilon_A(h \bullet a) = \varepsilon_{H_i}(h)\varepsilon_A(a)$ for any $a \in A, h \in H_i$.

In fact, if $\varepsilon_{A\Omega H}$ is an algebra map, then for any $a, b \in A, h, g \in H_i,$ we have

$\sum \varepsilon_{A\Omega H}((a\Omega h)(b\Omega g)) = \sum \varepsilon_{A\Omega H}(a(h_{1i} \bullet b)\Omega h_{2i} g) = \varepsilon_A(a)\varepsilon_A(b)\varepsilon_H(h)\varepsilon_H(g),$ we obtain $\varepsilon_A(ab) = \varepsilon_A(a)\varepsilon_A(b)$ by letting $h = g = 1_i$ and $\varepsilon_A(h \bullet a) = \varepsilon_A(a)\varepsilon_{H_i}(h)$ by taking $a = 1_A$ and $g = 1_i.$ Conversely it is easy to prove .

Next we claim that $\Delta_{\alpha,\beta}(1\Omega 1_{\alpha\beta}) = 1\Omega 1_\alpha \otimes 1\Omega 1_\beta$ if and only if $\rho(1_A) = 1_\alpha \otimes 1_A$ and $\Delta(1_A) = 1_A \otimes 1_A.$ Indeed, if $\Delta_{\alpha,\beta}(1\Omega 1_{\alpha\beta}) = 1\Omega 1_\alpha \otimes 1\Omega 1_\beta$, then

$$\sum 1_1 \Omega 1_{2((-1),\alpha)} \otimes 1_{2(0)} \Omega 1_i = 1\Omega 1_\alpha \otimes 1\Omega 1_i ,$$

we get $\rho(1_A) = 1_\alpha \otimes 1_A$ by applying $\varepsilon_A \otimes id_{H_\alpha} \otimes id_A \otimes \varepsilon_{H_i}$ and $\Delta(1_A) = 1_A \otimes 1_A$ by taking $\alpha = i$ and applying $id_A \otimes \varepsilon_{H_i} \otimes id_A \otimes \varepsilon_{H_i}$. Conversely it is trivial to prove .

Finally we claim that $\Delta_{A\Omega H}$ is commutative if and only if for any $a,b \in A$, $h,g \in H_{\alpha\beta}$, $\sum (a(h_{1i} \bullet b))_1 \Omega (a(h_{1i} \bullet b))_{2((-1),\alpha)} h_{2\alpha\beta 1\alpha} \otimes (a(h_{1i} \bullet b))_{2(0)} \Omega h_{2\alpha\beta 2\beta} = \sum a_1((a_{2((-1),}$ $_{,\alpha)} h_{1\alpha})_{1i} \bullet b)\ \Omega (a_{2((-1),\alpha)} h_{1\alpha})_{2\alpha} b_{2((-1),\alpha)} \otimes a_{2(0)} (h_{2\beta 1i} \bullet b_{(0)}) \Omega h_{2\beta 2\beta}$.

Since, on one hand we have

$$\Delta_{\alpha,\beta}((a\Omega h)(b\Omega g)) = \sum (a(h_{1i} \bullet b))_1 \Omega (a(h_{1i} \bullet b))_{2((-1),\alpha)} h_{2\alpha\beta 1\alpha} \otimes (a(h_{1i} \bullet b))_{2(0)} \Omega h_{2\alpha\beta 2\beta} ;$$

on other hand we also have

$$\Delta_{\alpha,\beta}(a\Omega h)\Delta_{\alpha,\beta}(b\Omega g) = \sum a_1((a_{2((-1),\alpha)} h_{1\alpha})_{1i} \bullet b_1)\ \Omega (a_{2((-1),\alpha)} h_{1\alpha})_{2\alpha}$$

$b_{2((-1),\alpha)} \otimes a_{2(0)} (h_{2\alpha\beta 1i} \bullet b_{2(0)}) \Omega h_{2\alpha\beta 2\beta}$. we get $\rho(ab) = \rho(a)\rho(b)$ and the equation （9） by letting $\beta = i, a = 1, h = 1_\alpha$ and applying $\varepsilon_A \otimes id_{H_\alpha} \otimes id_A \otimes \varepsilon_{H_i}$, and we get

$\sum (h \bullet b)_1 \otimes (h \bullet b)_2 = \sum h_{1i} \bullet b_1 \otimes h_{2i} \bullet b_2$ and the equation （8） by letting $\alpha = \beta = i$, $h = 1_i$ and applying $id_A \otimes \varepsilon_{H_i} \otimes id_A \otimes \varepsilon_{H_i}$. The sufficiency is straightforward.

Theorem 7 Let H be a Hopf group coalgebra and A a left H group comodule coalgebra and a left H group module algebra.

（I）If $A\Omega H$ is a Hopf group coalgebra, then id_A has an inverse in the convolution algebra. $Hom_k(A, A)$.

（II）If H is a Hopf group coalgebra with the antipode $S^H = \{S_\alpha^H\}_{\alpha\in\pi}$ and ω_A is the inverse of id_A, then $A\Omega H$ is a Hopf group coalgebra and the antipode given by

$$S = \{S_\alpha : A\Omega H_\alpha \to A\Omega H_{\alpha^{-1}}, S_\alpha(a\Omega h) = \sum (1\Omega S_\alpha^H(a_{((-1),\alpha)} h))(\omega(a_{(0)}\Omega 1))\}_{\alpha\in\pi}.$$

Proof. （I）Supposing S is the antipode of $A\Omega H$, we define $\overline{S} = \{\overline{S_\alpha} : H_\alpha \to H_{\alpha^{-1}}, \overline{S_\alpha}(h) = (\varepsilon_A \otimes id_{H_{\alpha^{-1}}}) S_\alpha(1\Omega h)\}_{\alpha\in\pi}$, then we claim that $\overline{S}$ is the antipode of H. Indeed, for any $h \in H_i$ we have

$$\sum \overline{S_{\alpha^{-1}}}(h_{1\alpha^{-1}}) h_{2\alpha} = \sum (\varepsilon_A \otimes id_{H_\alpha})\ S_{\alpha^{-1}}(\ (1\Omega h)_{1\alpha^{-1}})\ (1\Omega h)_{2\alpha} = (\varepsilon_A \otimes id_{H_\alpha})\varepsilon(1\Omega 1)$$

$= \varepsilon_H(1)1_\alpha$ and $\sum h_{1\alpha^{-1}}\overline{S_\alpha}(h_{2\alpha}) = \sum(\varepsilon_A \otimes id_{H_{\alpha^{-1}}})(1\Omega h)_{1\alpha^{-1}} S_\alpha((1\Omega h)_{2\alpha})$

$= (\varepsilon_A \otimes id_{H_{\alpha^{-1}}})\varepsilon(1\Omega 1) = \varepsilon_H(1)1_{\alpha^{-1}}$.

We define $\omega: A \to A, a \mapsto (id_A \otimes \varepsilon_{H_i})((1\Omega a_{((-1),i)})S_i(a_{(0)}\Omega 1))$, then we claim ω is the inverse of id_A. Since, for any $a \in A$ we have

$id_A * \omega(a) = \sum(id_A \otimes \varepsilon_{H_i})((a_1\Omega 1)(1\Omega\ a_{2((-1),i)}S_i(a_{2(0)}\Omega 1))$

$= \sum(id_A \otimes \varepsilon_{H_i})((a_1\Omega a_{2((-1),i)})S_i(a_{2(0)}\Omega 1)) = \sum(id_A \otimes \varepsilon_{H_i})((a\Omega 1)_{1i}$

$S_i((a\Omega 1)_{2i})) = \varepsilon_A(a)1$ and $(\omega * id_A)(a) = \sum(id_A \otimes \varepsilon_{H_i})((1\Omega a_{1((-1),i)})S_i(a_{1(0)}\Omega 1))a_2 =$

$= \sum(id_A \otimes \varepsilon_{H_i})(S_i(a_{1(0)}\Omega a_{1((-1),i)})(a_{2(0)}\Omega 1)) = \sum(id_A \otimes \varepsilon_{H_i})(S_i(a\Omega 1)_{1i})((a\Omega 1)_{2i}) =$

$\varepsilon_A(a)1$.

（II）Indeed, for any $a \in A, h \in H_i$ we have $\sum S_{\alpha^{-1}}(a_1\Omega a_{2((-1),\alpha^{-1})}h_{1\alpha^{-1}})(a_{2(0)}\Omega h_{2\alpha}) =$

$\sum(1\Omega S^H_{\alpha^{-1}}(a_{((-1),\alpha^{-1})}h_{1\alpha^{-1}}))(\omega(a_{(0)1})a_{(0)2}\Omega h_{2\alpha}) = \sum\varepsilon_A(a)(1\Omega S^H_{\alpha^{-1}}(h_{1\alpha^{-1}}))(1\Omega h_{2\alpha}) = \varepsilon_A(a)$

$\varepsilon_{H_i}(h)1\Omega 1$ and $\sum(a_1\Omega a_{2((-1),\alpha^{-1})}h_{1\alpha^{-1}})S_\alpha(a_{2(0)}\Omega h_{2\alpha}) = \sum(a_1\Omega a_{2((-1),\alpha^{-1})}h_{1\alpha^{-1}})\ (1\Omega S^H_\alpha($

$a_{2(0)((-1),\alpha)}h_{2\alpha})(\omega(a_{2(0)}\Omega 1)) = \sum(a_1((a_{2((-1),i)}h)_{1\alpha^{-1}1i}\bullet 1)\Omega(a_{2((-1),i)}h)_{1\alpha^{-1}2\alpha^{-1}}S^H_\alpha((a_{2((-1),i)}h)_{2\alpha}$

$))(\omega(a_{2(0)})\Omega 1) = (a_1\Omega 1)(\omega(a_2)\Omega 1) = \varepsilon_A(a)\varepsilon_{H_i}(h)(1\Omega 1)$. Thus we complete the proof.

Proposition 8 Let $A\Omega H$ as mentioned, then $A\Omega H$ is commutative if and only if A and H are commutative and the action of H_i on A is trivial.

Proof. We show the necessity holds. Indeed, for any $a,b \in A, h,g \in H_\alpha$, we have $(a\Omega h)(b\Omega g) = \sum a(h_{1i}\bullet b)\Omega h_{2\alpha}g = \sum b(g_{1i}\bullet a)\Omega g_{2\alpha}h = (b\Omega g)(a\Omega h)$, then we obtain $ab = ba$ by letting $\alpha = i, h = g = 1_i$ and applying $id_A \otimes \varepsilon_{H_i}$, and we get

$$h\bullet a = \varepsilon_{H_i}(h)a$$

by letting $b = 1, h = 1_i$ and applying $id_A \otimes \varepsilon_{H_i}$.

Proposition 9 Let $A\Omega H$ as mentioned,

(I) If $\lambda_A \in A^*$ and $\lambda_H = \{\lambda_{H_\alpha}\}_{\alpha\in\pi} \in H^*$ are right integrals, then $\overline{\lambda} = (\lambda_A \otimes \lambda_{H_\alpha})_{\alpha\in\pi}$ is a right integral of $A\Omega H$;

（II）If $A\Omega H$ is also a Hopf group coalgebra, then $A\Omega H$ is cosemisimple if and only if there are left integrals λ_A and λ_H such that $\lambda_A(1)=\lambda_{H_{\alpha^{-1}}}(1)=1$ and

$\sum a_{((-1),\alpha)}\lambda_A(a_{(0)})=\lambda_A(a)1$.

Proposition 10 Let $A\Omega H$ as mentioned, then $a\Omega h=(a\Omega h_\alpha)_{\alpha\in\pi}$ is a grouplike element of $A\Omega H$ if and only if a and In fact, for any are grouplike elements $\rho_\alpha(a)=1_\alpha\otimes a$.

Proof. In fact, for any $\alpha,\beta\in\pi$ we have $\Delta_{\alpha,\beta}(a\Omega h_{\alpha\beta})=\sum a_1\Omega a_{2((-1),\alpha)}h_{\alpha\beta 1\alpha}$ $\otimes a_{2(0)}\Omega h_{\alpha\beta 2\beta}=a\Omega h_\alpha\otimes a\Omega h_\beta$,so we get $\Delta(a)=a\otimes a$ by letting $\alpha=\beta=i,h=1_i$ and applying $id_A\otimes\varepsilon_{H_i}\otimes id_A\otimes\varepsilon_{H_i}$, and we get $\Delta_{\alpha\beta}(h_{\alpha\beta})=h_\alpha\otimes h_\beta$ by applying $\varepsilon_A\otimes id_{H_\alpha}\otimes\varepsilon_A\otimes id_{H_\beta}$. We obtain $\rho_\alpha(a)=1_\alpha\otimes a$ by letting $\beta=i,h=1_\alpha$ and applying, $\varepsilon_A\otimes id_{H_\alpha}\otimes\varepsilon_A\otimes id_{H_i}$.Conversely it is trivial to prove.

Refernces

[1] R. K. Molnar, Semidirect Products of Hopf Algebras, J. Algebra, 47(1977): 29-51.
[2] S. Montgonery, Hopf Algebras And Their Actions On Rings, CBMS Vol.82, AMS, (1993).
[3] Vladimir Turaev, *Homotopy Field Theory In Dimension 3 And Crossed Group-ca tegories*, Preprint GT/0005291S.
[4] Alexis Virelizier, Hopf Group-Coalgebras, Journal Of Pure And Applied Algebra, 171(2002):75-122.
[5] Alexis Virelizier, Crossed Hopf Group-Coalgebras and invariants of links and 3-manifold, in preparation.

doi:10.4028/www.scientific.net/AMM.50-51.823

The Comprehensive Evaluation of the Human All-Round Development Based on G1- Standard Deviation and Empirical Study

Libin Zhou[1,2,a], Gang Li[2,b], Guotai Chi[1,c]

[1]School of Management, Dalian University of Technology, Dalian 116024, China

[2]Department of Economics, Northeast University at Qinhuangdao, 066004, China

[a]zhoulibin50@yahoo.com.cn, [b]dlutligang@yahoo.com.cn, [c]chigt@dlut.edu.cn

Keywords: comprehensive evaluation of Human All-Round Development; index system of Human All-Round Development; G1; Standard Deviation

Abstract. According to the connotation of the Scientific Outlook on Development, this paper constructed the comprehensive Human All-Round Development evaluation index system, determined the optimal weights of indicators by combining G1 method and Standard Deviation method and analyzed the Human All-Round Development of 14 provinces. The contribution characteristics lie on two aspects. Firstly, through the evaluation of the quality of life, quality of human and human capital environment, economic development environment, social development and environment reflect the harmony between humans and nature, humans and the society, humans and humans. Secondly, to evaluate the human All-Round development by determining the weights of indicators in using G1 and Standard Deviation could reflect its own characteristics and indicator data expertise.

Introduction

The connotation of Scientific Outlook on Development is "insisting on human-oriented, comprehensive, coordinated and sustainable development". To realize people's development in an all-round way is the aim of insisting on human-oriented scientific development. The Comprehensive Evaluation of Chinese Human All-Round Development based on the Scientific Outlook on Development is to evaluate the development of Chinese Human All-Round Development, and then to summarize experience and lessons in order to establish the relative policy and provide the reference for the central government in working out economic policies.

There are some representative index systems of Human All-Round Development, such as: the conceptual Framework and Structure of European System of Social Indicators[1].the Human Development Index HDI which is from the United Nations Development Program UNDP[2]. the Living Standards Measurement of the Word Bank[3]. the Well-off Society which is from the institute of sociology, Chinese Academy of Social Sciences[4]. Mazumdar, K come up with the measuring the Well-being of the Developing Countries in 1999[5].

The representative Evaluation Methods of Human All-Round Development, such as the objective evaluation study of Factor Analysis and cluster analysis methods，Factor Analysis method，entropy weight method[6-7]. The subjective evaluation study use AHP method to evaluate the development of the well-off society[8].

This paper constructed the comprehensive Human All-Round Development evaluation index system, determined the optimal weights of indicators by combining G1 method and Standard Deviation method, and analyzed the Human All-Round Development of 14 provinces.

The construction of evaluation index system

It referred to 4 authoritative institutions[1-4] to build the Human All-Round Development evaluation indicator system. And using the quality of life, the quality of population and the human capital environment to implement the principle of the harmony between human and using the economic development environment and the society development, environment to implement the principle of

the harmony between human and society and using the ecology environment to implement the principle of the harmony between human and nature.

According to select GDP per capita, life expectancy at birth, Gini Coefficient, personal average net income, unemployed persons and Gross National Happiness established a comprehensive evaluation model of Human All-Round Development which has 6 criteria and 31 indexes. As shown in table 1.

Tab.1 The index system of Human All-Round Development

Criterion	Indicator	weight
Economic development	Per capita GDP	0.0545
	urban per capita disposable income	0.0428
	rural per capita net income	0.0318
	the personal spent able average income	0.0275
	The proportion of tertiary industry	0.0182
Social development	Gini Coefficient	0.0438
	Social security	0.0279
	Doctors per 10,000	0.0271
	lawyers per 10,000	0.0237
	place of entertainment per million	0.0194
Ecological	the per-capita water	0.0366
	success rate of three waste	0.0279
	greenbelt area per capita	0.0214
	forest coverage	0.0199
Human capital	Registered Unemployment Rate	0.0346
	Rural employment rate	0.0433
	urban unemployment	0.0321
	population burden coefficient	0.0293
Quality of life	Gross National Happiness	0.0491
	Engel coefficient	0.0608
	life expectancy at birth	0.0298
	the per-capita housing	0.0311
	infant mortality rate	0.0277
	domestic consumption	0.0286
Quality of human	average education	0.0437
	rate of compulsory education	0.0432
	illiteracy rate	0.0351
	students in universities	0.0274
	education investment/GDP	0.0201
	Scientific and Technical Personnel	0.0195
	Popular rate of computer per 10,000	0.0221

The Evaluation model Based on Standard Deviation and G1

The thought of the Comprehensive Evaluation model

The thought of using standard deviation to calculate weight

This paper used standard deviation method to determine the weight of the Evaluation model. The Standard Deviation method is one objective weighting method based on the Coefficient of Variation of indicator data to determine the weight of indicators. This method avoids effectively subjectivity of determining the weights of concerned appraisal indicators.

The thought of evaluation model based on G1

The G1 method is a subjective weight method. Its core idea is to reflect the importance of indicators by subjectively compositor. The more important the indicator is, the higher the weight is.

The standardization of the data

The positive index is the index that the bigger the better, such as the Per Capita GDP. The negative index is the index that the smaller the better, such as the Engel coefficient.

Let V_{ki} denote the i^{th} year and the k^{th} index, let x_{ki} denote the standardization of the i^{th} year and the k^{th} index,. Let n denote the numbers of the year.

The standardization of the positive and negative index is listed below [9]:

$$x_{ki} = \frac{V_{ki} - \min\limits_{1\le i\le n}(V_{ki})}{\max\limits_{1\le i\le n}(V_{ki}) - \min\limits_{1\le i\le n}(V_{ki})} \tag{1}$$

$$x_{ki} = \frac{\max\limits_{1\le i\le n}(V_{ki}) - V_{ki}}{\max\limits_{1\le i\le n}(V_{ki}) - \min\limits_{1\le i\le n}(V_{ki})} \tag{2}$$

In the method, evaluation indices are normalized to be maximizing ones.

Weighted the standardization data

(1) Using the Standard Deviation method to determine the objective weight of indexes

The Standard Deviation method is one objective weight method based on the Standard Deviation of the index data to determine the weight of the index.

Let μ_i denote the weight of the index, σ_i denote the Standard Deviation of the index. The weight of the index μ_i [10]:

$$\mu_i = \sigma_i / \sum_{i=1}^{n} \sigma_i \tag{3}$$

The comprehensive evaluation model based on factor analysis with objective weighting helps to avoid subjectivity in ordering indicators and setting weights. Making the evaluation more reasonable, avoids effectively subjectivity of determining the weights of concerned appraisal indexes.

(2) Using the G1 method to determine the subjective weight of indicators

Weights of indicator layers against standard layers

Let r_k denote the rational evaluation to the importance ratio of one factor to another,Let w_m denote the weight of the m^{th} indicator where m=1, 2, 3, …, m, as in [11].

$$v_m = (1+\sum_{k=2}^{m}\prod_{i=k}^{m} r_i)^{-1} \tag{4}$$

The weighs of $(m\text{-}1)^{th}$,$(m\text{-}2)^{th}$,…, second, first indicators are obtained through the G1 weight of the m^{th} indicator.

$$v_{k-1}=r_k v_k, \quad k=m,m\text{-}1,...,2,1 \tag{5}$$

Weights of indicators against destination layer

Let β_k denote the weight of the k^{th} indicator against the destination layer, v_k denote the weight of the k^{th} indicator against the j^{th} standard layer, $v^{(j)}$ denote the weight of the j^{th} standard layer against destination layer.

$$\beta_k=v_k\times v^{(j)} \tag{6}$$

(3) The calculation of combination weight coefficient

Let w_i denote the comprehensive weight of the i^{th} indicator to the general goal.

The combination weighting w_i:

$$w_i = \frac{\mu_i\beta_i}{\sum_{j=1}^{n}\mu_j\beta_j} \tag{7}$$

The benefit of this combination weighting model is that it not only reflects the influence of indicator score to the weight, but also dynamically reflects the indicator data to the weight, and then make up for the defects of other models in these two aspects.

(4) The evaluation model based on combination weighting model

Let D_j denotes the comprehensive score of the j^{th} year. According to the linear weighting comprehensive evaluation model, the comprehensive score D_j:

$$D_j = \sum_{i=1}^{m} d_{ij} w_i \tag{8}$$

Empirical Study

With the data of the China Statistical Yearbook[12], we used the Standard Deviation and G1 to get weights of the Human All-Round Development Evaluation. the results of the evaluation were gained as the following table 2.

Tab.2. the results of the evaluation

provinces	Economic developmen		Social developmen		Ecological		Human capital		Quality of life		Quality of human		Human All-Round Development	
	score	order	score	order	score	order	score	order	score	order	score	order	score	order
beijing	0.1602	2	0.1152	2	0.0502	6	0.1394	1	0.1926	1	0.1930	1	0.8504	1
shanxi	0.0217	6	0.0457	4	0.0158	14	0.0733	4	0.0924	7	0.1082	6	0.3571	7
liaoning	0.0387	5	0.0598	3	0.0380	10	0.0428	13	0.1003	5	0.1258	4	0.4053	5
heilongjiang	0.0199	7	0.0435	5	0.0532	5	0.0550	7	0.0864	8	0.1017	8	0.3597	6
shanghai	0.1618	1	0.1179	1	0.0336	11	0.0966	2	0.1453	2	0.1524	2	0.7077	2
jiangxi	0.0123	10	0.0326	8	0.0657	3	0.0433	12	0.0722	11	0.0878	11	0.3139	10
shandong	0.0424	4	0.0324	9	0.0437	9	0.0649	5	0.1228	3	0.1129	5	0.4191	4
henan	0.0138	8	0.0273	11	0.0314	12	0.0476	10	0.1011	4	0.1020	7	0.3234	8
guangdong	0.0759	3	0.0328	7	0.0756	2	0.0737	3	0.0928	6	0.1269	3	0.4777	3
guangxi	0.0120	11	0.0134	13	0.0798	1	0.0442	11	0.0508	13	0.0875	12	0.2876	12
sichuang	0.0117	12	0.0246	12	0.0455	8	0.0288	14	0.0616	12	0.0774	13	0.2496	13
yunnan	0.0083	14	0.0041	14	0.0536	4	0.0488	9	0.0296	14	0.0082	14	0.1526	14
shanxi	0.0101	13	0.0310	10	0.0284	13	0.0596	6	0.0812	9	0.1001	9	0.3104	11
xinjiang	0.0132	9	0.0410	6	0.0462	7	0.0547	8	0.0729	10	0.0903	10	0.3183	9

Based on the results of table 2, we got the figure 1, the scores and curve of Human All-Round Development comprehensive evaluation All-Round Development comprehensive evaluation.

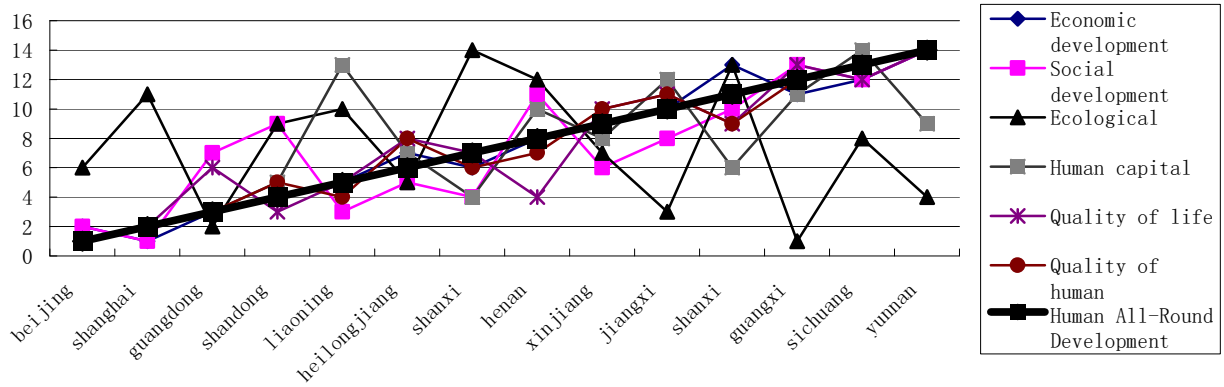


Fig.1 the scores and curve of 14 Provinces

Conclusion

In this paper, we through GDP per capita growth, the Gini coefficient, and other indicators that can provide data to calculate Gross National Happiness, to reflect the comprehensive development of the process of national sense of happiness. This paper has solved that Chinese Human All-Round Development evaluation does not reflect the scientific concept of development issues at the present stage because of lacking of Gross National Happiness. we use Standard Deviation and G1 method to determine the weights of indicators, and analyzed the Human All-Round Development of 14 provinces.

Acknowledgment

The research is supported by the National Social Natural Science Foundation of China as Major Program (the grant number is 06&ZD039), and is Supported by Social Sciences Foundation of Qinhuangdao(20101062) and is supported by the Dalian University of Technology as Major Program of Social Science Foundation (the grant number is DUTHS2007101). We thank both the National Social Natural Science Foundation of China and the Dalian University of Technology for their financial support.

References

[1] B. Schmitt. Conceptual Framework and Structure of an European System of Social Indicators[J]. *European Reporting Working Paper No.9. Social Indicators Department*, 2000.

[2] United Nations Development Programme(UNDP).Human Development Reports[R]. *New York: Oxford University press*,2002:141-162.

[3] World Bank. Living Standards Measurement Study of the Word Bank[M]. *Washington: World Bank press,*1980:271-305.

[4] L. Peilin, Zhu Qingfang. Task brief introduction: 《The evaluation index research of Chinese wll-off society》[N/OL].Economic information Daily, 2005.10.08[2008.03.20]. http://202.84.17.25/www/Article/200510810425-1.shtml.

[5] K.Mazumdar Measuring the Well-being of the Developing Countries:Achievement and Improvement Indices[J]. *Social Indicator Research*,1999(47):1-60.

[6] D. Xinling, Zhao Yanyun.The index system and Comprehensive Evaluation of the well-off society[J] *Statistics and Decision*,2007,(12):61-64.

[7] Z. Shunquan. A Study of Living-Standard Appraisal for Chinese Urban Residents[J].*Journal of Shanxi Finance and Economics University*,2004,26(3):35-37.

[8] Y. Wenzhou. The Statistic and evaluation index system of the well-off society[J].*Statistics and Decision*,2005,(7):28-29.

[9] Y. Pengfei, WANG Lina. The choice of sample data dimension less methods in Financial Early Warning [J]. *Friends of Accounting,* 2005,(12):43-44.

[10] W. Kun, Song Haizhou.The compare of three objective weights[J]. *Technoeconomics & Management Research*, 2003,(3):18-19.

[11] G. Yajun.The theory and methods of comprehensive evaluation [M].*Beijing, Science Press*, 2002.7-12.

[12] National bureau of statistic of China. China statistical year book.http://www.stats.gov.cn/tjsj/ndsj/. May 20th, 2010.

doi:10.4028/www.scientific.net/AMM.50-51.828

Modal Parameters Identification Based on Noise Cancellation from Measured Vibration Response Signals

Xingxian Bao [1, a] and Cuilin Li [2, b]

[1]Dept. of Marine Engineering and Fluid Mechanics, College of Petroleum Engineering, China University of Petroleum (East China), Qingdao, China

[2]Dept. of Ocean Engineering, College of Engineering, Ocean University of China, Qingdao, China

[a]baoxingxian@yahoo.com.cn, [b]lynnouc@gmail.com

Keywords: Modal paramenter identification, noise cancellation, singular value decomposition, vibration measuring, impulsive response function.

Abstract. Measured vibration response signals are inevitably contaminated with noise when a data acquisition system is used for an experimental measurement. This situation often leads to serious difficulties in identifying the modal parameters with proper accuracy. This paper presents a noise cancellation method for measured impulsive response functions (IRFs) based on structured low rank approximation (SLRA) so as to improve the accuracy of the modal parameters identification. Numerical study of a cantilever beam model is carried out to demonstrate the performance of the proposed method. The results show that this method can remove noise from measured IRFs efficiently, and the modal parameter identifications based on the filtered IRFs are very good.

Introduction

It is extremely desirable to acquire high-quality data from vibration tests for modal parameters identification. However, the measured vibration response signals are inevitably contaminated with noise when a data acquisition system is used for an experimental measurement. This situation often leads to serious difficulties in identifying the modal parameters with proper accuracy.

In the modern modal identification analysis, it does not seem to have a particular procedure intending to eliminating the noise. In fact, the “noise modes” are often absorbed and then other techniques are applied to distinguish the real modes and computational modes [1-3]. It is difficult to identify the real modes, especially when the signal-noise ratio is low.

This article proposed a procedure to remove much noise from the measured impulsive response function (IRF) before applying a time domain technique for modal parameters identification. For removing noise from the measured IRF to yield a filtered IRF, a structured low rank approximation (SLRA) method for the Hankel matrix is carried out, which has been published, mainly in the journals related to signal processing and numerical linear algebra [4-8]. Cadzow’s algorithm [9] is to be employed in this study to remove the noise. Once the filtered IRF is obtained, we can apply the complex exponential method to extract the modal parameters [1,2].

A finite element model of cantilever beam will be used to demonstrate the performance, and illustrate the procedure as well, of the proposed scheme in the application studies.

Structured Low Rank Approximation

Let $\{h_\ell\}$, $\ell = 0,\ldots,s$, represent an IRF (or a segment of an IRF) sequence of an N-degree of fredom (N-DOF) system, and be sequentially filled to form a rectangular Hankel matrix, $\mathbf{H}_{m\times n} \in \mathbb{R}^{m\times n}$,with $m,n \geq 2N$ as

$$\mathbf{H}_{m\times n} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \cdots & h_n \\ \vdots & \vdots & \ddots & \vdots \\ h_{m-1} & h_m & \cdots & h_s \end{bmatrix} \tag{1}$$

where $s = m + n - 2$.

Singular Value Decomposition

Typically, the rank of a matrix can be estimated by a classical method based on the singular value decomposition (SVD) of the matrix. Let the SVD of a matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ be expressed as:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{T}} \tag{2}$$

where $\mathbf{U} \in \mathbb{R}^{m\times m}$ and $\mathbf{V}^{\mathrm{T}} \in \mathbb{R}^{n\times n}$ are orthonormal matrices, and $\mathbf{\Sigma} \in \mathbb{R}^{m\times n}$ is a rectangular diagonal matrix whose diagonal elements are singular values, arranged in order of deceasing size. Because some of the singular values may be zero, we therefore partition $\mathbf{\Sigma}$ into a submatrix $\mathbf{\Sigma}_r$ of r nonzero singular values and several zero matrices as

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_r & 0 \\ 0 & 0 \end{bmatrix} \tag{3}$$

where $\mathbf{\Sigma}_r$ is a $r \times r$ diagonal matrix. This decomposition indicates the rank of $\mathbf{A}$ is r. For theoretical data, the singular values should go to zero when the rank of the matrix is exceeded. For measured data, however, due to random errors and small inconsistencies in the data, the singular values will not become zero but will become very small.

Cadzow's Algorithm

In mathematical terms, the noise cancellation problem here is a structured low rank approximation (SLRA) problem. A simple engineering algorithm is the iterative Cadzow's algorithm [9,10], which can be summarized by the following steps:

1. Use the singular value decomposition technique with an appropriate value of rank r estimated from the data to obtain a low rank approximation to the Hankel data matrix. Note that the resulting matrix $\hat{\mathbf{A}}$ will not be a Hankel matrix.

2. Rebuild a Hankel matrix $\hat{\mathbf{H}}$ from $\hat{\mathbf{A}}$ by replacing all elements of each anti-subdiagonal by the arithmetic average of the elements along the anti-subdiagonal. Note that the resulting Hankel matrix $\hat{\mathbf{H}}$ will not be rank r.

In Cadzow's algorithm, the low rank approximation and anti-subdiagonal averaging are alternated iteratively until a convergence test has been met. It has been proven that such iterative usage always converges [10].

Numerical Study

The performance of the proposed method is investigated by a cantilever beam modelled by ANSYS commercial software. The length of the steel beam is 1.5m, and the section is 57×25mm^2. The first two natural frequencies of this finite element model are 8.3512 and 52.319Hz, respectively. Whereas a proportional damping matrix $\mathbf{C} = \alpha\mathbf{M} + \beta\mathbf{K}$ is taken, the damping ratio ξ_r of the rth mode is related to the rth modal frequency ω_r (in rad/sec) and coefficients α and β as:

$$\xi_r = \frac{\alpha}{2\omega_r} + \frac{\beta\omega_r}{2} \tag{4}$$

Targeting ξ_1 and ξ_2 equal to 0.002 and 0.004, from the computed ω_1 and ω_2, we should set $\alpha =$ 0.14662 and $\beta = 2.2979\times10^{-5}$. Exciting the beam by an impulsive load on the free end, we recorded the corresponding acceleration signals with sampling rate 200 Hz. Without any loss of generality, only a segment about 5 seconds with 1024 sample points of the free end would be taken for the later analysis. Corrupted IRF is referred to as "measured" IRF in this section. It is generated by adding a Gaussian white noise to the exact (noise-free) IRF. The level of the additive white noise is quantified by a stated percentage, defined as the ratio of the standard deviation of the white noise to that of the exact IRF. Shown in Fig. 1 is the exact related frequency response function (FRF) and that with 3% white noise signal.

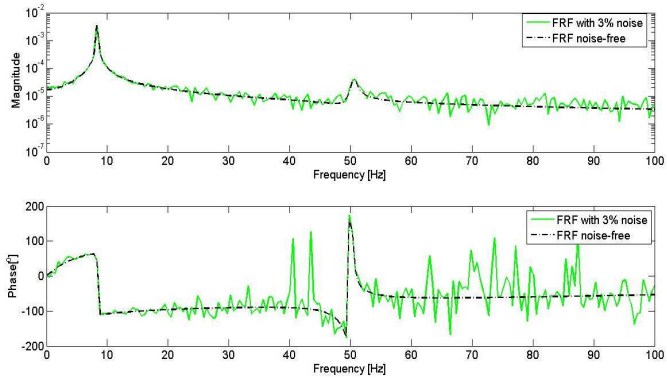


Figure 1: The corrupted and noise-free FRFs

Rank Determination

In SLRA method, we should know the appropriate value of rank of the Hankel matrix. That could be easily done if the singular values of the matrix have been ordered sequentially from the largest to the smallest, by plotting the normalised singular values and choosing the appropriate rank when the

normalised singular values approach an asymptote [3]. In this section, the 1024-point signal is referred to as the measured IRF, from which the corresponding 513×512 largest (nearly square) Hankel matrix could be produced. The normalized singular values of the Hankel matricx is shown in Fig. 2. Because the fifth normalized singular value drops to a much smaller value, the noise threshold, i.e. the effective rank of the Hankel data matrix is set to 4.

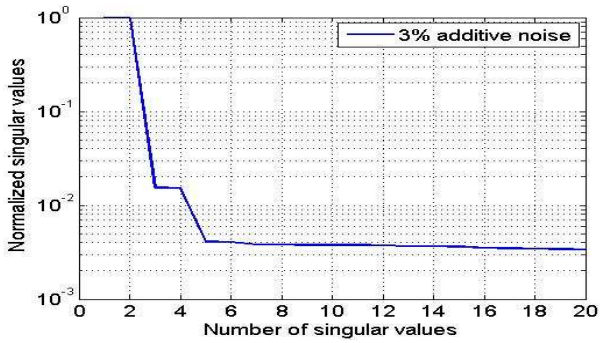


Figure 2: Normalized singular values of the Hankel matrix.

Noise Cancellation

Cadzow's algorithm alternates iteratively between two steps: (unstructured) low rank approximation and anti-subdiagonal averaging. Working on the Hankel matrix based on rank reduction 4, we obtain an excellent result of the noise removal.

The evaluation of the performance of the noise cancellation is done by comparing exact and filterd FRFs (see Fig. 3). The filtered FRF has 2 sharp peaks within the frequency range 0-100 Hz, and the visible level of noise at the measured FRF has been successfully removed.

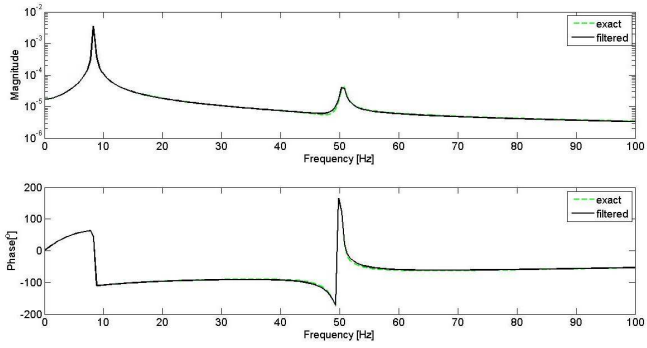


Figure 3: Performance of noise cancellation of the measured signal.

Modal Parameters Identification

After carrying out the noise cancellation to obtain the filtered IRFs, we apply complex exponential method for estimating modal parameters. The first two modal frequencies and damping ratios estimated from the exact, measured and filtered signals are listed in Table 1, where the values of finite element (FE) model are also included for comparison. The two modal frequencies and damping ratios obtained from the exact data basically agree with those of the FE model, recognizing that discrepancy always exists between the true modal parameter and its estimated from truncated signals.

As exhibited in Table 1, the two order modal parameters estimated from the filtered IRFs are in excellent agreement with that from the exact IRFs, respectively. However, the modal frequencies

and damping ratios estimated from the measured IRF are worse, and the second modal parameters are not identified due to the noise.

Table 1: The first two modal frequencies and damping ratios estimated from different signal

Mode	FEM		Exact		Measured		Filtered	
	f_t	ξ_t	f_e	ξ_e	f_m	ξ_m	f_f	ξ_f
1	8.3512	0.002	8.3435	0.0022581	9.5903	0.0010498	8.3431	0.0022008
2	52.319	0.004	52.548	0.0043376	——	——	52.494	0.0045397

Concluding Remarks

Due to the noise in measured vibration response signals, accurately identifying modal parameters has been a challenging task. This paper presents a noise cancellation procedure for IRFs based on structured low rank approximation so as to improve the accuracy of the modal identification.

A numerical model of cantilever beam, measured IRF with 3% noise was investigated. We could accurately identify the modal parameters (frequencies and damping ratios) of the first two order modes by using the proposed method. We observed significant improvement in modal parameters identification when the filtered, rather than measured, IRF was used.

Reference

[1] D.J. Ewins: *Modal Testing: Theory, Practice and Applications*, (Research Studies Press, Baldock, Hertfordshire, England, 2000).

[2] N. Maia, J. Silva, J. He, N. Lieven, R.-M. Lin, G. Skingle, et al.: *Theoretical and Experimental Modal Analysis*, (Research Studies Press, Taunton, Somerset, England, 1997).

[3] R.J. Allemang and D.L. Brown: A unified matrix polynomial approach to modal identification, Journal of Sound and Vibration, vol. 211 (1998), p. 301–322.

[4] B. De Moor: Structured total least squares and L2 approximation problems, Linear Algebra and its Applications (1993), p. 188-189: 163–207.

[5] I. Markovsky and S.V. Huffel: Overview of total least squares methods, Signal Processing, vol. 87 (2007), p. 2283–2302.

[6] I. Markovsky: Structured low-rank approximation and its applications, Automatica, vol. 44 (2008), p. 891–909.

[7] M. Schuermans, P. Lemmerling and S. V. Huffel: Block-row hankel weighted low rank approximation, Numerical Linear Algebra with Applications, vol. 13 (2006), p. 293–302.

[8] M. T. Chu, R. E. Funderlic and R. J. Plemmons: Structured low rank approximation, Linear Algebra and its Applications, vol. 366 (2003), p. 157–172.

[9] J.A. Cadzow: Signal enhancement—A composite property mapping algorithm, IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 36 (1988), p. 49–62.

[10] D. Tufts and A. Shah: Estimation of a signal waveform from noisy data using low-rank approximation to a data matrix, IEEE Transactions on Signal Processing, vol. 41 (1993), p. 1716–1721.

doi:10.4028/www.scientific.net/AMM.50-51.833

A New Method for Predicting Unsteady Performance of Integrative Propulsion system

M . He [1, a],S . Huang[1]

1. College of Shipbuilding Engineering, Harbin Engineering University, Harbin 150001, China
[a]miaomiao591213@yahoo.com.cn

Keywords:integrative propulsion system;unsteady hydrodynamic performance; integral solution; influence coefficiets

Abstract. In order to investigate unsteady hydrodynamic performance of integrative propulsion system, integral solution based on panel method is used to calculate the system. Interaction between propeller and rudder is considered by recalculating influence coefficents. Influence of propeller wake to rudder is studied, and then propulsion performance of integrative propulsion system is calculated in steady flow. Calculation results is compared with steady value and indicates that it's feasible to use this method for predicting the unsteady performance of integrative propulsion system.

Introduction

Integrative propulsion system is a new type of energy-saving propulsive system. The rudder-ball with plat head properly connecte to hub. The flap rudder used in the system can manipulate the hull by swing flap. For rudder is closely behind propeller, interaction between them should not be ignored. And the system is in the ship wake which has great influence in fluctuating pressure, noises and cavitation performance which are usually predict ed by unsteady calculation.

Potential based panel method is wildly used to solve unsteady hydrodynamic performance problems of propeller-rudder systems[1-6]. Usually interaction between them are considered by induced velocity from one to the other, and convergent results are obtained by iteration in time domain. Xie Xuesen(2009)[7] developed integral solution method to predicet performance of propeller-pod system, and the interaction between propeller and pod is considered by influence coeffieints variation. In his work, interaction is also examined by influence coeffieints, and the unsteady performance of integrative propulsion is gained using improved integral solution method.

Prediction method

Numerical formulation

Fluid is assumed to be incompressible and inviscid. The flow is considered to be irrotational. Propeller rotate with a constant angular velocity $\boldsymbol{\omega}$, calculation begins at time t, and the time step is Δt ,so the rotation angle can be expressed as $\Delta\theta = \omega\Delta t$. It assumed that at original time $t_0 = 0$, blades are at angle $\theta = 0$. AT time step k_t ,the time can be expressed as $\Delta\theta = \omega\Delta t$, and the panel coordinate rotate $k_t\Delta\theta$.

The perturbation velocity potential $\varphi(t)$,normal dipole jump $\Delta\varphi(t)$ and $[\boldsymbol{V}_0(t)\cdot\boldsymbol{n}]$ are assumed to be equally distributed on every panel surface and control points are at centroid. At k_t ,the linear

system of algebrac equations for the unkown φ_i as shown:

$$\sum_{j=1}^{N_P+N_R}(\delta_{ij}-C_{ij}(k_t\Delta t))\varphi_j(k_t\Delta t)-\sum_{m=1}^{N_r}W_{im1}(k_t\Delta t)\Delta\varphi_{m1}(k\Delta t)=-\sum_{j=1}^{N_P+N_R}B_{ij}(k_t\Delta t)[\boldsymbol{V}_j(k_t\Delta t)\cdot\boldsymbol{n}_j]+$$
$$\sum_{m=1}^{N_r}\sum_{l=2}^{N_w}W_{iml}(k\Delta t)\Delta\varphi_{ml}(k\Delta t) \qquad i=1,2,\cdots,N_P+N_R \tag{1}$$

Where, N_P and N_R are total panel number of propeller and rudder. N_r is wake strips number behind training edge of all propeller blades and rudder, and N_w is the number of panels on one strip. $C_{ij}(k_t\Delta t)$, $B_{ij}(k_t\Delta t)$ and $W_{iml}(k_t\Delta t)$ are influence coefficients, and should be recalculated every time step.

The force and torque can be obtained by integral on body surface. The total thrust and torque coefficients of propeller and rudder are expressed as follows:

$$K_t=\frac{T}{\rho n^2D^4},\quad K_q=\frac{Q}{\rho n^2D^5},\quad J=\frac{V_0}{nD},\quad \eta_0=\frac{J}{2\pi}\frac{K_t}{K_q}$$

Changing rules of influence coefficients

In the unsteady case, the propeller rotate while the rudder keep its position. It is necessary to recalculate the influence coefficientes between propeller and rudder. The influence coefficientes related to geometric position, so panels of propeller surface and wake should be rotated before calculating influence coefficients which are defined by Morino's analytic formula. For the same rules of $C_{ij}(k\Delta t)$ and $B_{ij}(k\Delta t)$, only the changing rules of $C_{ij}(k\Delta t)$ are listed here:

1) If $i\le N_P$ and $j\le N_P$, then $C_{ij}(k_t\Delta t)=C_{ij}(\Delta t)$, it presents influence coefficients from propeller panels to propeller panels. They keep the same for relative geometric position of its panels does not change.

2) If $i\le N_P$ and $j\ge N_P+1$, then

$$C_{ij}(k_t\Delta t)=\begin{cases}C_{ij}(k_t\Delta t) & k_t<T\\ C_{ij}(k_t\Delta t-T) & k_t\ge T\end{cases}$$

It presents influence coefficients from rudder panels to propeller panels. The value changes in one circle, and shows periodicity when the propeller rotates more than one circle.

3) If $i\ge N_P+1$ and $j\le N_P$, then

$$C_{ij}(k_t\Delta t)=\begin{cases}C_{ij}(k_t\Delta t) & k_t<T\\ C_{ij}(k_t\Delta t-T) & k_t\ge T\end{cases}$$

It presents influence coefficients from propeller panels to rudder panels and has periodicity, too.

4) If $i\ge N_P+1$ and $j\ge N_P+1$, then $C_{ij}(k_t\Delta t)=C_{ij}(\Delta t)$

It presents influence coefficients from rudder panels to rudder panels.

Unsteady hydrodynamic performance of integrative propulsion system in uniform flow

According to theory above, program is combined using *FORTRAN*. By altering relative parameter, it is easy to calcultate unsteady hydrodynamic performance of single propeller working in nonuniform flow and propeller and rudder system working in uniform and nonuniform flow.

Influence of propeller wake to rudder

Singular points exsit at the place where propeller wake panels cross rudder panels. Before calculation, those singular points should be discussed.

For great demand of computation time for propeller with 5 blades, propeller with 3 blades is investigated. While calculating influence coeffients W_{ij}, singular points exist at the control points of propeller wake infinitely close to the control points of rudder panels,as shown in Fig.1. And the problem leads to distribution of dipole strength on propeller panels incorrect. The Dipole intensity variation of a panel at different angular position is shown in Fig.2.

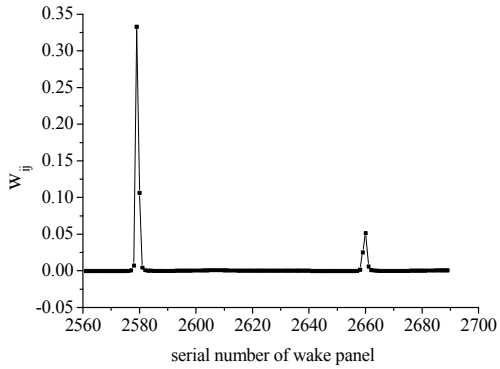


Fig.1 Propeller wake influence to rudder

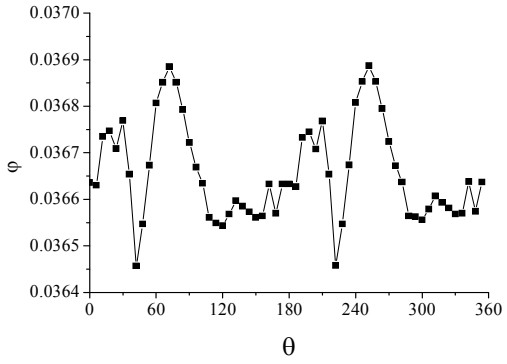


Fig.2 Dipole intensity variation of panel at different angular position

The paper adopted two methods. Firstly, the singular points are shielded. The influence coefficients of propeller wake return to 0 at singular points. Fig.3 shows dipole intensity distribution of the panel which is the same as the panel in Fig.2. Secondly, the influence coefficients of sigular points are spline interpolated. The dipole strength is shown in Fig.4.

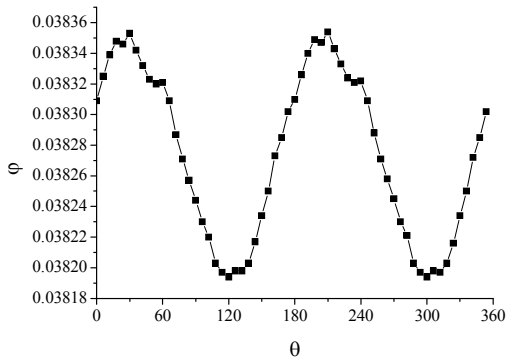


Fig.3 Dipole intensity variation of panel after shielding

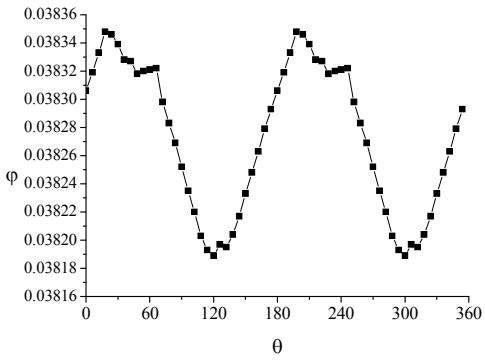


Fig.4 Dipole intensity variation of panel after interpolating

The results indicated that both methods make the dipole intensity curve fairing, but there are still problem in crest of curves for the methods are not accurate numerical treatment.

Unsteady hydrodynamic performance of integral propulsion system in uniform flow

In this section, unsteady hydrodynamic performance of integrative propulsion system in

uniform flow is calculated. The propeller and rudder model is in reference[8].

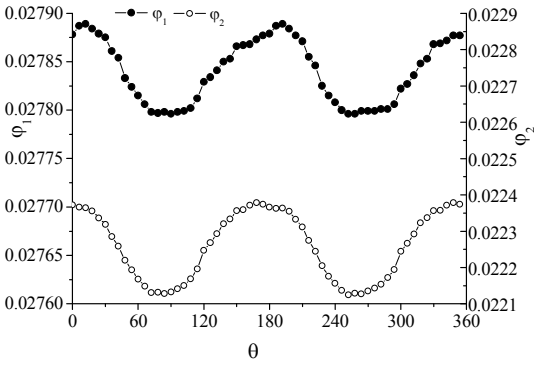


Fig.5 Dipole intensity variation of panel at different angular positons

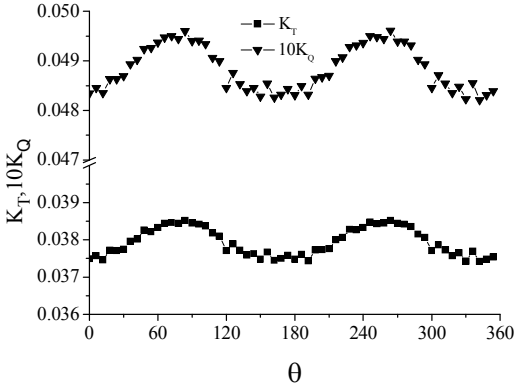


Fig.6Axial force and moment of NO.1 blade at different angular position

The dipole intensity variation of panel at different angular positions is shown in Fig.5. For symmetry of the rudder, there are two peak in a circle. The dipole strength curve is not so smooth for the singular points of propeller wake to rudder discussed before. For the integrative propulsion system, it is important to consider the effect of rudder to propeller. And because of the existence of rudder, the propeller working in a nonuniform wake field. As shown in Fig.6, axial force and moment of a single blade has two periods which illuminated that the propeller worked in a symmetrical flow field and the symmetry face is the vertical profile of the rudder. As is known, thrust and torque of blade in high wake region is larger comparing with blade working in low wake region. By studying the curves in Fig.6, it can be concluded that low wake aera is right in front of the leading edge of rudder.

Then average values of unsteady calculation are compared with steady values which were validated by experments[8]. The results indicate that it's feasible to use this method for predicting unsteady performance of integrative propulsion system.

Table 1 Comparasion results of steady and unsteady calculate

	Steady cal. value	Average value of unsteady cal.	error（%）
Kt	0.18969	0.19263	1.526
10kq	0.24426	0.2519	3.035

Conclusion

A computational method has been used to predict the hydrodynamic performace of the integral propulsion system. All panels on propeller and rudder are participated in calculation, and interference between propeller and rudder are calculted by variation of influence coefficents. We can conclude from the results that this method can predict unsteady performace of integrative propulsion system well. Singlular points is important in this method, and author's next work is focusing on numerical treatment of them.

References

[1] WANG Guoqiang, LIU Xiaolong. A potential based panel method for prediction of steady and unsteady performances of ducted propeller with stators[J]. *Journal of Ship Mechanics*,2007,11;333-334

[2] Jiang Shaojian. Hydrodynamic performance of propeller-rudder interaction. Ph. D. *Thesis,*

Shanghai jiaotong University.Vol:97-100, 1993

[3] Hassan Ghassemi,Parviz Ghadimi. Computational hydrodynamic analysis of the propeller-rudder and the AZIPOD systems[J]. *Ocean Engineering*,2008,35;123-125

[4] Liu Xiao-long. A Potential Based Panel Method for Prediction of Steady and Unsteady Performances of Contra-rotating Propellers. *First International Symposium on Marine Propulsors smp'09, Trondheim, Norway,* June 2009

[5] YE Jin-ming, XIONG Ying. Prediction of podded propeller cavitation using an unsteady surface panel method.*Journal of hydrodynamics*. 20(6), 790-796, 2008

[6] Spyros A. Kinnas, HanSeong Lee, and Yin L. Young Modeling of Unsteady Sheet Cavitation on Marine Propeller Blades. *International Journal of Rotating Machinery*, 9: 263–277, 2003

[7] Xie Xuesen.Numerical simulation of propulsion and cavitation performanc of podded propusor. Ph. D. Thesis, *Harbin Engineering University.* Vol:83-85,2009

[8] Li xin. A Hydrodynamic Analysis on Energy-Saving Rudder Ball Behind Propeller. *Master's Thesis, Harbin Engneering University*. Vol:38.2009.

doi:10.4028/www.scientific.net/AMM.50-51.838

Influence of Absorber's Stiffness and Installation Site on Bit Trajectory

X.H. Zhu[1, a], Y.J. Jia[2, b], J.H. Li[3, c]

[1,2,3]Mechatronic Engineering College of Southwest Petroleum University, China

[a]zxhth113@163.com, [b]jiayj007@163.com, [c]lijinhe1996@126.com

Keywords: Drillstring vibration; Absorber; Stiffness; Installation site; Well path

Abstract. A model of drillstring system dynamics was derived based on the drillstring's Dynamics Equation, System Dynamics Theory and the optimization of absorber's stiffness and installation site. A numerical calculation approach for solving drillstring system and bit movement was presented according to the drilling parameters. And this paper employed the simulation calculations to analyze and optimize absorber's stiffness and installation site on different conditions. The results show that the absorber's stiffness has significant impact on the movement frequencies of bit in the fourth quadrant when the bit next to the borehole wall drills ahead, and absorber has preferable ability of controlling deflection under stiffness for 3950*kN/m* and 6200*kN/m*; while the installation site of absorber has acute influence on the bit movement in the first quadrant, the third quadrant and the fourth quadrant, and the optimum position of the absorber is 0.6*m* distant from the bit. Parameters of absorber's stiffness and installation site recommended in this study could improve the effectiveness of penetration and are good for controlling deflection and well path.

Introduction

Parameters of absorber's stiffness and installation site have significant impact on the bit trajectory owing to the impact on the drillstring vibration performances. Scholars studied the parameters impacted on bit trajectory, considering the longitudinal vibration, the transverse vibration and the torsional vibration separately and using the differential equation analytics[1], the transfer function method[2], the finite element method(FEM)[3,4] and the elastic wave method[5]. Absorber's damping coefficient is closely related with the deformation of upper drillstring[6]. Therefore it should consider the deformation and movement of the whole drillstring system in real well to put forward more serviceable conclusions

This paper derives the dynamics model of slick assembly with absorber under the condition of the real well parameters, which is based on the previous studies. And an approach for simulating calculation is studied using the ABAQUS software. The quantitative evaluation is finished, focusing on the point of view that the bit's ability of controlling deflection under the condition of transverse vibration, on different conditions of absorber's stiffnesses and installation sites. This study has significance to predict the well path and optimizes the drilling assembly.

Dynamics model of the drillstring in deep vertical well

The drillstring vibration system is simplified to multiple-degree of freedom system using the Spring-Mass-Damping(S-M-D) method [7, 8].This model could easily consider the drilling fluid factor, wall boundary and the top-and-bottom conditions.

The hypotheses as follows are introduced according to the structure characters of the whole drillstring system: (1) the drillstring system is inhomogeneous elastic beam, and the drillstring parameters such as geometry size, material properties are constant. The threads, local slots and effects of temperature is neglected; (2) the weight on bit (WOB) is the superposition of steady and transient harmonic load, and it follws the sine harmonic regular; (3) absorber is counted as Flexibility-Damping unit; (4) the borehole is always circular cross section, and the damping of mud and friction between the drillstring and wall are considered.

The dynamic balance equation of the whole drillstring system is established based on the fundamental principle of the rotor dynamics [9, 10]. The axis of the borehole is in homogeneously discrete for several units based on the drill string geometric shape and the two-nodes-three-dimension elastic beam element. The unit dynamic balance equation is shown in Eq. (1).

$$[M^e]\ddot{u}_e(t)+[C^e]\dot{u}_e(t)+[K^e]u_e(t)=[P^e(t)] \tag{1}$$

The stiffness matrix in local coordinate system is shown in Eq. (2).

$$[K^e]=[K_E^e]+[K_G^e] \tag{2}$$

Rayleigh damping method can solve the problem of the friction between drillstring and wall as well as the viscous force of mud. It is shown in Eq. (3).

$$[C^e]=\alpha[M^e]+\beta[K^e] \tag{3}$$

The influence could be approximately simulated under the conditions mentioned above by selecting the proper value of α and β [11].

The unit load is shown in Eq. (4)

$$[P^e(t)]=[P^e]+\mu[P^e]\sin(\omega t)+[K]\delta^e(t) \tag{4}$$

The n-degree-freedom totality dynamic balance equation, which is shown in Eq. (5), of the whole drillstring is derived by compositing all of the units' dynamic balance equation.

$$[M]\ddot{u}(t)+[C]\dot{u}(t)+[K]u(t)=[P(t)] \tag{5}$$

The dynamic characters of the whole drillstring system could be obtained depending on solving the Eq. (5) under the given boundary conditions.

Finite element Model and its numerical calculation

The drillstring system diagram is drawn using develop tools of CAD based on the real well conditions and parameters of drilling. In order to keep coincidence with real drilling, the simulation model is treated as Mass-Spring-Damping.

The real drilling assembly of one Chuandong oil field is as follows: Φ241.3*mm* bit + back pressure valve + (Φ177.8*mm* absorber) + one Φ178*mm* NMDC+ twenty Φ177.8*mm* DCs + JZ-T178 bumper + bypass valve + twelve Φ139.7*mm* HWDPs+ Φ139.7*mm* DPs + kelly stem

The nonlinear finite element models of slick assembly and upper drillstring are established in Cartesian coordinates. The drillstring 3-D model is shown in Fig.1.

The drilling parameters are as follows: $2.2\times10^5 N$ WOB,77*r/min* speed on bit(SOB) ,$2.0\times10^3 kg/m^3$ mud density,60*s* viscosity. The parameters of borehole are shown in Tab.1.There are 30 cases shown in Tab.2 as a result of different parameters of absorber's stiffness and placement.

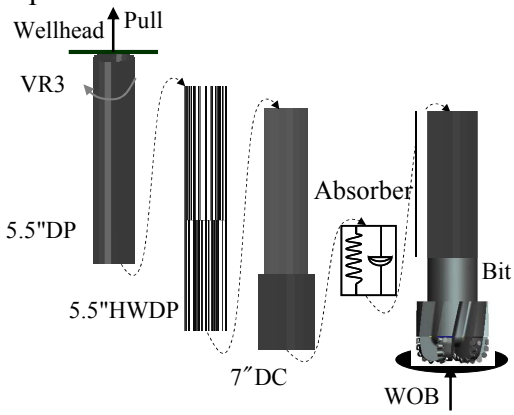


Fig.1. the drillstring of slick assembly entity 3-D model

Tab.1. Parameters of borehole

Hole section(*m*)		Hole depth(*m*)	Diameter of bit(*mm*)	Diameter of borehole(*mm*)	
begin	end				
0.00	2053	2053	316.5	Casing pipe	273.1
2053	4693	2640	241.3	Bare hole	275.6

Tab.2. Parameters of absorber's stiffness and placement

Parameters	1	2	3	4	5	6
Stiffness (kN/m)	3200	3950	4700	5450	6200	--
Site from bit(m)	0.6	5	10	18	27	40

Results and discussion

The collision time is 25*ms* which is obtained from the calculation of the coefficients of restitution between drillstring and rock. In this study, there are 4000 outputs over 100*s* on the condition that the step size is 15*ms*.

There are four quadrants in the range of bit's movement, and the direction $+y$ ($+U2$) is the magnetic north. The results show that the bit is approximately next to the wall in drilling, which led bit to appear in each of the four quadrants at different probabilities, thus it affects the cutting of rock and extends to impact the well path.

Parameters of absorber's stiffness and placement have significant influence on bit movement, as are shown in Fig.2-Fig.6.

The influence of absorber's stiffness on bit trajectory. The movement of the bit under different parameters of absorber's stiffness (the installation site from bit is 0.6*m*) is shown in Fig.2 and Fig.3.The frequencies of bit's appearance in the center of the borehole and of the transverse motion are awfully low when the parameter of absorber's stiffness is 3950*kN/m*; the frequencies of bit transverse motion is high when the parameter of absorber's stiffness is 4700*kN/m*.The frequencies of bit's appearance in the second quadrant under the two conditions are clearly different, and the later is lower.

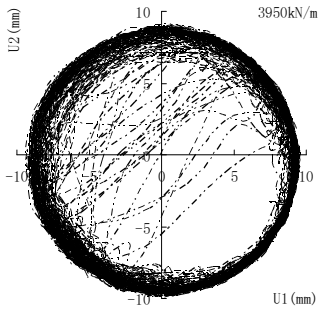
The curve of bit vibration under five different parameters of absorber's stiffness is shown in Fig.5.The frequencies of bit movement as well as the fluctuation range are higher in the fourth quadrant than in others. The frequency difference reaches the maximum in the second and the fourth quadrants when the stiffness is 4700*kN/m;* and the frequency difference reaches the minimum when the parameters of absorber's stiffness are 3950*kN/m* and 6200*kN/m*, under which the bit cuts the rock uniformly. Consequently, both the two parameters, 3950*kN/m* and the *6200kN/m*, are recommended in the slick assembly to control the well path.

The influence of absorber's placement on bit trajectory. The curve of bit vibration under different parameters of absorber's installation site, but the parameter of absorber's stiffness is 3950*kN/m*, is shown in Fig.2 and Fig.4. Comparing the two figures, the absorber's installation site has little relationship with the frequencies of bit movement in the second quadrant; while the frequency fluctuation range is greater in the first, the third and the fourth quadrant when absorber's site is 18*m* or 0.6*m* from the bit.

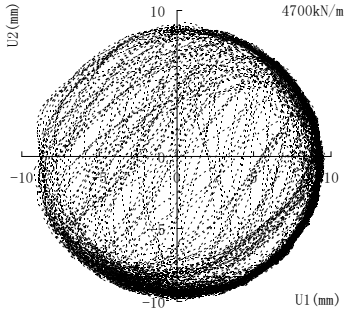
The influences of different parameters of absorber's placement on bit movement under the two parameters of absorber's stiffness, 3950*kN/m* and 6200*kN/m*, are shown in Fig.6 and Fig.7, respectively. Comparing the 0.6*m* of absorber's site from the bit with the other five conditions, under the former condition, bit cuts rock uniformly in the four quadrants under the two absorber's stiffness.

The frequencies of bit movement in the second and the third quadrants are low when the distance of absorber's placement from the bit is over 5*m*, as are shown in Fig.6, which goes against to the deflection control; the over-distance between the bit and the absorber's placement goes against to the deflection control ,as is shown in Fig.7.

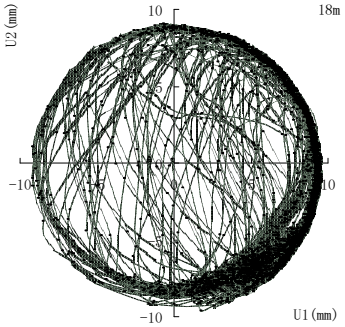
Comprehensively, the optimum parameters of absorber's stiffness are 3950kN/m or 6200kN/m, and placement from the bit is 0.6*m* under the goal of cutting uniformly and preventing deflection.



Stiffness: 3950kN/m; Distance: 0.6m
Fig.2. Bit trajectory of transverse movement



Stiffness: 4700kN/m; Distance: 0.6m
Fig.3. Bit trajectory of transverse movement



Stiffness: 4700kN/m; Distance:18m
Fig.4. Bit trajectory of transverse movement

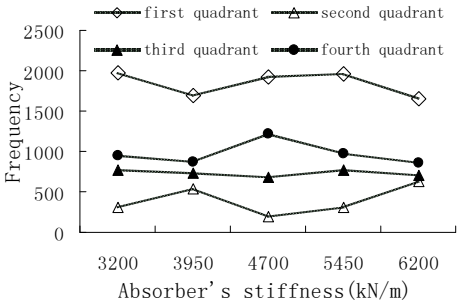


Fig.5. the influence of absorber's stiffness on bit trajectory (0.6m)

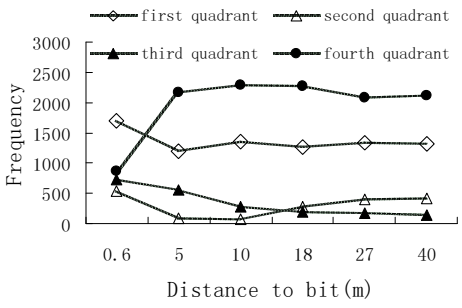


Fig.6. the influence of absorber's placement on bit trajectory (3950kN/m)

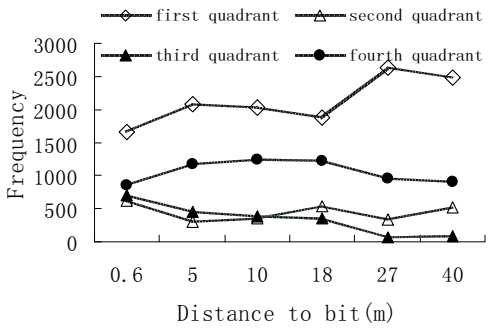


Fig.7. the influence of absorber's placement on bit trajectory (6200kN/m)

Conclusions

This paper presented the model of drillstring system dynamics and from the view that whether bit has the ability of cutting rock uniformly or not, the quantitative evaluation on bit trajectory has been finished under different absorber's stiffness and installation sites. The conclusions as follows:

(1) Parameter of absorber's stiffness has significant impact on the bit trajectory, thus the optimum of absorber's stiffness contribute to controlling the well path. The slick assembly possesses stronger ability of controlling deflection when the absorber's stiffness is 3950*kN/m* or 6200*kN/m*.

(2) Parameter of absorber's installation site has greater influence on the bit trajectory than the absorber's stiffness. Improper absorber's installation site lead dropping assembly to become building assembly. Comprehensively, it would be best for the slick assembly to install the absorber on the position of 0.6*m* distant from the bit.

Nomenclature

$[M^e]$ = mass matrix of element

$[C^e]$ = damping matrix of element

$[K^e]$= stiffness matrix of element

$[P^e(t)]$= disturbing force vector of element

(u_e), $(\dot{u}_e)$, $(\ddot{u}_e)$ = node displacement vector, velocity vector, acceleration vector of element, respectively

$[K_E^e]$ = element elastic stiffness matrix

$[K_G^e]$ = element geometric stiffness matrix

$[P^e]$ = permanent load of element

$\mu[P^e]\sin(\omega t)$ =wave force of element in bit on condition of SOB vibration frequencies [12, 13]

$[K]\delta^e(t)$ = instantaneous collision load of element

References

[1] Arrestad T V, Kyllingstad A: An Experimental and Theoretical Study of Coupled Boundary Conditions for Vibrations in Drillings. SPE 15563.

[2] Z.Y. Wang, M.T. Xu. Computing Forced Longitudinal Vibration of Drill String by Generalized Transfer Matrix Method [J].*Acta Petrolei Sinica*, 11(3):107~113(1990).

[3] Z.F. Li. Analysis of Longitudinal and Torsion Vibration of Drillstring[J]. *Engineering Mechanicas,* 21(6):203~210(2004).

[4] J.B. Liu. Calculation of The Optimal Placement location of Vibration Damper for Deep Hole Drill String [J]. NATURAL GAS INDUSTRY, 17(5):44~48(1997).

[5] Z.Z. Xie. Elastic Wave in a Drill-pipe String [J]. *Acta Petrolei Sinica*, 13(3):94~101(1992).

[6] Z.J. Cui. Analysis of Dynamic Characteristic of An Absorber with Considering The Influence of Upper Tubular Column Deformation [J]. *Journal Of Huazhong University Of Science And Technology.Nature Science*, 36(9):106~108(2008).

[7] X.H. Zhu, Q.Y. Liu, H. Tong. Research on Dynamics Model of Full Hole Drill-string System with Three-dimensional Trajectory [J].*Acta Petrolei Sinica*, 29(2):288~291(2008).

[8] Y.C. Kuang, D.K. Ma, Q.Y. Liu. Computer Simulation on Dynamic Action of Drillstring-Bit-Rock System [J]. *Acat Petrolei Sinica*, 22(3):81~85(2001).

[9] Dykstra Mark W. Nonlinear Drill String Dynamics. Ph. D. Dissertation, the University of Tulsa, (1996).

[10] C.C. Zhu, D.H. Feng, B. Lu. Nonlineaer Study on Dynamic Action of Integrated Drill String-well Rock System [J].*Journal Of Mechanical Engineering*, 43(5):145~149(2007,).

[11] J.L. Yao.Dynamic Characteristics and Fsilure Mechanism of Drill String in Straight Hole with Air Drilling [J].*Journal Of China University Of Petroleum. Nature Science*, 2008, 32(3):63~67(2008).

[12] X.H. Zhu,Qingyou Liu,Hua Tong.Emulation Research on Dynamics of Roller Cone Bit[J].*Acta Petrolei Sinica*, 25(4):96~100(2004).

[13] Y.S. Jiao, Z.X. Cai, Zongda Yan. Chaotic Motion of a Drillstring in Vertical Section [J].*Journal Of Vibration Engineering*, 13(2):314~318(2000).

[14] X.Z. Yan. Contact-impact Dynamic Simulation of Drillstring and Formation Based on Lagrangial Element Method [J]. *Journal Of Mechanical Strength,* 28(3):341~345(2006).

[15] G. H. Zhao, Z. Liang. Analysis Of Collision Between Drillstring And Well Sidewall. *International Journal Of Modern Physics B*. Vol. 22, Nos. 31 & 32 (2008) 5459~5464.

© (2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.843

Dynamic Analysis of Sandwich Beams Filled with ER Elastomers

Quan Bai [1,a], Kexiang Wei [1,b], Wenming Zhang[2,c]

1.Department of Mechanical Engineering, Hunan Institute of Engineering, Xiangtan,China
2 State Key Laboratory of Mechanical System and Vibration, Shanghai JiaoTong University, Shanghai, China
[a] e-mail：baiquan1119@163.com,[b] e-mail:wei_kx@163.com,[c] wenmingz@sjtu.edu.cn

Keywords：ER elastomers; sandwich beam; dynamic characteristics; vibration control

Abstract. Considered electrorheological (ER) elastomers as the visco-elasticity material, a finite element model of a sandwich beam filled with ER elastomers was developed based on Hamilton's principle and sandwich beam's theory. Then its dynamic characteristics were analyzed. Simulation results show that natural frequencies of the sandwich beam increase and vibration amplitudes of the beam decrease as the intensity of applied electric field increases. Increased the thickness of the ER elastomers layer, natural frequencies of the beam decrease and loss factors increase. Those indicate that the dynamic characteristic of ER elastomers sandwich beams is similar as that of ER fluids beam, which can be used for vibration control of flexible structures by applied a electric field.

Introduction

Electrorheological (ER) fluids is a kind of intelligent materials whose rheological property could be controlled by applied a electric filed [1]. The sandwich structures with ER fluids have a adaptive control capability of varying their damping and stiffness when changing the intensities of the electric field applied on them. Many papers about these sandwich structures have been published in the past years[2~5]. However, in the most of study, the ER fluids used are suspension liquid materials composed of basic liquids and disperse particals which are inevitable the problems of easily subside, the bad stability and seal difficulty for long use[6].

Electrorheological (ER) elastomers is a new branch of the electrorheological materials which is a solid analog of ER fluids. The polarizable solid particles and the elastomers are mixed and solidfied with applied a electric field. So the chains or fibrous structures are formed by solid particles in the two poles for ER effect [7]. The polarizable pellets are fixed in the substrate so that ER elastomers have the same controllablity, rapid responsibility and revesibility as ER fluids and the well stability and simple designing structure are both their unique advantages[8]. Therefore, ER elastomers have more potential in application for the vabration control of flexible structure than that of ER fluids. But now, the studies on ER elstomers applied in vibration conrol of flexible beams are rare.

In this paper, the controllable stiffness and damping characteristics of a sandwich beam filled with ER elastomers are studied. Consider ER elastomers as a viscoelastic damping material, a finite element model of the sandwich beam is developed based on Hamilton's principle and sandwich beam's theory. Then its dynamic characteristics for various applied elctric field and structures dimensions are numerically analyzed.

2 Finite element modeling of ER elastomers sandwich beam

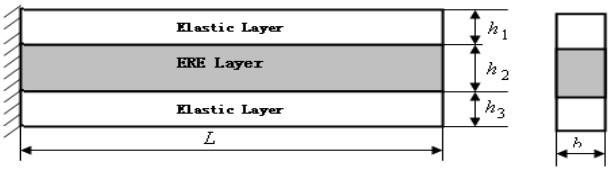


Fig. 1 The sandwich beam with an ER elastomers core and a constrained layer

Consider a three-layer sandwich beam as shown in Figure 1, in which an ER elastomers layer is restricted by two elastic layers. When applied different voltages to the elastic layers, it will prouduce an electric field in the area between these layers and consquently cause the physics properties of the core layer to be changed. The mathematical model of this structure is based on some following assumptions:(1) Considered electrorheological elastomers as the visco-elasticity material, the damping of structure is determined by the layer of ER elastomers.(2) No slipping between the elastic layers and ER elsatomers layer occurs.(3)The each layer of the sandwich beam shares the same longitudinal deformations.(4) It exists no normal stress in the ER elastomers layer, while it exists no shear strain in the elastic layer either.

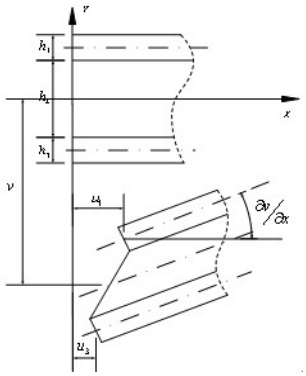


Fig.2 Undeformed and deformed configurations of a sandwich beam

As shown in Fig.2, the strain-displacement defromation can be expressed as

$$\frac{\partial u^i}{\partial x}=\varepsilon_i=\frac{\partial u_i}{\partial x}-z_i\frac{\partial^2 w}{\partial x^2},i=1,3 \tag{1}$$

where u^i is the displacement of layer i, u_i is the axial displacement of the mid-plane of layer i, and z_i is the distance of the mid-height of layer i from the reference plane.

The longitudinal normal force is negligible.

$$E_1A_1\frac{\partial u_1}{\partial x}+E_3A_3\frac{\partial u_3}{\partial x}=0\ ,u_3=-eu_1,\left(e=\frac{E_1A_1}{E_3A_3}=1\right) \tag{2}$$

From Fig. 2, the shear deformation of the core layer can be written as

$$\gamma=\frac{\partial w}{\partial x}+\frac{\partial u_2}{\partial z} \tag{3}$$

where w is the transverse displacement and u_2 is the axial displacement of the ER elastomers layer, and the following relationship among the u_1, u_3 and $\partial w/\partial x$ of the face-plate can be obtained

$$\frac{\partial u_2}{\partial z}=\frac{u_1-u_3}{h_2}+\frac{h_1+h_3}{2h_2}\frac{\partial w}{\partial x} \tag{4}$$

substituting the above equation into Equation (3), it can be obtained

$$\gamma = \frac{\partial w}{\partial x} + \frac{\partial u_2}{\partial z} = \frac{u_1 - u_3}{h_2} + \frac{h_1 + 2h_2 + h_3}{2h_2}\frac{\partial w}{\partial x} = \frac{u_1 - u_3}{h_2} + \frac{d}{h_2}\frac{\partial w}{\partial x} \tag{5}$$

where $d = h_2 + 1/2(h_1 + h_3)$.

In this study, a standard beam element with two nodes and four degrees of freedom per node, including the top longitudinal displacement u_1 , the base longitudinal displacement u_3 , the transverse displacement w and angular rotation of θ . The $q(t) = [w_1 \quad \theta_1 \quad u_{1i} \quad u_{3i} \quad w_2 \quad \theta_2 \quad u_{1j} \quad u_{3j}]$ is the vector of node desplacement.Thus, longitudinal and transverse displacements of the beam can be written in terms of a nodal displacement vector and a shape function as

$$\{\Lambda^{(e)}\} = \left[u_1 \quad u_2 \quad w \quad \frac{dw}{dx}\right]^T = [N_1 \quad N_2 \quad N_3 \quad N_4]^T \{q(t)\} \tag{6}$$

With

$$[N_1] = [1-\zeta \quad 0 \quad 0 \quad 0 \quad \zeta \quad 0 \quad 0 \quad 0]$$
$$[N_2] = [0 \quad 1-\zeta \quad 0 \quad 0 \quad 0 \quad \zeta \quad 0 \quad 0]$$
$$[N_3] = \left[1-3\zeta^2+2\zeta^3 \quad (\zeta - 2\zeta^2 + \zeta^3)l \quad 0 \quad 0 \quad 3\zeta^2 - 2\zeta^3 \quad (-\zeta^2 + \zeta^3)l \quad 0 \quad 0\right]$$
$$[N_4] = N_{3,x}$$

where $\zeta = x/l, l$ is the length of element, $[N_1], [N_2], [N_3]$ and $[N_4]$ are the finite element shape functions.

The elemental potential energy of the sandwich beam with the ER elastomers can be rewritten in terms of nodal displacement variables as follows

$$\begin{aligned} U &= \frac{1}{2}\int_0^l (E_1 A_1 u_{1,x}^2 + E_3 A_3 u_{3,x}^2)dx + \frac{1}{2}\int_0^l (E_1 I_1 + E_3 I_3) w_{,xx}^2 dx + \frac{1}{2}\int_0^l G_2 A_2 \gamma^2 dx \\ &= \frac{1}{2}\{q(t)\}^T ([K_{ei}] + [K_{bi}] + [K_{si}])\{q(t)\} \end{aligned} \tag{7}$$

With

$$[K_{ei}] = E_1 A_1 \int_0^l [N_{1,x}]^T [N_{1,x}]dx + E_3 A_3 \int_0^l [N_{2,x}]^T [N_{2,x}]dx$$
$$[K_{bi}] = (E_1 I_1 + E_3 I_3)\int_0^l [N_{3,xx}]^T [N_{3,xx}]dx$$
$$[K_{si}] = G_2 A_2 / h_2^2 \int_0^l [N_1 - N_2 + hN_{3,x}]^T [N_1 - N_2 + hN_{3,x}]dx$$

where E_1 and E_3 are the elastic modulus of the top and the base, G_2 is the complex modulus, A is the cross section area and and I is the inertia of the layer of the ER elastomers, $[K_{ei}]$ is the stretching deformation energy, $[K_{bi}]$ is flexural deformation energy and $[K_{si}]$ is shearing deformation energy.

The elemental kinetic energy of the sandwich beam includes the kinetic energys of the every layer associated with flexural deformational movement and stretching deformational movement is

$$T = \frac{1}{2}\int_0^l \left[\sum_{k=1}^3 \rho_k A_k \left(\dot{u}_k^2 + \dot{w}_k^2\right)\right]dx = \frac{1}{2}\left\{\dot{q}(t)\right\}^T [M_i]\left\{\dot{q}(t)\right\} \tag{8}$$

Where ρ_k is the material density of the k th layer and the above spots of $\dot{u}_k$, $\dot{w}_k$ and $\{\dot{q}(t)\}$ are denoted the derivation of t , $[M_i]$ is elemental mass matrix and can be expressed as

$$[M_i] = \int_0^l \sum_{k=1}^3 \rho_k h_k b [N_3]^T [N_3]dx + \rho_1 h_1 b \int_0^l [N_1]^T [N_1]dx + \rho_3 h_3 b \int_0^l [N_2]^T [N_2]dx + \frac{1}{4}\rho_2 h_2 b \int_0^l \left[N_1 + N_2 + \frac{h_1 - h_3}{2}N_{3,x}\right]^T \left[N_1 + N_2 + \frac{h_1 - h_3}{2}N_{3,x}\right]dx$$

Considering the situation of a sandwich beam element as $\{F\} = [K_{Fi}]\{q(t)\}$, in which $[K_{Fi}]$ is the stiffness matrix, the work caused by the force can be written as

$$W = \frac{1}{2}\{q(t)\}^T \{q(t)\} = \{q(t)\}^T [K_{Fi}]\{q(t)\} \tag{9}$$

According to the Hamilton's principle, the equation(7), the equation(8)and the euation(9) are substituted into $\delta\int_{t_1}^{t_2}(T - U + W)dt = 0$. Then the dynamic equation of k th the element can be obtained

$$[M^e]\{\ddot{q}(t)\} + [K^e]\{q(t)\} = 0 \tag{10}$$

where $[M^e]$ is element mass matrix, and $[K^e]=[K_{ei}]+[K_{bi}]+[K_{si}]+[K_{Fi}]$ is element stiffness matrix.

Assembling each element, the governing equations of the sandwich beam can be expressed as

$$[M]\{\ddot{q}(t)\}+[K]\{q(t)\}=0 \tag{11}$$

The dynamic charactristics of the beam can be got by the eigenvalue problem

$$(-\omega^2[M]+[K])\{q\}=\{0\} \tag{12}$$

Then the natural frequency and damping loss factor can be obtained by solving the eigenvalue of the above equation. where η is the natural frequency and ω is the damping loss factor.

$$\omega=|\omega|(1+i\eta) \qquad |\omega|=\sqrt{\mathrm{Re}\,\omega^2+\mathrm{Im}\,\omega^2} \qquad \eta=\mathrm{Im}\,\omega^2/\mathrm{Re}\,\omega^2 \tag{13}$$

Numerical example and discussion

This research decribes ER materials by a complex shear modulus, which was adopted by most other literatures[1~5].As presented in equation (14), in the pre-yield regime, the shear stress τ is related to shear strain γ by complex shear modulus G^*

$$\tau=G^*\gamma \tag{14}$$

The complex shear modulus G^* is written in the form

$$G^*=G'+G''i=G'(1+i\eta_{ER}) \tag{15}$$

where G' is the storage modulus, G'' is the loss modulus and $\eta_{ER}=G''/G'$ is the loss factor.

The geomerty sizes and material parameters of sandwich beam filled with ER elastomers are list in Table 1. The shear modulus of the beam with applied electric field and without applied electric field were measured by Hao et al [8] and are shown in Table 2.

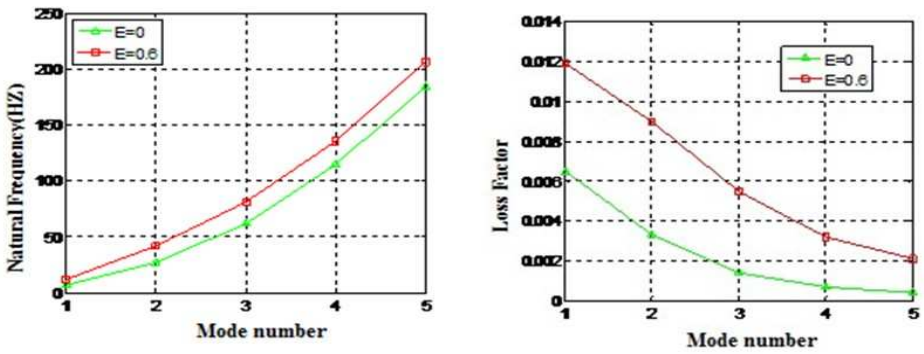
Table 1 Geometry sizes and material parameters of sandwich beam filled with ER elastomers

Parameters of the beam	$L=250mm, b=30mm, h_1=h_3=0.5mm, h_2=3mm$
Parameters of the elastic layer	$\rho_1=\rho_3=2800kg/m^3, E_1=E_3=70\times10^9 Pa$
Parameters of the ER elastomers layer	$\rho_2=1470kg/m^3$

Table 2 Shear modulus of sandwich beam filled with ER elastomers with different electric fields

E=0	$G=(37100+800i)Pa$
E=0.6	$G=(127900+4400i)Pa$

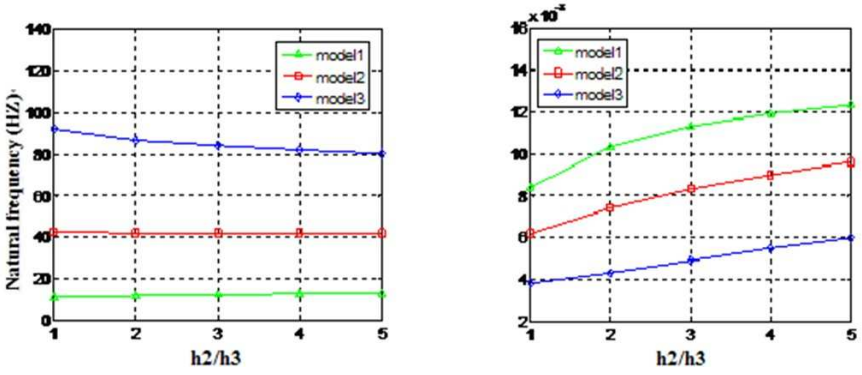
Figure 4 illustrates the natural frequency and loss factor variations of up to five modes. As shown in Figure 4(a), the natural frequencies shift to higher frequencies as the applied electric field increases. Figure 4(b) shows effect of electric field on the loss factors of the adaptive beam for E=0.6kv/mm and E=0kv/mm. The loss factors significantly decrease as the mode number increase. Because of the electric field effect, it makes the resonant frequecy of the system increased and the whole of the stiffness improved. These results demonstrate that the beam can suppress the vibration of the structure and the low mode has a greater inhibition effect to the high mode.



（a）Natural frequcies of the beam （b）Loss factors of the beam

Fig.4 Effect of the elctric field on the natural frequencies and loss factors of the sandwich beam filled with ER elastomers

A comparison of the natural frequencies of beam structures with different thickness of the ER elastomers layer is shown in figure 5(a), where a 0.6kv/mm electric field level is applied and the thickness of the top h_1=0.0005mm is not changed. The results show that the first two mode of natural frequencies are no significant change with the increase of the mutipile. The first step natural frequencies are about 13HZ and the second step natural frequencies are about 42HZ, but the third step natural frequencies are decreased as the mode nuber increase. It can be seen from the figure 5 (b) that the relationship between the loss factors and thickness rotio of h_2/h_3 is almost increased linerly. The loss factors of low mode is larger than that of the high mode in the same thickness ratio. The results demonstrate that the sandwich beam filled with ER elastomers might have better suppresing effect at low mode.



(a) （b）

Fig.5 Dependence of the natural frequencies and loss factors of the first three order of the beam at $E = 0.6kv/mm$ with changing thickness of the ER elastomers layer

Conclusion

The sandwich beams filled with ER elastomers with different applied electric field strength and different structural parameters have been investigated. Their dynamic charateristics have been simulated and analyzed to study the influence of the natural frequencies and loss factors for various applied electric field strength and different structural parameters. The results show that natural frequencies of the ER elastomers sandwich beam increase and vibration amplitudes of the beam decrease as the intensity of applied electric field increases. Increased the thickness of the ER elastomers layer, natural frequencies of the beam decrease and loss factors increase. Those indicate that the dynamic characteristic of ER elastomers sandwich beams is similar as that of ER fluids beam, which can be used for vibration control of flexible structures by applied a electric field.

Refrence

[1] R Sakurai, H See, T Saito, M Sumita. Effect of matrix viscoelasticity on the electrorheological properties of particle suspensions, Journal of Non-Newtonian Fluid Mechnics, 81(1999),p235-250

[2] Yalcintas M, Coulter J P. Adaptive beam model with electrorheological material based applications. Journal of Intelligent Material Systems and Structures, 1995, 6(4),p498-507.

[3] Rahn C D, Joshi. Modeling and control of an electrorheological sandwich beam, Journal of Vibration and Acoustics, Transactions of the ASME, 1998, 120(1),p 221-227.

[4] Young K K, Kim J, Choi S B. Passive and active damping characteristics of smart electro-rheological composite beams, Journal of Smart Materials and Structures, 2001, 10(4),p 724-729.

[5] Yeh J Y, Chen L W,Wang C C. Dynamic stability of a sandwich beam with a constrained layer and electrorheological fluid core, Journal of Composite Structures, 2004, 64(1),p47-54.

[6] Gao L X, Zhao X P. The behavior responsing to electric field of starch/gelatin/glycerin aqueous electrorheological elastomers,Journal of Function Material,2004,4(35),p426-431(in Chinese).

[7] Gao L X. Preparation and properties of composite hydrous electrorheological elastomers, [Ph.D.Thesis].Shanxi:Northwestern Polytechnical University,2006(in Chinese)

[8] Hao L M, Shi Z H, Zhao X P. Mechanical behavior of starch/silicone/silicone rubber hybrid electric elastomer, Journal of Reactive & Funcyional Polymers ,2009, (69),p165-169

Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.849

Modeling and Rendering Realistic Trees

Xingquan Cai [a], Baoxin Qian and Jinhong Li

College of Information Engineering, North China University of Technology, Beijing, 100144

[a] xingquancai@126.com

Keywords: realistic trees, recursive fractals, forest scenery, real-time rendering

Abstract. In this paper, we present an efficient modeling and rendering realistic trees method based on recursive fractals. Firstly, according to the growth characteristics of poplar trees, we divide the trees model into three parts, trunks, branches and leaves. We provide a multi-branches tree topological structure to organize all the trunks and the branches, and add leaves to leaf nodes of our multi-branches tree structure. Then we define the recursive fractals iterator to generate trunks model and side branches model, and use billboard to simulate the tree leaves. After that, we render the trees. Finally, we implement our method, and model and render different shape trees. The experiments prove that our method is feasible and high performance.

Introduction

Because of the complexity and diversity of the natural plants, modeling and rendering plants, especially modeling and rendering trees, has been one of the most difficult tasks in computer graphics. And modeling and rendering plants is one of the hot topics of the simulation of realistic natural sceneries.

In this paper, we present an efficient modeling and rendering realistic trees method based on recursive fractals. Firstly, according to the growth characteristics of poplar trees, we divide the trees model into three parts, trunks, branches and leaves. We provide a multi-branches tree topological structure to organize all the trunks and the branches, and add leaves to leaf nodes of our multi-branches tree structure. Then we define the recursive fractals iterator to generate trunks model and side branches model, and use billboard to simulate the tree leaves. After that, we render the trees. Finally, we implement our method, and model and render different shape trees. The experiments prove that our method is feasible and high performance. Our method could be used in photorealistic games, computer animations, computer generated films, virtual garden design, etc.

In this paper, after exploring the related work of modeling and rendering trees, we present our modeling and rendering realistic trees method based on recursive fractals. Then we provide the results using our method. Finally, we draw the conclusion.

Related work

Modeling and rendering trees has interested computer graphics researcher over the last decades. Many approaches have been proposed to it [1]. They can be roughly classified into several types, image based method, volume based method, and rule based method.

Image based method, especially Image-Based Rendering (IBR), has become a major technique for rendering real-world scenes. Meyer [2] proposed a method to improve the view quality of billboards by pre-computing a large set of views. Nevertheless, the drawback of billboards will be visible in animations when walking around a tree and walking through or over-viewing at a forest. In 2007, Tan et al [3] used this approach to construct more complexity trees. Despite the low cost of an image, Zhang [4] proposed many limitations of image based rendering methods around diversity, lighting, and animation because of the static nature and the relatively high memory cost of an image.

As to volume based method, Reche et al [5] used the volume based method to reconstruct the tree from photographs. Nevertheless, the method just works for the tree with sparse leaves, rather than

dense forest, because single pixels contain the blended projection of numerous leaves, branches and background.

Rule-based method is to use rules or syntax to create 3D trees, such as L-system, Iterated Function Systems (IFS), particle system, reference axis technique, etc. Lindenmayer [6] proposed the formalism of L-systems as an original worked for plant modeling. Reeves [7] presented particle system and rendered complex trees using a set of small disks represent the leaves. Other work has attempted to reconstruct 3D tree geometry from 2D sketch. Makoto [8] presented a system for designing 3D models of botanical trees quickly and easily using freehand sketches and additional example-based editing operations. So in this paper, we present an efficient modeling and rendering realistic trees method based on recursive fractals.

Modeling and rendering realistic tree based on fractal recursion

After we observe and study the trees in our life, according to the growth characteristics of poplar trees, we divide the trees model into three parts, trunks, branches and leaves. One tree is composed with a lot of trunks, branches and leaves. So tree can be recursively defined that one tree has many trunks, one trunk has several branches, and leaves appear at the end of the branches. In this way, we can get the self-similar tree model by recursive fractals.

Multi-branches tree structure. We have a multi-branches tree topological structure to organize all the trunks and the branches. Usually, the bottom radius of trunk is equal to its father's top radius, while the bottom radius of branch is smaller than its father's top radius. In our multi-branches tree structure, the left node is the trunk node. But one trunk node might have two child trunk nodes, this case is rare. We can deal with this case by adding a flag in our structure, $m\_bIsTrunk$. If the value of $m\_bIsTrunk$ is true, this node is trunk. And if the value of $m\_bIsTrunk$ is false, this node is branch.

Just as Fig. 1 shows, the left is the trunks and the branches. AB, BC, CD are the trunks, and EF, GH, JK are the branches. The right of Fig. 1 is our multi-branches tree structure hierarchy chart. We define one structure to organize the data of trunks and branches, just as Table 1. In Table 1, m_pFather is used to point the father-branch node, m_pChild is used to store the first child-branch node information, and m_pNext is used to point the brother-branch node. In this way, we can organize the trunks and the branches.

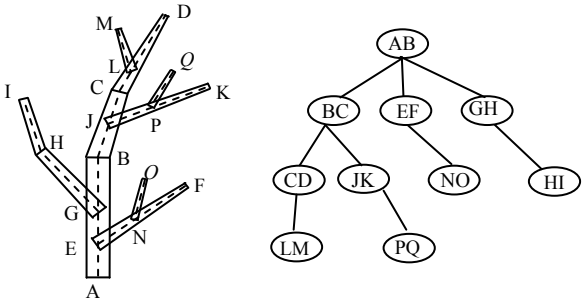


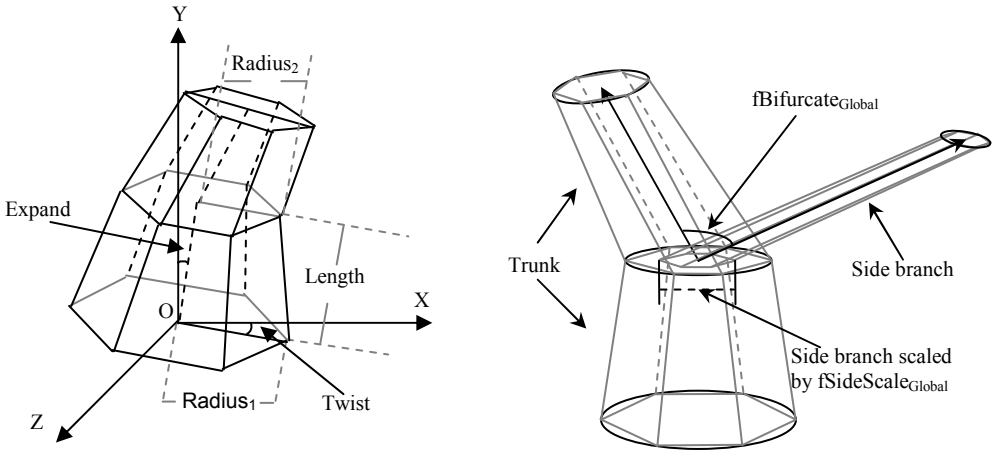
Figure 1. Static area and dynamic area

Branches modeling. Usually, the appearance of branches of trees is cylinder. It is difficult to render so many cylinders. So we select the multi-side-quads frustum to present the branches, and all the control points of the multi-side- quads frustum are on the bottom circle and the top circle. The branch can be presented by three to ten side-quads frustum. In order to present the detail of the branch and don't increase the complexity, we select the six-side-quads frustum. And the bottom and top of the frustum are regularhexagons.

Just as Fig. 2a) shows, in our method, the branch has five attributes, Radius1, Radius2, Length, Twist and Expand. Radius1 is the radius of the bottom regularhexagon, and Radius2 is the radius of the top regularhexagon. Twist is the angle which the branch rotates the Y-axis, and Expand is the angle which the branch rotates the Z-axis. Radius1, Radius2, and Length control the size of the branch. Twist and Expand control the global distortion of the branch.

Table 1. Structure of the branch

```
struct SBranch {
    bool m_bIsTrunk;
    float m_fTwist;
    float m_fExpand;
    float m_fLength;
    float m_fRadiusBottom;
    float m_fRadiusTop;
    SBranch* m_pFather;
    SBranch* m_pChild;
    SBranch* m_pNext;
    int m_iChildrenNumber
};
```



a) b)

Figure 2. Model of the branch and model of the side branch

We can select recursive fractals to generate the self-similar tree trunks model, so we define the iterator as Equation (1). In order to increase the randomness of the tree model, we add some random parameters in our iterator. In Equation (1), r_0, r_1, r_2, r_3 are the random Integer parameters in $\{0,1,2\}$, and they control the Twist, Expand, Length, Radius of the branch. In this way, we can generate the tree trunk model.

$$\begin{cases} fTwist_n = fTwist_{n-1} + fTwist_{Global} \cdot (0.9 + 0.05r_0) \\ fExpand_n = fExpand_{Global} \cdot (0.9 + 0.05r_1) \\ fLength_n = fLength_{n-1} \cdot (0.75 + 0.1r_2) \\ fRadiusBottom_n = fRadiusTop_{n-1} \\ fRadiusTop_n = fRadiusTop_{n-1} \cdot (0.75 + 0.1r_3) \\ iDepth_n = iDepth_{n-1} - 1 \end{cases} \quad n \in [0, iDepth_0] \tag{1}$$

$$\begin{cases} fTwist_{side} = fTwist_{Base} + fTwist_{Global} \cdot (0.9 + 0.05r_0) + 180\deg \\ fExpand_{Side} = (fBifurcate_{Global} - fExpand_{Base}) \cdot (0.9 + 0.05r_1) \\ fLength_{Side} = fLength_{Base} \cdot fSideScale_{Global} \cdot (0.75 + 0.1r_2) \\ fRadiusBottom_{Side} = fRadiusTop_{Base} \cdot fSideScale_{Global} \\ fRadiusTop_{Side} = fRadiusTop_{Base} \cdot fSideScale_{Global} \cdot (0.75 + 0.1r_3) \\ iDepth_{Side} = Min(iDepth_{Base} - 1, iDepth_{Base} \cdot fDepthScale_{Global}) \end{cases} \tag{2}$$

As to the side branches, we also define the iterator to generate the tree side branches model. Usually, the side branch is slenderer than the trunk, just as Fig. 2b) shows. So we add three parameters to control the shape of the side branch, $fSideScale$, $fBifurcate$ and $iDepth$. $fSideScale$ is used to control the

radius of the side branch, and $fBifurcate$ is used to compute the Expand angle of the side branch. One side branch might have some child branches, so we use $iDepth$ to control the depth of the side branch. We define the iterator as Equation (2) to generate the side branch model. In Equation (2), r_0, r_1, r_2, r_3 also are the random Integer parameters in $\{0,1,2\}$, and control the Twist, Expand, Length, Radius of the side branch.

Leaves modeling. Leaves are the important characteristic of the trees, and different types of trees have different leaves shape. Usually, the number of leaves is so large, and the edge information of leaves is so rich, so it is difficult to render them in detail as actual. We use billboard to simulate the tree leaves, and one quad shape billboard represents one leaf. Irregular border of leaf is visualized via Alpha Testing. It is also parameterized for different appearance.

Just as Fig. 3 shows, in our method, the leaf has four attributes, Scale, ScaleNoise, Expand, and Twist. Twist and Expand control the global distortion of the leaf. Scale controls the size of the leaf. In order to increase the variety of leaves and make the tree leaves look reality, we add a noise parameter, ScaleNoise. The noise sequence is produced by a random function. The initial position of leaves is at each branch twig.

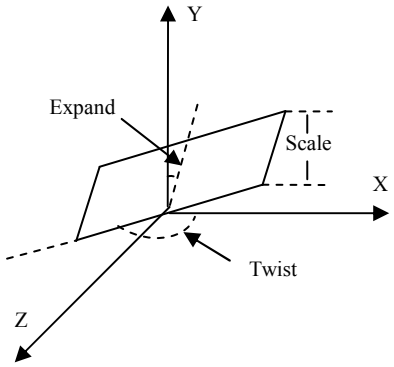


Figure 3. Model of the leaf

Rendering realistic tree. Since we have generated the model of the tree and have a structure hierarchy chart, when rendering the tree, we just need to traverse the multi-branches tree structure and shade the trunks, branches and leaves. In this way, we get the realistic trees. We also can add the shadow in our scene.

Results

We have implemented our method to model and render realistic trees. Our implementations are running on a Intel Core2 Duo 3.0GHz computer with 2GB RAM, and NVIDIA GeForce8400 SE graphics card with 256M RAM, under Windows 7 system, Visual C++ 8.0, and OpenGL environment. The rendering system has a real viewport size of 1024×768.

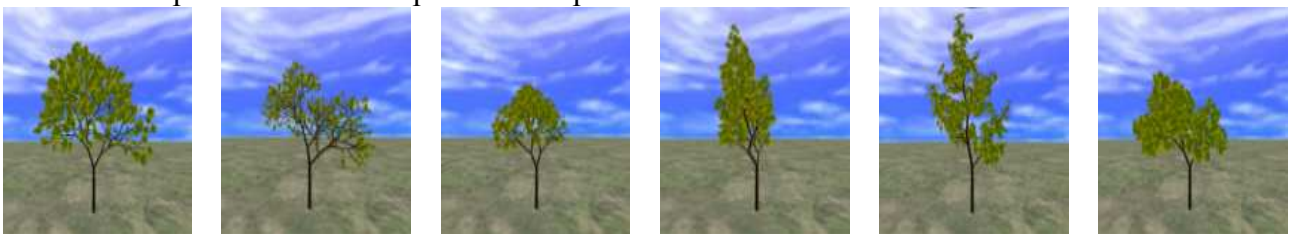
In order to test and verify our method, we implement our method to model and render single tree. We can change the branches parameters and leaves parameters of the tree model, and obtain the different shape trees. Just as Table 2, we set the different branches parameters and leaves parameters, and we also set the random noise. Fig. 4 shows the result trees in turn, and the shape of the result trees is different. Since we can get so many different trees, our method can be used to model and render large scale forest scene.

Conclusions and future work

In this paper, we present an efficient modeling and rendering realistic trees method based on recursive fractals. Firstly, according to the growth characteristics of poplar trees, we divide the trees model into three parts, trunks, branches and leaves. We provide a multi-branches tree topological structure to organize all the trunks and the branches, and add leaves to leaf nodes of our multi-branches tree

structure. Then we define the recursive fractals iterator to generate trunks model and side branches model, and use billboard to simulate the tree leaves. After that, we render the trees. Finally, we implement our method, and model and render different shape trees.

As a future possibility, we are working on adding other effects to our modeling trees and using our method to implement other complex natural phenomenon.



a) b) c) d) e) f)

Figure 4. Model of the single tree

Table 2. Different parameters of different shape trees

No.	Branches				Leaves	
	Expand	Twist	Radius	Length	Size	Expand
1)	21.78	105.84	0.76	0.88	3.36	7.65
2)	9.81	58.86	0.80	0.85	2.04	3.24
3)	32.22	128.70	0.76	0.82	2.39	47.97
4)	43.65	134.10	0.90	0.89	2.60	13.68
5)	13.05	98.10	0.91	0.91	2.84	81.27
6)	28.89	72.00	0.80	0.85	3.41	20.70

Acknowledgment

Supported by PHR(IHLB) Grant (PHR20100509, PHR201008202), Funding Project of Beijing Outstanding Academic Human Resource Program (No. 20081D0500200114), Funding Project of Beijing Municipal Education Committee (No. KM201010001027).

References

[1] Zhang Qi-Long, Pang Ming-Yong. A Survey of Modeling and Rendering Trees. *In Proceedings of Edutainment 2008*, pp. 757-764.

[2] Meyer A., Neyret F., Poulin P. Interactive Rendering of Trees with Shading and Shadows. *In Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, 2001, pp. 183-196.

[3] Ping Tan, Gang Zeng, et al. *Image based Tree Modeling. In Proceedings of SIGGRAPH 2007*.

[4] Zhang C., Chen T. A survey on image-based rendering representation, sampling and compression. *Signal Processing: Image Communication*, vol. 19, no. 1, 2004, pp. 1-28.

[5] Reche, A., Martin, I., Drettakis, G. Volumetric reconstruction and interactive rendering of trees from photographs. *In Proceedings of SIGGRAPH 2004,* pp. 720-727.

[6] Lindenmayer A. Developmental Systems without Cellular Interaction, their Languages and Grammars. *Journal of Theoretical Biology*, 1971, pp. 455-484.

[7] Reeves W. T., Blau R.. Approximate and Probabilistic Algorithms for Shading and Rendering Structured Particle Systems. *In Proceedings of SIGGRAPH 1985*, pp. 313-322.

[8] Makoto Okabe, Shigeru Owada, et al. Interactive design of botanical trees using freehand sketches and example-based editing. *In Proceedings of SIGGRAPH 2006,* pp. 487-496.

doi:10.4028/www.scientific.net/AMM.50-51.854

Comparison and Analysis on Four International Seismic Isolation Design Codes for Bridge

Tieyi Zhong [1, a], Jiying Lv[1] and Junhui Feng[2]

[1] School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China

[2] Guiyang Municipality Development and Reform Commission, Guiyang 550081, China

[a]tyzhong2008@163.com

Keywords: bridge, seismic isolation, lead rubber bearing, code comparison.

Abstract. As one kind of primary passive vibration control technology, the seismic isolation technology provides a new way for anti-seismic of bridge and has got initial application in practical bridge engineering. Some standard clauses have been added into the international codes including Eurocode 8, codes of AASHTO and Japan Highway as well as China Highway. In this paper, using the finite element program Sap2000, a three spans continuous beam bridge is taken as practical calculation example for analysis and comparison based on four codes above. The research results provide important reference for further research and practical application of bridge vibration reduction and isolation.

Introduction

The seismic isolated bearing installed in bridge can escape or mitigate the bridge damage and provides the effective protection for bridge during earthquake. Therefore, the researchers from many countries have carried out a lot of important research work on seismic isolation technology in the past twenty years [1-4]. At the same time, good anti-seismic performance has been verified in some earthquakes. At present, many countries and areas have increased the standard clauses related with seismic isolation in bridge anti-seismic design code including Europe, America, Japan and China, etc.

In this paper, based on the finite element program Sap2000, using three artificial earthquake waves, a three spans continuous beam bridge is taken as practical example for calculating the dynamic responses of seismic isolated bridge according to the four codes of Eurocode 8 [5], AASHTO [6], Japan Highway [7] and China Highway [8] respectively. By comparing and analyzing, some meaningful research results are acquired and will provide important reference for both further research and practical application of bridge vibration reduction and isolation.

Parameters of beam and pier

In this paper, a 120m length, three-span continuous bridge is taken as the practical example, and the span combination is 33m+54m+33m respectively. The calculation model is shown in figure 1. For the beam, C50 concrete is used , and the cross section area of the beam is 12.9m^2, and the lateral bending rigidity is 48.70m^4, as well as vertical bending rigidity is 3.43 m^4. The piers are rectangle sections with 2.2m in width and 2.8m in length as well as 9-10m in height, and the C45 concrete is used in piers.

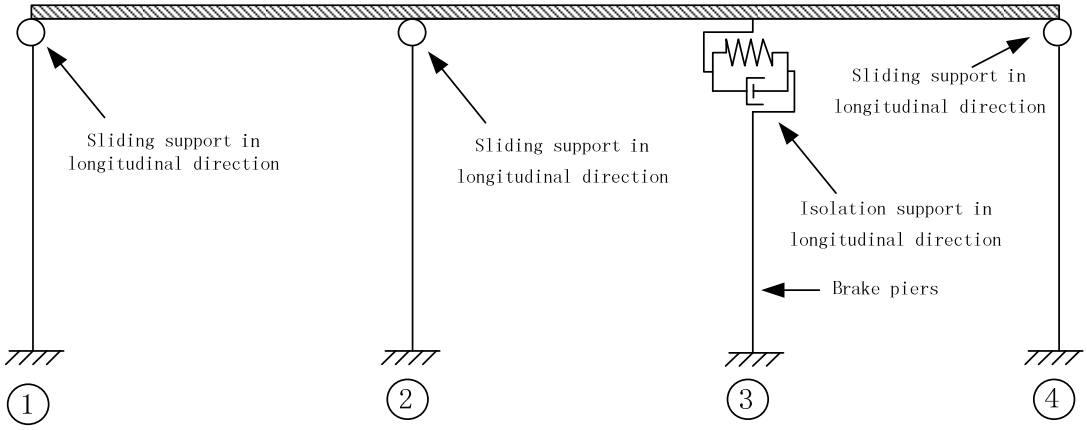


Figure 1 Sketch of the bridge model

Boundary conditions and damping. The 3# pier is brake pier, and the 1#, 2#, 4# piers are sliding piers whose bearings' restraints are fixed in transverse direction and free in longitudinal direction. For the isolated bridge system, the brake pier is isolated in longitudinal direction and fixed in transverse direction, and the constraints of the other piers are unchanged. The Rayleigh damping is used in this research.

Seismic excitations. The seismic excitations adopt three artificial seismic waves, which are generated considering the site conditions and the influences of random phase. The peak values of each wave are adjusted according to seismic fortification intensity of 8 degree and of 9 degree respectively, and the time history analyses of bridge seismic responses are done under the excitation of three artificial seismic waves corresponding to seismic intensity of 8 degree and of 9 degree, and the calculated results are compared and analyzed. The acceleration curves of three artificial seismic waves are shown in figure 2 to figure 4.

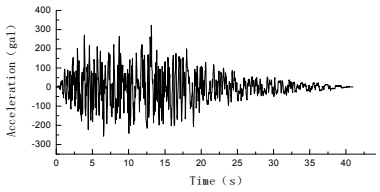


Figure 2 Artificial seismic waves No.1

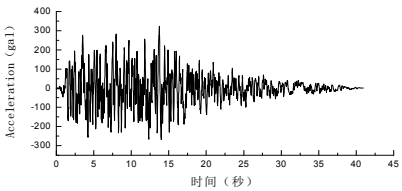


Figure 3 Artificial seismic waves No.2

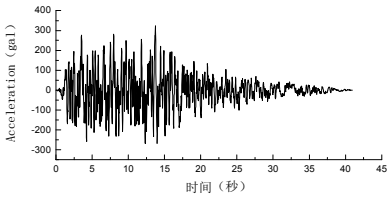


Figure 4 Artificial seismic waves No.3

Mechanical parameters of LRB. LRB is used only at brake pier of bridge for seismic isolation. The bilinear mechanical model of LRB is shown as figure 5.

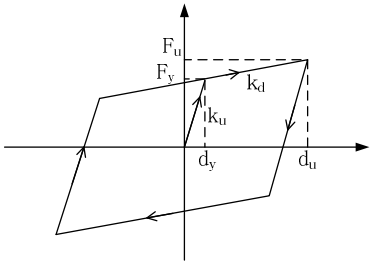


Figure 5 Mechanical models of LRB

The mechanical parameters of LRB used in this paper are shown as follows, the yield strength F_y is 4.8e5N, and the initial stiffness K_u is 1.6e8N/m, and the yield displacement d_i is 3mm, and the hardening ratio α is 0.1. The design displacement of LRB is determined as 90mm.

Equivalent stiffness and equivalent damping ratio as well as mechanical parameters of LRB according to different seismic isolation design codes are shown in table 1.

Table 1 Parameter of LRB

Code	Equivalent stiffness (N/m)	Equivalent damping ratio	Initial stiffness (N/m)	Yield displacement (mm)	Yield strength (N)	Hardening ratio
American	2.82e7	0.19	1.6e8	3	4.8e5	0.1
Japanese	2.64e7	0.23	1.6e8	3	4.8e5	0.1
European	2.48e7	0.23	1.6e8	3	4.8e5	0.1
Chinese	2.24e7	0.24	1.6e8	3	4.8e5	0.1

Results and analyses

Based on the LRB parameters in Table 1, finite element model of the isolated bridge is established and non-linear time history analyses are carried out. The results corresponding to the four different codes are compared and analyzed as follow.

Comparison of fundamental periods between models with and without isolation. The fundamental period of the model without isolation and the periods of the models with isolation for four codes are given in Table 2.

Table 2 Comparison of fundamental periods with and without isolation

Period without isolation (s)	Period with isolation (s)			
	AASHTO	Eurocode 8	Japan Highway	China Highway
0.63	1.601	1.692	1.647	1.769

It can be seen from Table 2 that the fundamental periods of isolated bridge are two times larger than that of non-isolated bridge. The calculation result according to the China Highway Code is the largest and the results are very close for the rest three codes.

Comparison of seismic isolation effect. Considering the excitations of seismic intensities of 8 degree and 9 degree and using three artificial earthquake waves, the seismic responses of isolated bridge are computed and compared according to the four codes aforementioned. For the calculated results under three artificial earthquake waves excitation, the maximum of seismic responses for the three seismic waves is chosen to be compared.

Fig. 6 and 7 give out the time history displacement responses of pier top under different seismic intensities corresponding to four codes.

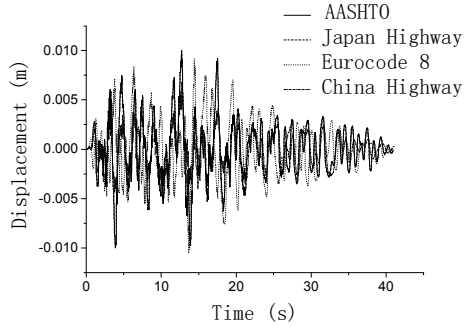


Figure 6 Time history of displacement responses of pier top under intensity 8

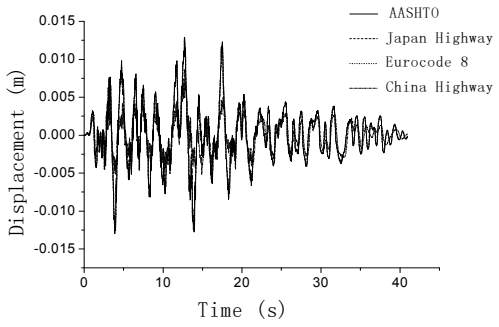


Figure 7 Time history of displacement responses of pier top under intensity 9

It can be seen from Fig.6 and Fig.7 that under the same excitation, the maximum displacement responses of pier top corresponding to the four codes are all different and it is due to that the effective

stiffness of LRB are different when calculated according to the four codes, which leads to the variations of vibration properties of the bridge. The peak responses and isolation ratios of the seismic isolated bridge corresponding to four codes are mainly compared and analyzed in the following.

Comparison of peak responses and isolation ratios corresponding to four codes. Comparison of peak seismic responses corresponding to four codes under seismic intensities of 8 degree and 9 degree are shown in Table 3 and Table 4 respectively.

Table.3 Comparison of peak responses corresponding to four codes (8 degree)

Seismic responses	Without isolation	With isolation			
		AASHTO	Eurocode 8	Japan Highway	China Highway
Shear force at pier bottom (N)	2.49E+07	4.70E+06	4.61E+06	4.58E+06	3.05E+06
Moment at per bottom (N·m)	1.48E+08	3.88E+07	3.78E+07	3.76E+07	2.46E+07
Displacement of pier top (m)	3.50E-02	1.10E-02	9.00E-03	1.00E-02	6.84E-03
Displacement of beam (m)	3.77E-02	1.30E-01	1.40E-01	1.70E-01	1.89E-01

Table.4 Comparison of peak responses corresponding to four codes (9 degree)

Seismic responses	Without isolation	With isolation			
		AASHTO	Eurocode 8	Japan Highway	China Highway
Shear force at pier bottom (N)	3.3E+07	5.57E+06	5.47E+06	5.36E+06	3.54E+06
Moment at per bottom (N·m)	1.97E+08	4.6E+07	4.47E+07	4.44E+07	3.1E+07
Displacement of pier top (m)	5E-02	1.3E-02	1.24E-02	1.27E-02	8.67E-03
Displacement of beam (m)	4.6E-02	1.67E-01	1.58E-01	2.07E-01	2.67E-01

It is shown in Table 3 that, under the seismic intensity of 8 degree, the isolation ratios of shear forces at pier bottom are 81.1%, 81.5%, 81.6% and 87.7% and the isolation ratios of moments at pier bottom are 73.8%, 74.5%, 74.6% and 83.4%, as well as the isolation ratios of displacements of pier top are 68.6%, 74.3%, 71.4% and 80.4% corresponding to the code of AASHTO, Eurocode 8, Japan Highway and China Highway respectively. At the same time, the increase ratios of displacement of beam are 71%, 73.1%, 77.8% and 80.1%.

Also, it can be seen from Table 4 that, under the seismic intensity of 9 degree, the isolation ratios of shear forces at pier bottom are 82.7%, 83.4%, 83.7% and 89.27% and the isolation ratios of moments at pier bottom are 76.6%, 77.3%, 77.5% and 84.3% as well as the isolation ratios of displacements of pier top are 74%, 75.2%, 74.6% and 82.7% corresponding to the code of AASHTO, Eurocode 8, Japan Highway and China Highway respectively. At the same time, the increase ratios of displacement of beam are 72.4%, 70.9%, 77.8% and 82.8%.

Apparently from the comparison above, under the seismic intensities of 8 degree and 9 degree, the isolation ratios of shear forces at pier bottom, moments at pier bottom and displacements of pier top and the increase ratio of displacement of beam calculated according to Chinese code are the largest in four codes and the other three codes are close. Thus, to this example, comparing to Chinese code, Eurocode 8, AASHTO and Japan Highway code tend to be more conservative.

Conclusions

In this paper, based on the finite element program Sap2000, the finite element model of a three spans isolated continuous beam bridge is established considering isolated bearing only installed on brake pier, and the calculation of dynamic characteristics and responses of above bridge are carried out according to the codes of Eurocode 8, AASHTO, Japan Highway and China Highway respectively. Some valuable results have been achieved as follow: (1) The fundamental periods of isolated bridge are two times larger than that of non-isolated bridge, and the calculation results is the largest according to the China Highway Code and the results are very close for the rest three codes. (2) For seismic intensity being 8 degree and 9 degree, using three artificial earthquake waves, the seismic responses of bridge are calculated with and without isolation. The results indicate that the ratios of vibration reduction of bridge pier including the moment and shear force at pier bottom as well as displacement on top of pier and the increasing ratio of beam displacement are all the largest for China Highway Code, and those results of Eurocode 8, AASHTO and Japan Highway code are very close, tending to be more conservative.

Acknowledgment

This study is sponsored by the Science & Technology Research Development Project of China Railway Ministry (Grant No. 2008G010-A) and the Science & Technology Research Development Project of China Railway Ministry (Grant No. 2008G028-C).

References

[1] Buckle, I. G. and Mayes, R. L.: Seismic Isolation: History, Application, and Performance-A World View, Earthquake Spectra, Vol. 6, No.2, 1990.

[2] Panos Tsopelas and Michael. C. Constantinou. Study on Elastoplastic Bridges Seismic Isolation System, Journal of Structural Engineering, 123 (4): 489-498, 2000.

[3] T.Y. Zhong, F.L. Yang and H. Xia. The Design Method of the Seismically Isolated Bridge by Lead Rubber Bearing Based on Energy Method, China Railway Science, 30(2): 43-48, 2009.

[4] Kyu-Sik Park, Hyung-Jo Jung and In-Won Lee. A Comparative Study on A Seismic Performance of Base Isolation Systems for Multi-span Continuous Bridge [J]. Engineering Structures, 2002,24: 1001-1013.

[5] Eurocode 8: Design of structures for earthquake resistance-Part 2: Bridges. 1998-2: 2006-2

[6] AASHTO: Guide Specifications for Seismic Isolation Design, 1999

[7] Standard Specification for Seismic Design of Road and Bridge V. Tokyo: Japan Society of Road, 1996

[8] Guidelines for Seismic Design of Highway Bridges. (JTG/T B02-01-2008), 2008. Beijing, China.

doi:10.4028/www.scientific.net/AMM.50-51.859

Time-dependent Load bearing Capacity of Corroded R.C. Columns in Atmosphere Environment

LI Qiang [1,2,a], JIN Nanguo[1], JIN Xianyu[1]

[1] Department of Civil Engineering, Zhejiang University, Hangzhou 310027, P.R. China

[2] Zhejiang College of Construction, Hangzhou, 311231, P.R. China

[a]liqiangtzh@126.com

Keywords: reinforced concrete, columns, corrosion, time-dependent behaviour, bearing capacity

Abstract. As a supporting member of concrete structures, the reliability of the columns is of critical importance. To have models targeted to estimating the residual load-bearing capacity of corroded reinforced concrete (R.C.) columns so that inspection procedures and maintenance interventions can be well conducted, a time-dependent capacity model for corroded R.C. column is presented. The model was based on equilibrium equations on which load-bearing capacity of R.C. column at given time (year) can be predicted considering the overall deteriorations of rebar, concrete section, mechanical properties and bond behavior between corroded reinforcement and concrete. The model was verified by published literature test results of corroded R.C. columns served in atmosphere environment for years. The comparison of the model predictions with the test results shows the validity of the model. In the end, considerations for use of the model were suggested.

Introduction

Corrosion of reinforcement embedded in concrete is a worldwide problem. It not only reduces the bar section and the mechanical properties of reinforcement, but also generates concrete cracking and changes bond strength between reinforcement and concrete due to the expansion of corrosion products. Synergistically, load-bearing capacity and mechanical behavior of RC elements are affected by the corrosion of reinforcement [1]. Worldwide attentions have been paid to the corrosion issues such as the mechanisms of corrosion, effects of corrosion on concrete structures, testing methods to corrosion monitoring, measures to corrosion prevention, and effectiveness of repair and strengthening of corroded structures. However, the majority of the previous work has been focused on corroded R.C. beams. Works on the mechanical behavior and load capacity of corroded R.C. columns are comparatively limited compared with works on beams. It is well known that collapse of column may lead to very high risks of building collapses. As a supporting member of concrete structures, the reliability of the columns is of critical importance.

Fortunately, a number of researchers [2-7] have made contributions to our understanding of behavior of corroded R.C. columns in recent years. Numerical and analytical models have also been developed and calibrated with the experimental results to predict the bearing capacity of corroded R.C. columns [8-9]. Ref. [10] reviewed the previous studies carried out to investigate the behavior of Corroded R.C. columns. Generally speaking, the previous studies mainly concentrated on the mechanical properties of corroded members at a given time or given corrosion. Few models consider the overall deteriorations of rebar and concrete section, materials mechanical properties and the relationship of bond-slip. The time-dependent model of life cycle behavior for RC column is not well established. It is therefore extremely important to have models targeted to estimating the residual load-bearing capacity of columns at any given time so that inspection procedures and maintenance interventions can be well conducted. Although some models of time-variant degradation of material properties are not well established yet, authors of this paper still have an intention of taking their first step towards the time-dependent load-bearing capacity of corroded R.C. columns. This paper, therefore, based on the Code for Design of Concrete Structures (GB 50010-2002, China) [11] and Standard for Durability Assessment of Concrete Structures (CECS 220) [12], try to present a

simplified model, taking into account all of the degradations mentioned above, in particular, the relationship between the corrosion degree and the load bearing capacity of corroded R.C. columns in atmosphere environment. The results of the model computation are verified by two experimental results [14, 15]. For simplicity, only even corrosion is considered in this paper.

Modelling the Time-Dependent Load-Bearing Capacity of Corroded R.C. Columns

Load-bearing capacity model of corroded R.C. columns. Fig.1 shows the simplified sketch of compressive load-bearing capacity for large eccentric compression column with rectangular section. The time-dependent compressive load-bearing capacity for eccentric compression column of rectangular section shall conform to the following stipulations.

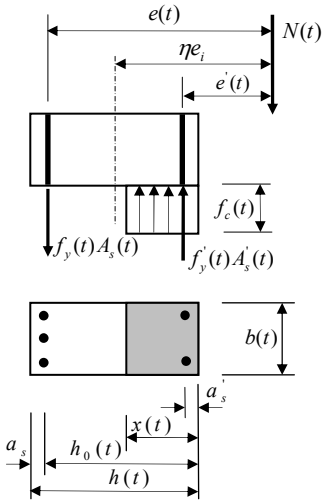


Fig.1 Simplified sketch of compressive load-bearing capacity for large eccentric compression members with rectangular section

$$N(t) = \alpha_1 f_c(t) b(t) x(t) + f_y'(t) A_s'(t) - f_y(t) A_s(t) \ . \tag{1}$$

$$N(t)e(t) \le \alpha_1 f_c(t) b(t) x(t) \left(h_0(t) - \frac{x(t)}{2} \right) + f_y'(t) A_s'(t) \left[h_0(t) - a_s' \right] . \tag{2}$$

Where, $N(t)$ is the time-dependent eccentric compressive force, α_1 is coefficient calculated according to Ref. [11]. $x(t)$ is the time-dependent depth of compression zone of equivalent rectangular stress diagram. $b(t)$ and $h_0(t)$ are the time-dependent width and effective depth of rectangular section respectively. $h_0(t)$ is the distance from the point for resultant of longitudinal tension steel reinforcement to the surface of compression zone.

$$b(t) = b - 2\alpha c \ . \tag{3}$$

$$h_0(t) = h - c - d/2 - \alpha c \ . \tag{4}$$

Where, h is sectional height; c is the concrete cover, here we assuming that the concrete cover is the same for one column. if c is different for tension and compression zone, the real thickness of the concrete cover should be taken; d is the initial diameter of reinforcement. when $0 < \omega < 2mm$, $\alpha = 0.15\omega$, when $\omega = 2mm \sim 3mm$, $\alpha = 0.45$ [13], ω is the width of the crack.

$$e(t) = \eta e_i + \frac{h_0(t)}{2} - \alpha_s \cdot \tag{5}$$

$$e_i = e_0 + e_a \ . \tag{6}$$

Where $e(t)$ is the distance between acting point of axial compressive force and point for resultant of forces of longitudinal ordinary steel bars. η is the eccentricity enlargement coefficient for axial compressive force of eccentric compression member, considering the influence of second bending moment, calculated according to the Ref. [11]. e_i is the initial eccentricity. a_s is the distance from the point for resultant of forces of longitudinal ordinary steel bars to the near edge of section. e_0 is the eccentricity of axial compressive force with respect to centroid of section. e_a is additional eccentricity, determined according to the code [12].

Time-variant material properties. $f_c(t)$ is the time- dependent axial compressive strength of concrete, which can be determined by equation (7), f_{c0} is 28-day concrete cubic strength [12].

$$f_c(t) = 0.737 f_{c0} \bullet 1.4529 e^{-0.0246(\ln t - 1.7154)^2} \ . \tag{7}$$

$$A_s(t) = A_s\left[1 - \eta_s(t)\right] \ . \tag{8}$$

$$f_y(t) = \xi_s(t) f_y = \frac{\left[1 - 1.077\eta_s(t)\right]}{\left[1 - \eta_s(t)\right]} f_y \ . \tag{9}$$

A_s and f_y are the initial section area and tensile yield strength of ordinary steel bars, respectively; $\xi_s(t)$ is time-dependent strength reduction factor of reinforcement, as suggested in Ref.[12]; $\eta_s(t)$ is the time-dependent sectional area loss ratio of corroded reinforcement, for even corrosion, it can be evaluated by:

$$\eta_s(t) = \frac{A_s - A_s(t)}{A_s} = \frac{4\Delta d(t)}{d} - \frac{4\Delta d^2(t)}{d^2} \cong \frac{4\Delta d(t)}{A_s} \ . \tag{10}$$

Where, $\Delta d(t)$ is the reduction in diameter of reinforcement due to corrosion, as suggested in Ref. [12].

$$\Delta d(t) = \begin{cases} 0 & for \ \ t < t_i \\ v_0(t - t_{in}) & for \ \ t_{in} \le t \le t_{cr} \\ v_0(t_{cr} - t_{in}) + v_1(t - t_{cr}) & for \ \ t_{cr} < t \end{cases} \ . \tag{11}$$

t_{in} is the corrosion initiation time (years), t_{cr} is the time for concrete cover cracks caused by corrosion. t_{in} and t_{cr} are determined by:

$$t_{in} = (c - x_0)^2 / k^2 \ . \tag{12}$$

$$t_{cr} = t_{in} + \frac{\Delta d_{cr}}{v_0} \ . \tag{13}$$

$$x_0 = (1.2 - 0.35k^{0.5}) \bullet c - \frac{6}{k_{ce} + 1.6}(1.5 + 0.84k) \cdot \tag{14}$$

$$k = 3K_{co_2} \bullet K_{kl} \bullet K_{kt} \bullet K_{ks} \bullet RH^{1.5}(1 - RH) \bullet T^{0.25}(\frac{58}{f_{cu}} - 0.76) \cdot \tag{15}$$

Where, Δd_{cr} is the critical corrosion depth of reinforcement when corrosion-induced crack is 0.1mm, determined by equation (16) and (17) for corner bar and normal bar respectively [12]. x_0 is Carbonation remains, k is

$$\Delta_{cr} = 0.012c/d + 0.00084 f_{cuk} + 0.018 \cdot \tag{16}$$

$$\Delta_{cr} = 0.015(c/d)^{1.55} + 0.0014 f_{cuk} + 0.016 \cdot \tag{17}$$

v_0 and v_1 are the corrosion rates before and after concrete cover cracking caused by corrosion respectively, determined by Ref. [12]. Equation (18) and (19) are for outdoors and indoors respectively.

$$v_0 = 7.53K_{cl} \bullet m \bullet (0.75 + 0.0125T)(RH - 0.45)^{2/3} \bullet c^{-0.675} \bullet f_{cuk}^{-1.8} \cdot \tag{18}$$

$$v_0 = 5.38K_{cl} \bullet m \bullet (0.75 + 0.0125T)(RH - 0.5)^{2/3} \bullet c^{-0.675} \bullet f_{cuk}^{-1.8} \cdot \tag{19}$$

$$v_1 = (4.5 - 340v_0)v_0 \cdot \tag{20}$$

where, K_{cl} is the position coefficient of steel bar, for corner bar K_{cl}=1.6, else K_{cl} =1.0; m is the coefficient of local environment; T is environmental temperature (℃); RH is the annual average relative humidity (%), when RH>80%, then RH=80%; c is cover depth (mm); f_{cuk} is the concrete cubic strength (MPa).

The calculation of time-variant material property $f_y'(t)$ and $A_s'(t)$ is same as that of $f_y(t)$ and $A_s(t)$ as mentioned above, which will not be listed here.

Model Verification

The model was verified by published literature test results of corroded R.C. columns that had served 24 years [14] and 36 years [15] in the atmosphere environment. For the limitation of space, the detail of the calculation is not listed here. Table 1 shows the comparison of the predictions with the tested ones. It can be seen from table 1 that the predictions are smaller than the test ones.

Table 1.The comparison of the predictions with tested ones

Ref.	Columns dimensions	Service(year)	Predicted/Tested
Ref.[14]	2900mm×240mm×240mm	24	0.79
Ref.[15]	2150mm×200mm×200mm	36	0.84

Conclusions

It is feasible to predict the time-dependent load bearing capacity of concrete columns at given time (year) by considering the overall deteriorations of rebar and concrete section, mechanical properties and bond behavior between corroded reinforcement and concrete. However, the deteriorations of the

concrete member are such a complexity that it is difficult to predict precisely the time-dependent behavior in life cycle time simply using models. Own to the limitation of the test results, the validation of the model would be further researched. It is greatly recommended to combine the set inspection with the prediction of the model when maintenance interventions of R.C. structure are conducted.

Acknowledgements

The support of National Natural Science Foundation of China (Grant No.50838008) and National Basic Research Program of China (973 Program) (Grant No. 2009CB623200) is gratefully acknowledged.

References

[1] J.A. Gonza'lez. Some questions on the corrosion of steel in concrete. Part 1: when, how and how much steel corrodes. Materials and Structures, Vol. 29, January-February (1996), p.40-46

[2] Shi Qing-xuan. Experimental study of bearing capacity of corroded reinforced concrete eccentric compressive members. Industrial Construction, Vol.31, No.5 (2001), p.14-17

[3] XU Shanhua, The models of deterioration and durability evaluation of reinforced concrete structure, Xi'anUniversity of Architecture and Technology, (2003)

[4] Fu-Xue Sun. Theoretical and Experimental Studies on Predicting the ServieeLife of Subsea Tunnel Lining Structures, [D]. TongJi University, (2006)

[5] LEI Guoqiang. Experimental research on the carrying capacity of corroded R.C. columns, [D]. Hunan University, (2007)

[6] YI Weijian. Experimental Research on the Bearing Capacity of Corroded Reinforced Concrete Columns. Journal of Hunan University (Natural Sciences), No. 3(2008), p.6-10

[7] Xiao-HuiWang, Fa-Yun Liang. Performance of RC columns with partial length corrosion, Nuclear Engineering and Design, No.238 (2008), p.3194–3202

[8] WEI Jun, ZHANG Hua. Experimental research on the deterioration of the flexural stiffness of eccentric compressive members due to the corroded bar. Journal of Wuhan University of Technology, No. 7 (2009), p.60-64.

[9] Xue-bin LU. Buckling property of corroded steel rebar and lts influences on reinforced conerete column seetional strength [D], TongJi University, (2007)

[10] Li, Qiang. Bearing capacity of corroded RC columns, ICIC 2010-3rd International Conference on Information and Computing. Wuxi, Jiang Su, China, (2010), p.233-236

[11] GB50010-2002. Code for Design of Concrete Structures (in Chinese), Beijing: China Architecture & Building Press, China, (2002)

[12] CECS 220, Standard for durability assessment of concrete structures (in Chinese), Beijing: China Engineering Construction Standard (CECS) Association, China Planning Press, China, (2007)

[13] Hui Yunling, Li Rong. Experimental studies on the property before and after corrosion of rebar in basic concrete members. Industrial Construction, No.6, (1997), p.14-18

[14] PAN Yi, CHEN Zhaohui. Experimental studies on the property of corrosion in basic concrete members. Sichuan Building Science, No. 3, (2004), p.71 -74.

[15] Tao Feng, Wang Linke. Experimental study on the bearing capacity of existing reinforced concrete members. Industrial Construction. No. 4, (1996), p.17-20.

Online: 2010-12-31 Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.864

New Method of Reliability Analysis for Deep Tunnel

Chen Guoqing[1, a], Su Guoshao[2,b] and Li Tianbin[1,c]

[1] State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu, China

[2] College of Architecture and Civil Engineering,Guangxi University,Nanning, China

[a]chgq1982@126.com, [b]suguoshao@163.com, [c]ltb@cdut.edu.cn

Keywords: deep tunnel; reliability; monte carlo; support vector machine; particle swarm optimization

Abstract. The performance function of deep tunnel in the reliability analysis can not be explicitly expressed and has non-linear characteristics, a new method of reliability analysis is proposed by combining intelligent methods and Monte Carlo principle. The support vector machine(SVM) model is employed to establish the nonlinear mapping relationship between basic random variables and structural response action, particle swarm optimization is used to optimize the parameters of SVM model. The learned SVM model is integrated with Monte Carlo method to calculate the failure probability and reliability. At last, the validity of the proposed method is verified by two typical engineering examples, the reliability index of wall rock along Jinping Hydropower Station diversion tunnels is calculated，the results are in good agreement with the actual situations. The proposed method could benefits other similar projects.

1. Introduction

Deep tunnel will be built frequently in many projects such as “south-north diversion project” and “west-east power transmission project” in China, many accident in practical engineering break out such as collapse, rockburst, karst collapse, debris flow and so on. The reason is not that the accuracy of the calculation method, but because of the deterministic methods of analysis with the calculation failed to take account of uncertainties. With the development of the theory of reliability [1-2], the study of structural reliability has reached a practical level. Now, in the reliability analyses, first-order second-moment method or JC method are recommended widely, however, these methods can only solve explicit performance function of a simple structural reliability [3]. For the implicit limit state equation with the complex structural reliability calculation, response surface method are proposed [4], but the nonlinear function approximation ability of the response surface polynomial form is limited, the accuracy of response surface can not solve the problem.

The performance function of limit state for deep tunnel is nonlinear equations, or even can not be clearly expressed the limit state equation, the reliability analysis of such issues would be difficult, it is necessary to study efficient and accurate method of calculating reliability. Monte Carlo method is an exact algorithm for reliability calculation, and can solve non-linear equation of limit state fundamentally [5]. Therefore, this article based on the Monte Carlo principle, intelligent methods (SVM-PSO) used to solving complex engineering systems in great advantages [6]. The trained SVM model can instead of numerical calculation in the iterative calculation of Monte Carlo process, greatly improving the calculation efficiency.

2. Fundamental principles

The Monte Carlo method is based on law of large numbers, firstly, large random sampling the random variable are conducted, then put these sample value into the performance function, determine whether the structure are failure, the more sampling times, the more accurate results, so it has the ability of solve complex problems directly.

SVM is based on statistic learning theory and the parameters selection is important to it. The optimization problem for SVM is the following:

$$\min J(\omega,\varepsilon)=\frac{1}{2}\omega^T\omega+\gamma\frac{1}{2}\sum_{i=1}^{N}\varepsilon_i^2 \tag{1}$$

Where: the fitting error is denoted by ε_i. Regularization parameter γ controls the trade-off between the smoothness of the function. This optimization problem leads to a solution:

$$\hat{f}(x)=\sum_{i=1}^{N}\alpha_i K(x,x_i)+b \tag{2}$$

Where: α_i are the coefficients and $K(x,x_i)=\varphi^T(x)\varphi(x_i)$ is the kernel.

Insensitive factor ε and penalty factor C are the main factors affecting the performance of SVM model [7]. Based on PSO-SVM, the paper proposes a short-term traffic flow predication model, PSO algorithm measure the superiority and inferiority of the i particle according to the fitness value. The optimal location named P_{id} in the history of each particle and the entire history of particle swarm optimal location named P_{gd} will be recorded in iterative calculation process. Mathematical description of PSO algorithm is as follows:

$$v_{id}=wv_{id}+c_1r_1(p_{id}-x_{id})+c_2r_2(p_{gd}-x_{id}) \tag{3}$$

$$x_{id}=x_{id}+v_{id} \tag{4}$$

Where: v_{id} is speed of the i particle in the d dimension, x_{id} is the location of the i particle in the d dimension, r_1、r_2 is uniformly distributed random number between zero and one, c_1 and c_2 are the learning factor, and w is inertia weight .

3. Method of calculation

For the deep tunnel project, due to various factors is complex and changeable, in this paper stress of surrounding rock, deformation and other action-effect are obtained by FLAC3D software. The software can inject a special non-linear constitutive model, obtaining very strong function of random parameters.

A new method of reliability calculation combining PSO-SVM-Monte Carlo is presented, it is integration of intelligent method, numerical calculation and Monte Carlo principle, will integrate the three elements for the reliability analysis system for deep tunnel reliability analysis. The specific process is as follows:

(1) Uniform design method are used to select study sample size, designed the range of random variable interval [$\mu-3\sigma$, $\mu+3\sigma$], μ and σ respectively are the average of all random variables and standard deviation. The establishment of engineering structures numerical analysis of the numerical model can provide the role of the structure effect (displacement, stress, etc.), as SVM learning samples.

(2) By using PSO to optimize the parameters of SVM, it can improve the accuracy of prediction and reduce the blindness of trial. The algorithm is summarized below:

① Initialize PSO, including the particle size, maximum velocity v, particle vector represents the parameters of SVM, which are penalty factor C, insensitive factor ε.

② Evaluate the fitness of each particle, compare the fitness value and determine the personal best particle and global best particle. Change the velocity for each particle according to Eq. (3) and then update the position for each particle according to Eq. (4).

③If it convergence, the global best particle is the optimized parameter.

(3) A sufficient number of random samples for Monte Carlo simulation can directly be provided, stress, displacement can be calculated through the random variable has been input trained SVM model. Finally, the effect of random variables and the role of engineering structures by substituting into the function of limit state function, statistic the number of function value which is less than 0, according to the Monte Carlo principle, the reliability of engineering structures can be calculated.

4. Verification of the method

Hypothesis the performance function of one structure is Z_1=18.46-7.48X_1/X_2, the distribution of parameter variable is X_1～N(10, 22), X_2～N(2.5, 0.3752), solving structural reliability index β.

The number of samples are selected as 15 accordance to uniform design(Table 1). Argument on behalf of the sample into the PSO-SVM model, training SVM, then the nonlinear mapping relationship between the function value and the variable is established. Stochastic simulation is 100,000 times, reliability calculation in the Table 2, comparison between JC and direct MC method with high precision, it shows that the method can be applied to reliability calculation of deep tunnel project.

Table 1 Samples selected by uniform design

Parameters	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X_1	11.6	15.7	14.9	10.7	8.2	4.8	13.	14	12.	7.3	4	9.8	5.7	6.5	9.0
X_2	1.5	2.6	1.8	3.0	3.5	3.2	2.2	3.3	2.9	1.9	2.3	2.1	1.6	2.8	2.5
Z_1	-7.2	11.9	-1.4	1.6	17	17	9.2	15.	14.7	10.7	16	10	8.8	16	14

Table 2 The calculated results and comparison

Method	JC	In this paper	MC
β	2.33	2.37	2.36

5. Case study

Taking Jinping Ⅱ hydropower station tunnel project as a example. The project is located in the Yalong River in Sichuan Province Jinping river bend(Fig.1), total installed capacity is 4800MW, through the excavation of four water diversion tunnel built power plants to generate electricity. There are two auxiliary tunnel, A and B. A hole cross-section width 5.5m, high 5.7m; B hole cross-section width 6.0m, clear high 5.0m. The maximum depth of the overlying strate of tunnel is about 2525m, and have large embedded depth, long tunnel axis, big diameter selection characteristics, it become the deep tunnel in the world. Depth greater than 1500m length is 73% of total length(Fig. 1).

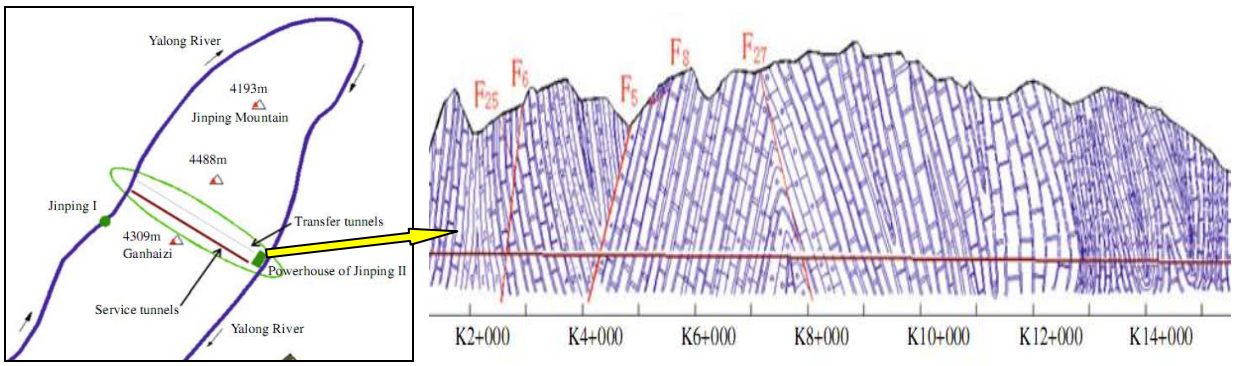


Fig.1 A position of Jinping II hydropower tunnels and assistant tunnels, b longitudinal section of assistant tunnels

The reliability of AK2+000, AK4+000, AK9+000, AK14+000, AK15+000 section were calculated, three random variables(E, C and Φ) are considered. The statistical characteristics of variables refer to field geological survey recommended range. Because of space limitations, just show the parameter distribution of AK9+000 (Table 3).

Table 3 Probability distribution of random variables

Parameters	Average	Standard deviation	Dispersion pattern	Unit
E	25	2.5	Normal	GPa
C	15	1.5	Normal	MPa
Φ	30	3.0	Lognormal	°

CWFS constitutive model in FLAC3D numerical calculation is adopt [8], the model can better simulate mechanical behavior of deep tunnel, in-situ stress field is: σ_{xx} =55MPa, σ_{yy} =70MPa,

σ_{zz} =65MPa, σ_{xy} =3.18MPa, σ_{yz} =8.9MPa, σ_{zx} =5.97MPa. Parameter setting: Poisson's ratio 0.2, peak friction angle value is two times the initial, residual cohesion is 10% of initial cohesion, critical plastic strain is 1e-3, the monitoring points and measurement line layout in tunnel show in Fig. 2.

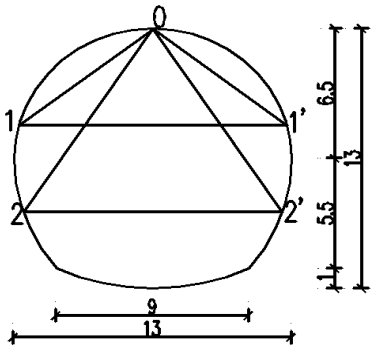


Fig.2 Line layout

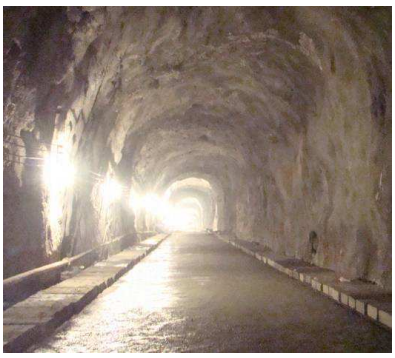


Fig.3 Stability of K15+000


Fig.4 Unstability of K9+000

Damage of rock is the changes of stress-strain relationship, but damage of surrounding rock are function failure, local destruction of surrounding rock does not mean that the destruction of the whole project. So, it is unreasonable that treating the rock failure criterion as performance function. Surface displacement as a predictor of the basic physical quantities, not only it is measured conveniently, more importantly, the measurement data is reliable, and also follows the corresponding standard norms. Therefore, this article uses the displacement of surrounding rock as stability criterion, according to the statistical 197 measurement section in 61 tunnel in China(on the basis of SUN Jun Academician), when relative convergence values are 2%, the tunnel is failure. Convergence 2% of the relative value of surrounding rock displacement is a criterion to judgment the surrounding rock stability limit. Performance function Z is showed in Eq. 5.

$$Z = [\lambda]_{\max} - \frac{1}{6}\sum_{i=1}^{6}\lambda_i \tag{5}$$

Where: $[\lambda]_{\max}$ is 2%, $\frac{1}{6}\sum_{i=1}^{6}\lambda_i$ is the average value of six line convergence in Fig. 2.

Thirty samples are generated by uniform design method, the convergence displacements are calculated by FLAC3D. Substituting into the calculation program, different reliability indexs of surrounding rock are got in different buried depth using the calculation method proposed in this paper(Table 4). If the numerical calculation times is 0.1 million only using FLAC3D software, the computation is very large, it will take several years to complete the calculation.

Table 4 Reliability indexs at different buried depth

Station	K2+000	K4+000	K9+000	K14+000	K15+000
Buried depth/m	1400	1650	2500	1700	1250
Rock classification	II	III	II	III	III
Lithology	Chloriteschist	Sandstone	Marble	Limestone	Cipolino
β	4.3	3.7	2.9	3.5	4.7

The security level of Jinping Ⅱ hydropower station is I level, category of destruction is B damage(sudden destruction), compared between the results and the current specification of the target(Table 5), when the reliability index is more than 4.2, it can not meet regulatory requirements. Fig. 3 and Fig. 4 are the A auxiliary tunnel's Rock-site conditions, K15+000 has overall stability of a good paragraph, but there are many landslides and other damage in K9+000, so the reliability evaluation index can provide the necessary reference and a basis for engineering design and construction. This paper presents a new way to characterize the reliability index obtained the safety level of the tunnel rock, it can reflect the actual situation of underground works, and can be certainty analysis of the expansion and additional. It is easier for people to accept.

Table 5 β

Security levels		I	II	III
Damage type	A	3.7	3.2	2.7
	B	4.2	3.7	3.2

6. Conclusion and Discussion

(1) The performance function of deep tunnel in the reliability analysis can not be explicitly expressed and has non-linear characteristics, in this paper, a new way of reliability analysis is proposed combining intelligent methods and principles of Monte Carlo, it overcomes some disadvantages of the traditional calculation method such as poor precision and low efficiency.
(2) In this method, the uniform design method are used to construct learning samples, and the software FLAC3D numerical are used to calculate effects of deep tunnel project, and then use SVM model to build the non-linear mapping relationship between the sample parameters and the structural effects, and use PSO to optimize parameters of SVM model, a non-linear approximation implicit performance function is simulated in the process of Monte Carlo.
(3) Reliability index which indicates that the safety of surrounding rock of deep tunnel can further reflect the actual situation of underground project. It is expanation and complement of deterministic evaluation methods, it is easier for people to accept.
(4) The study of uncertainty and spatial variability of surrounding rock parameters was not enough, hindered the application of reliability analysis method in deep tunnel, it needs to do more widely research work.

Acknowledgment

This work is supported by Reserch fund of State Key Laboratory of Geohazard Prevention and Geoenvironment Protection(Chengdu University of Technology), under grant NO. SKLGP2009Z015, and supported by Scientific Reserch Fund of SiChuan Provincial Education Department(NO. 09ZC026).

This work is also supported by the Open Research Fund of State Key Laboratory of Geomechanics and Geotechnical Engineering, Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, under grant NO. Z010909, and supported by National Natural Science Foundation of China under Grant No. 41002110.

References

[1] J. M. Duncan: Factor of safety and reliability in geotechnical engineering[J]. Journal of Geotechnical and Geoenvironmental Engineering，Vol. 126 (2000)，p. 307-316.
[2] K. K. Phoon: Reliability-based design in geotechnical engineering： computations and applications[M]. [S. l.] (2008).
[3] B. K. Low: Reliability analysis of rock slopes involving correlated nonnormals[J]. International Journal of Rock Mechanics and Mining Sciences，Vol. 44 (2007)，p.922-935.
[4] S. P. Huang， B. Liang and K. K. Phoon: Geotechnical probabilistic analysis by collocation-based stochastic response surface method-an EXCEL add-in implementation[J]. Georisk，Vol. 3(2009)，p. 75-86.
[5] S. Tamimi，B. Amadei and D. M. Frangopol: Monte Carlosimulation of rock slope reliability[J]. Computers and Structures，Vol. 33(1989)，p. 1495-1505.
[6] C. L. Huang, J. F. Dun: A distributed PSO–SVM hybrid system with feature selection and parameter optimization[J]. Applied Soft Computing，Vol. 10(2008)，p. 1381-1391.
[7] H.Y. Xiang and S.Y. Zhu:Optimal Time Interval and Lookback Interval of Short-term Traffic Flow Prediction Under Free Traffic condition[J]. Journal of Wuhan University of Technology (Transportation Science & Engineering) , Vol. 30(2006)，p. 326-628.
[8] R. S. Read: 20 years of excavation response studies at AECL's Underground Research Laboratory[J]. Int. J. Rock Mech. and Min. Sci, Vol. 41(2004)，p. 1251-1275.

doi:10.4028/www.scientific.net/AMM.50-51.869

Producing Mechanism and Distribution Laws of Remote Cracks for Geotechnical Engineering Structure

Houquan Zhang[1, a], Xinggen Huang[2, b], Libing Xue[2, c], Yanfeng Zhang[1, d]

[1]State Key Laboratory for Geomechanics and Deep Underground Engineering, School of Mechanics and Civil Engineering, China University of Mining and Technology, Xuzhou 221008, P.R. China

[2] Yuncheng Coal Mine, Heze city, Shandong 274718, P.R. China

[a]52zhq@163.com, [b]huangxinggen@yahoo.cn, [c]lbxue@126.com, [d]yfzhang@cumt.edu.cn

Keywords: Producing mechanism, Distribution laws, Remote cracks, Lateral confinement.

Abstract. The purpose of this paper is to investigate producing mechanism and distribution laws of remote cracks for geotechnical engineering structure. The fracture modes of geotechnical material containing a hole with different lateral confining pressure coefficients of 0, 0.05, 0.1, 0.15, 0.25, 0.33, 0.75 and 1 were numerically simulated using RFPA^{2D} (rock failure process analysis) code. The results indicate that under a certain confining pressure, three types of cracks, i.e., primary crack, remote crack and shear crack, can be observed simultaneously in the same sample. When the lateral pressure coefficient is satisfied that $0<k<0.33$, the remote cracks in en echelon forms occur in the remote region from the hole. The minimum distance of remote crack from the hole boundary is more than 0.5r ("r" denotes the radius of hole). Remote cracks are resulted by the connection of many micro vertical tensile cracks. The distribution laws of remote cracks are followed by the modulus $\left(\sqrt{2}\right)^{i-1}$ r (i=1, 2, 3, 4). A primary tensile crack is one necessary condition for the occurrence of remote cracks for circular geotechnical engineering structure.

Introduction

Geotechnical material is a kind of common material in civil engineering. Due to its special formation and embedment conditions, it is always heterogeneous and under triaxial compression of in-situ stress. It is known that the fracturing of geotechnical material is a progressive process and the unstable failure may come into different patterns due to the difference of fracturing process. Although the great difference exists, the fracturing always comes forwards to form the best loading structure whose stress concentration is the minimum, for example, the circular frame.

With the deep exploitation of mineral resources, the complex geological environment and high in-situ stress make the surrounding rock mass be fractured again and again. Therefore, a great deal work on the fractures of material surrounding a circle hole has already been done to study the failure of underground engineering structure and general results [1-6] are obtained as follows: (1) primary cracks, which appear in tensile stress concentration zone at the top and bottom of the hole; (2) remote cracks, which come forth in the area remote from the hole; (3) shear cracks, which can be observed in shear stress concentration region along hole sidewall. The distribution of three kinds of cracks is shown in Fig. 1.

Although a lot of field observations, experimental and numerical research results show the existence of remote cracks and great dangerousness for the engineering safety, there seldom exists reasonable theoretical explanation for the producing mechanism the distribution laws of remote cracks so far [6]. Only when these two aspects of remote cracks are well-known by geotechnical technologists, the reasonable and economical support can be designed. In the present study, this question will be discussed using numerical analysis code, called RFPA^{2D} (Rock Failure Process Analysis) again, which may be very helpful for understanding the remote cracks of rock surrounding a tunnel.

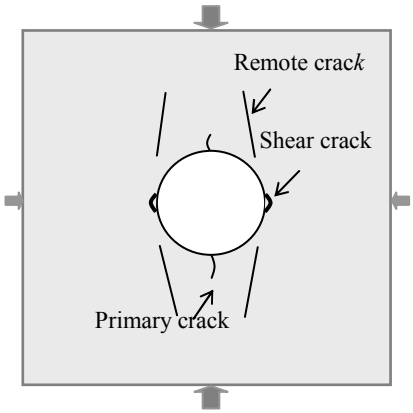


Fig. 1: Three types of cracks

Brief introduction of RFPA^{2D}

Herein, Rock Failure Process Analysis code (abbreviated as RFPA^{2D}) [7] are introduced and adopted for the failure analysis of material surrounding a circular hole. This numerical method of RFPA^{2D} has also validated by the reasonable agreement between Hoek's experiments and numerical results for the failure process and fracture mode of material surrounding a circular hole [6].

The basic elements of RFPA^{2D} model can be generalized as follows [8]: (1) by introducing heterogeneity of rock material properties into the model, RFPA^{2D} can simulate nonlinear deformation of a quasi-brittle rock with an ideal brittle constitutive law for local material; (2) by introducing a reduction of material parameters after element failure, RFPA^{2D} can simulate strain-softening and discontinuum mechanics problems in a continuum mechanics mode; (3) by recording the event-rate of failed elements, RFPA^{2D} can simulate seismicities associated with progressive fracture process. For more detailed descriptions of RFPA^{2D} model, please refer to related published literature.

Numerical model descriptions

Numerical model, which is similar to Fig. 1, is abstracted from a horizontal circle practical tunnel. There just exists one difference that the bottom of numerical model is fixed so as not to move in the vertical direction. Eight models with different confining pressure coefficients (Ratio between horizontal / vertical applied stress, labeled as k in this paper) of 0, 0.05, 0.1, 0.15, 0.25, 0.33, 0.75 and 1 were conducted to model the progressive fracturing processes for surrounding rock of a tunnel.

The numerical model in two-dimensional square with a hole in the center was discretized into 240×240=57,600 elements. The geometry of the block is 150mm × 150 mm in size while the diameter of the hole is 35 mm.

On account of the heterogeneity of rock material, the element local mechanical parameters such as uniaxial compressive strength σ_c, elastic modulus E_0, etc. (as listed in Table 1) are assumed to follow Weibull's distribution [9] which is defined as follows:

$$\phi(\alpha,m)=\frac{m}{\alpha_0}\left(\frac{\alpha}{\alpha_0}\right)^{m-1}\exp\left[-\left(\frac{\alpha}{\alpha_0}\right)^{m}\right] \tag{1}$$

Where α is a certain single element parameter, α_0 is the mean value of the element parameter and m is the shape parameter of distribution function which is referred to as homogeneity index [7]. A larger m implies a more homogeneous material and vice versa. Element failure is judged by a Coulomb criterion envelope with a tension-off, which is called as revised-Coulomb criterion [10]. When stress in an element satisfies with this failure criterion, the element becomes damaged either in tension or shear mode. A load with a value of 5 MPa is firstly applied in the vertical and then an increment with a value of 1 MPa is added on it step by step while the initial value and increment of horizontal lateral pressure are assigned according to confining pressure coefficient k.

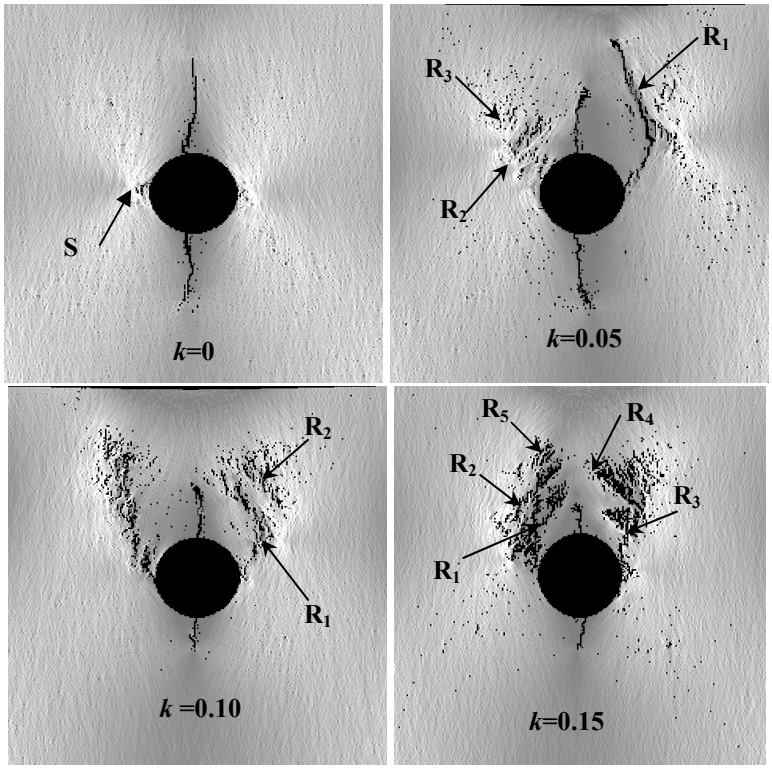
Table 1. Material mechanical property-mean value

Parameter	value
Homogeneity index (m)	4
Poisson's ratio (μ)	0.25
Elastic modulus (E_0)	60,000 (MPa)
Compressive strength (σ_c)	200 (MPa)
Tensile strength (σ_t)	20 (MPa)
Friction angle (φ)	30°
Tension cut-off	10%
Residual length coefficient (λ)	0.1

Numerical results of failure patterns

The shear stress distribution of surrounding material around a hole with different confining pressure coefficients are shown in Fig. 2. The degree of brightness in the picture indicates the value of shear stress. The brighter the element appears, the larger the stress is. In the Fig. 2, "S" denotes the shear crack. "R_i" denotes the remote cracks far away from the hole boundary and the subscription represents the time sequence of occurrence, that is to say, remote crack "R_2" occurs after remote crack "R_1".

From the Fig.2, it can be found that ① the primary tensile cracks occurs in the roof and floor part of hole when the lateral pressure coefficient k is less than 0.33; ② when $0<k<0.33$, after the primary tensile cracks propagate for a certain distance, the remote cracks occur in the remote region from the hole. The minimum distance of remote crack from the hole boundary is more than 0.5r ("r" denotes the radius of hole). The remote cracks are arrayed in en echelon forms, propagating inside into the deep intact part of rock mass. ③ when $0.33 \le k < 1.0$, no tensile cracks occur in the roof and floor while dog-eared shear cracks occurs on two sides of hole, the high concentration zones of shear stress. Then the shear cracks develop into the deep part of rock mass and the hourglass structure is formed after the whole failure; ④ when k=1.0, dog-eared shear cracks happen all around the hole.



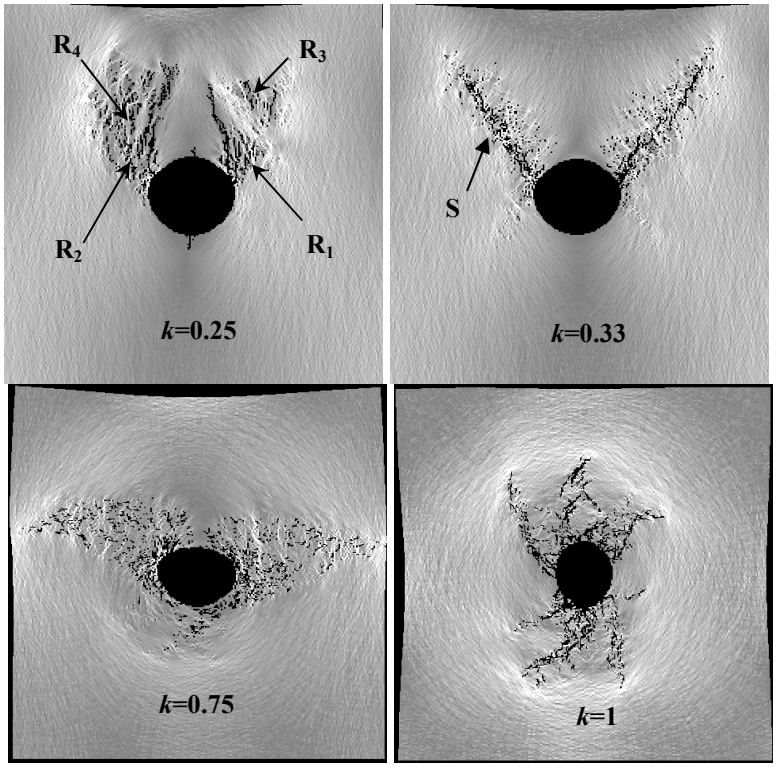


Fig. 2: Numerical results of failure patterns under different lateral pressure coefficients

Discussion on producing mechanism and distribution laws of remote cracks

Remote cracks are resulted by the connection of many micro vertical tensile cracks and some rather intact parts exist between micro tensile cracks, denoted by the arrowhead in the Fig. 3. Therefore, the remote cracks are considered to be induced by tensile stress; however, several micro cracks parallel to the shear stress plane exist. So the failure surface of remote cracks is coarser than that of shear fracture. Consequently, it can be concluded that the remote cracks are a kind of tensile-shear cracks. Lajtai et al. [2] had named it as secondary tensile and inclined shear cracks according to their experiment results.

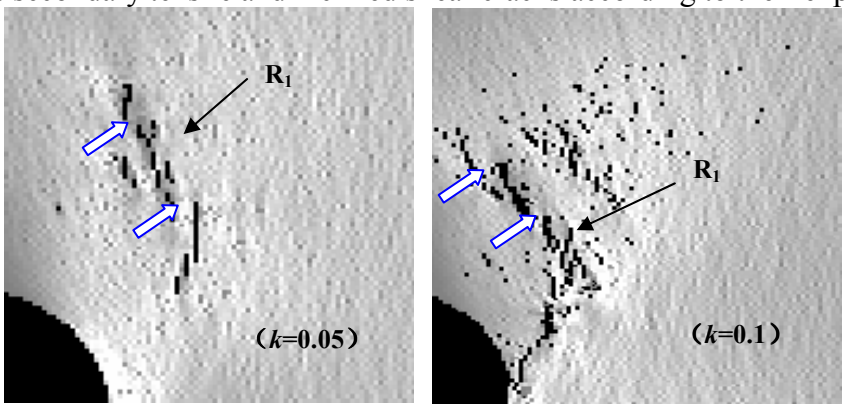


Fig. 3: Local magnification of remote cracks

Table 2 summarizes the geometrical distribution parameters of remote cracks under different lateral confinement. Fig. 4 represents the sketch map of the distance of remote cracks from the circular hole boundary.

It can be found from Fig. 4 that the distance of remote cracks from the hole center are about 1.5r, 2r, 2.8r except the models with lower lateral confinement of k=0.05 and k=0.1.

It is interesting that Курленя М.В. and Опарин В.Н. [11] also found that the location of disintegration radii through in-situ field observation are followed by the distribution laws of $\left(\sqrt{2}\right)^{i-1}$r (i=1, 2, 3, 4) ("r" denotes the radius of hole). Li et al [12] and Xu et al. [13] also found the distribution results of zonal disintegration by the method of boring television imager and electrical resistivity. But

Table 2. Geometrical distribution parameters of remote cracks under different lateral pressures

Geometrical parameters \ Lateral pressure coefficients	k=0.05	k=0.1	k=0.15	k=0.25
d_1: the distance of remote crack R_1 from the hole center	1.98r *	2.02r	1.53r	1.57r
d_2: the distance of remote crack $R2_1$ from the hole center	1.92r	2.71r	2.19r	1.70r
d_3: the distance of remote crack R_3 from the hole center	2.64r		1.61r	2.18r
d_4: the distance of remote crack R_4 from the hole center			2.14r	2.28
d_5: the distance of remote crack R_5 from the hole center			2.86r	
β_1: the angle of remote crack R_1 deviated from vertical direction	22.5 º	38.6 º	45º	38.6 º
β_2:the angle of remote crack R_2 deviated from vertical direction	45º	45 º	45º	26.5 º
β_3: the angle of remote crack R_3 deviated from vertical direction	45º		45º	45º
β_4: the angle of remote crack R_4 deviated from vertical direction			45º	45º
β_5: the angle of remote crack R_5 deviated from vertical direction			45º	

* r denotes the hole radius

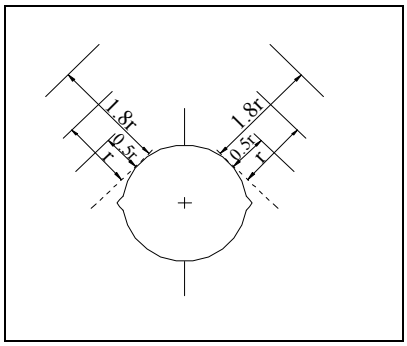


Fig. 4: Sketch map of distribution location of remote cracks

what is reason for this special distribution laws, it has not be known so far. Rock scholars should do more research on it.

From Fig. 2, it can also be found that primary tensile crack is one necessary condition for the occurrence of remote cracks for circular geotechnical engineering structure. When the tensile cracks occur, remote cracks formed later in the remote tensile and shear stress regions under a certain confinement.

Conclusions

Through the numerical simulations on fracture modes of geotechnical material containing a hole under different lateral confinements, it can be found that:

① Under a certain confining pressure, primary crack, remote crack and shear crack can be observed simultaneously in the same sample.

② When the lateral pressure coefficient is satisfied that $0<k<0.33$, the remote cracks in en echelon forms occur in the remote region from the hole.

③ Remote cracks are induced by the connection of many micro vertical tensile cracks. The distribution laws of remote cracks are followed by the modulus $\left(\sqrt{2}\right)^{i-1}$ r (i=1, 2, 3, 4) ("r" denotes the radius of hole).

Acknowledgements

This study was supported by NSFC (Nos: 50804046, 50774082 and 50490273), National Basic Research Program of China (2007CB209400) and Scientific Research Fund for youths of CUMT (No: 0B080240).

References

[1] Hoek, E. and Brown, E.T. Underground Excavations in Rock. *London: The Institution of Ming and Metallurgy*, 1980.

[2] Lajtai, E.Z. and Lajtai, V.N. The collapse of cavities. *Int. J. Rock, Min. Sci. & Geomech. Abst.*, 12(1) (1975) 81-86.

[3] Carter, B.J. and Lajtai, E.Z. Stress and time-dependent fracture around cavities in physical models of Potash salt rock. *In Hudson J. A. ed. Proc. ISRM Symposium: Eurock'92, Chester, Rock characterization, Imperial College, London,* (1992) 269-274

[4] Martin, C.D. Seventeenth Canadian geotechnical colloquium: the effect of cohesion loss and stress path on brittle rock strength. *Can. Geotech. J.*, 34(5) (1997) 159-168

[5] Fu, Y.F., Huang, M.L., Ren, F.Y. and Tang, C.A. Numerical analysis of crack evolution around borehole in rock sample subjected to confining pressures. *Chinese Journal of Rock Mechanics and Engineering,* 19(5) (2000) 577-583

[6] Zhang H.Q., Zhao D.S., et al. Numerical Study on Fracture in Rock Surrounding a Circular Tunnel under Different Confining Pressure, *In: Ohnishi Y, Aoki K, eds. Contribution of Rock Mechanics to the New Century. Rotterdam: Millpress*, (2004) 537-540.

[7] Tang, C.A. Numerical simulation on progressive failure leading to collapse and associated seismicity. *Int. J. Rock Mech. and Min. Sci.* 34(1997) 249-261.

[8] Tang, C.A., Liu, H, Lee, PKK, Tsui, Y. and Tham L.G. Numerical studies of the influence of microstructure on rock failure in unaxial compression-part I: effects of heterogeneity. *Int. J. Rock Mech. Min. Sci.,* 37(2000) 555-569.

[9] Hudson, J.A. and Fairhurst C. Tensile strength, Weibull's theory and a general statistical approach to rock failure. *The proceedings of the Civil Engineering Materials Conference, Southampton* (1969) 901-904.

[10] Brady, B.H.G. & Brown, E.T. *Rock Mechanics for Underground Mining,* Second Edition. London: Chapama & Hall Press, (1993).

[11] КУРЛЕНЯ М. В., ОПАРИН В.Н. Проблемы нелинейной геомеханики. *Ч. I. ФТПРПИ*, 3(1999) 12-23.

[12] Li S. C., Wang H.P., Qian Q.H., et al. In-situ monitoring research on zonal disintegration of surrounding rock mass in deep mine roadways. Chinese Journal of Rock Mechanics and Engineering, 27(8)(2008) 1545-1553.

[13] Xu H.F., Qian Q.H., Wang F.J., et al. Application of electric resistivity method to zonal disintegration exploration of deep roadway. Chinese Journal of Rock Mechanics and Engineering, 28(1)(2009) 111-119.

Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.875

Detection of Damage Extension in Cantilever Beams Using Change Ratio of Frequency Response Functions

Hailei Jia [1, a], Yin Zhao [1,b]

[1]Department of Engineering Mechanics,HoHai University,Nanjing 210098, China

[a] jhl1024@hhu.edu.cn, [b] zhaoyin01@hhu.edu.cn

Keywords: change ratio of frequency response functions, damage extension, damage detection, experimental modal analysis, cantilever beams.

Abstract. Frequency response function (FRF) is a fundamental dynamic index, which is capable of reflecting structural dynamic properties using full-spectrum information. In spite of distinct merits over conventional modal parameters, the FRF has an observable drawback of multi-dimensionality, unsuited for damage characterization. Such a situation motivates an interesting subject, i.e., extracting low-dimensional, high-sensitivity damage index from the FRF. This study focuses on developing a valid damage index, called FRF change ratio, to detect extension of damage. An experiment towards cantilever beams is systemically conducted. The results show that the FRF change ratio can effectively reflects damage extension, and it is more sensitive than conventional natural frequencies. This new damage index holds promise for practical damage detection in beam-like structures.

Introduction

Structural health monitoring and damage detection have become a research foucs in the last decade in civil community. Detection of damage extension is of important significance for avoiding catastrophic incidents[1]. Nowadays, the advanced technologies of sensing, signal processing and dynamic system identification provides a chance to develop new methods for detection extension of structural damage. Damages may change structural physical properties (mass, stiffness and damping), resulting in changes in structural dynamic properties[2]. Frequently used damage indices include FRF, natural frequency, mode shape, strain mode shape, curvature mode shape, modal assurance criterion (MAC) and coordinate modal assurance criterion (COMAC), et al.[3]. Among these features, the most fundamental one is the FRF, from which the other damage indices are yielded with some reduction of damage information. Natural frequency is a conventional damage feature with the major advantages of high precision and lower cost in acquirement and disadvantages of insensitivity to slight damage[4-5]. Joshi [6] explored the relation between natural frequency and structural damage under various boundary conditions. Salawue [7] provided a comprehensive review on existing theories and methodologies for use of natural frequency in damage identification.

The FRF carries more damage information than natural frequency. Schulz M J [8] and Park N G [9] successfully detected the location of a damage in beams by using FRF; Z S Wang [10] performed detection on multiple damages in beams based on the FRF. In spite of more damage information in FRF than other damage indices, multi-dimensionality is still an obstacle decreasing its applicability in damage identification applications. This situation poses an interesting research subject, i.e., how to extract low-dimensional, high-sensitivity damage features from FRF. The goal of this study is to establish a damage index derived from FRF for reflecting the extension of damage. An experiment of detecting extension of a crack in cantilever beams is conducted in detail to validate the damage index proposed.

Formulation of Damage Index

With a complex harmonic force $\mathbf{f}e^{j\omega t}$, the equation of motion for a multi-degree of freedom (DOF) is given as

$$M\mathbf{x} + C\mathbf{x} + K\mathbf{x} = \mathbf{f}e^{j\omega t} \quad (1)$$

where M, C and K denote the mass, damping and stiffness matrices, respectively, x is the dynamic response. Represent $\mathrm{x}(t) = \mathrm{u}e^{j\omega t}$. Substituting x into Eq. (1) yields

$$(K - \omega^2 M + j\omega C)\,\mathbf{u} = \mathbf{f} \quad (2)$$

Solving Eq. (2) gives

$$\mathbf{u} = (K - \omega^2 M + j\omega C)^{-1}\,\mathbf{f} \quad (3)$$

Get FRF expression from Eq. (3):

$$H(\omega) = (K - \omega^2 M + j\omega C)^{-1} \quad (4)$$

Clearly, $H(\omega)$ is related to K, M and C. This fact implies the change in structural properties necessarily alter the FRF and therefore the alternation of FRF can be used to identify the change of K, M and C.

Generally, assume damage alters structural stiffness from K to $(K-\delta)$, leading to

$$H^d(\omega) = ((K-\delta) - \omega^2 M + j\omega C)^{-1} \quad (5)$$

where the superscript *d* indicates the condition of damage.

The effect of damage is represented as

$$\Delta H(\omega) = \left|H(\omega) - H^d(\omega)\right| \quad (6)$$

ΔH reflects damage over the continuous spectrum range of interest, including modal and non-modal information. A damage index, called FRF change ratio, is defined as:

$$D(\omega) = \frac{1}{N}\sum_{\omega=0}^{\infty}\left|\frac{\Delta H(\omega)}{H(\omega)}\right| \quad (7)$$

where N represents the number of spectrum line.

For experimental modal analysis, a FRF is expressed as:

$$H_{ij}(\omega) = \frac{X_i(\omega)}{F_j(\omega)} \quad (8)$$

where $F(\omega)$ and $X(\omega)$ are the Fourier transform of excitation and acceleration dynamic response, respectively.

In practical damage detection, a limited range of spectrum, $[0\,\bar{\omega}]$, is available. Correspondingly, the FRF change ratio is written as

$$D_{ij}(\omega) = \frac{1}{N}\sum_{\omega=0}^{\bar{\omega}}\left|\frac{H_{ij}(\omega) - H_{ij}^d(\omega)}{H_{ij}(\omega)}\right| \quad (9)$$

When damage extends progressively, the FRF change ratios at different damage stages are in the form

$$D_{ij}^k(\omega) = \frac{1}{N}\sum_{\omega=0}^{\bar{\omega}}\left|\frac{H_{ij}(\omega) - H_{ij}^{d,k}(\omega)}{H_{ij}(\omega)}\right| \quad (10)$$

where $H_{ij}^{d,k}(\omega)$ and $D_{ij}^k(\omega)$ are the FRF and the FRF change ratio at damage level *k*. $D_{ij}^k(\omega)$, $k = 1, 2, \cdots$, reflect the extension of the damage.

Experimental Procedure

A rectangular cantilever beam with a uniform section is used as experimental model. Each beam specimen of Steel Q235 is specified with Length (L) 0.6m, width 0.012m, height 0.019m, density 7750 kg/m^3, elastic modulus 1.84×10^{11} Pa, and Poisson ratio 0.3. Numerical-control wire cutting machine is employed to generate cracks with tungsten filament of 0.16mm. The actual cracks spread between 0.16mm to 0.18mm due to some uncertainties in the experiment. A single crack with various severities in a beam is considered to simulate the extension of the crack. Cracks with crack location

ratios (β) from the clamped end, 3.3%, 26.7%, 46.7%, 66.7% and 86.7%, are considered, and for each crack, the crack depth ratios (α) of 10%, 20%, 30% and 40% are considered.

The lengths of the fixed and cantilevered portions of the beam are 100 mm and 500mm, respectively. Hammer impulse imposed near the free end of the beam is carried out to generate acceleration responses. The acceleration responses are collected with sampling frequency 12800 Hz and 65536 sampling (5.12s). The resulting resolution of frequency 0.1953Hz, which is adequate for damage identification. The FRFs acquired are denoised using wavelet technique, which gives rist to denoising effect is illustrated in Figs. 3 and 4.

Fig. 1 Data collecting system

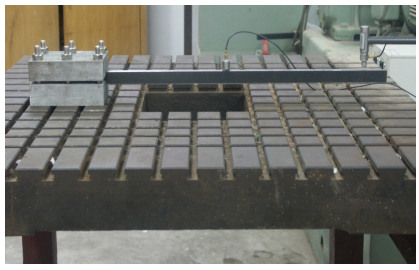


Fig. 2 Set-up for testing

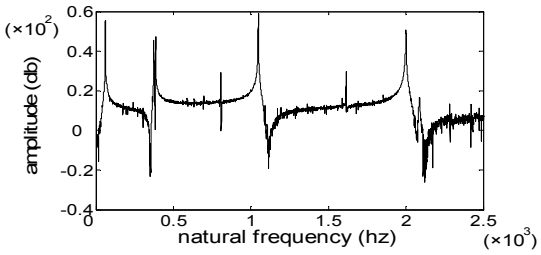


Fig. 3 FRF before denoising

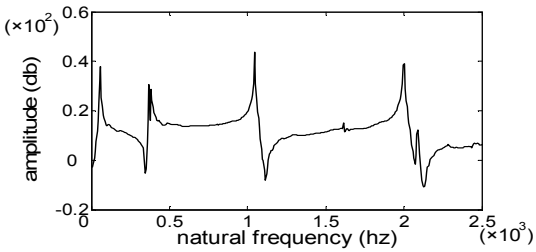


Fig. 4 FRF after denoising

Detection of Damage Extension

A sequence of FRF change ratio $D_{ij}^{k}(\omega)$ are used to depict extension of a crack. The supremum of ω is $\bar{\omega} = 2100\,\text{Hz}$, covering modes from order 1 to order 4. The variations of FRF due to progressive increase of crack depth ratios from $\alpha = 10\%$ to $\alpha = 40\%$ are illustrated in Figs. 5 and 6. Clearly, the crack extension can be characterized by the increased deviations between $H_{ij}^{d,k}(\omega)$ and $H_{ij}(\omega)$.

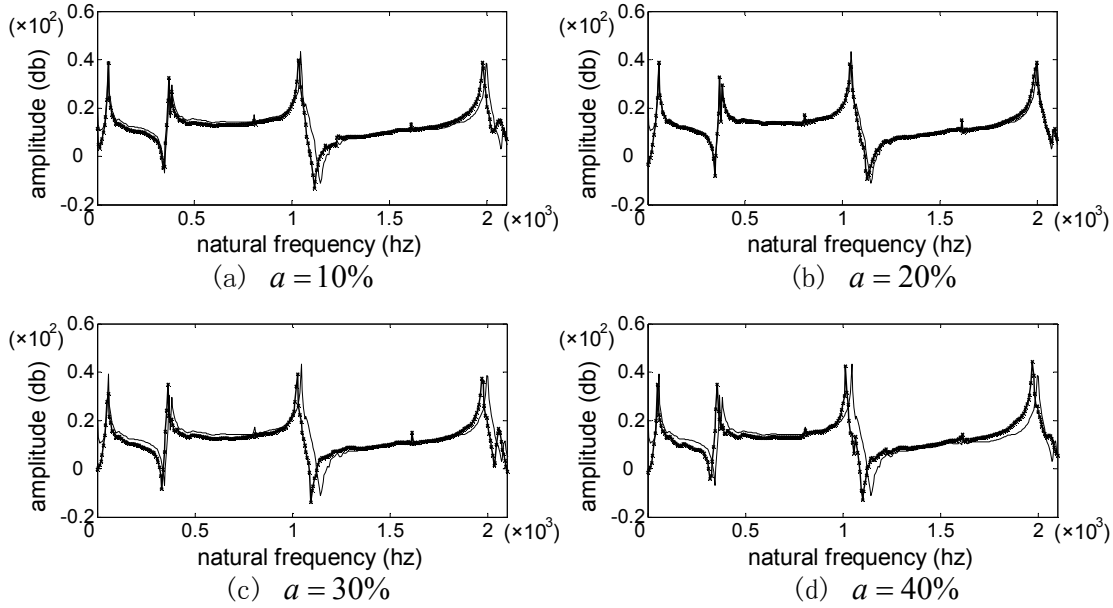


Fig. 5 Variations of of FRFs due to crack extension for crack case of $\beta_1 = 3.3\%$.

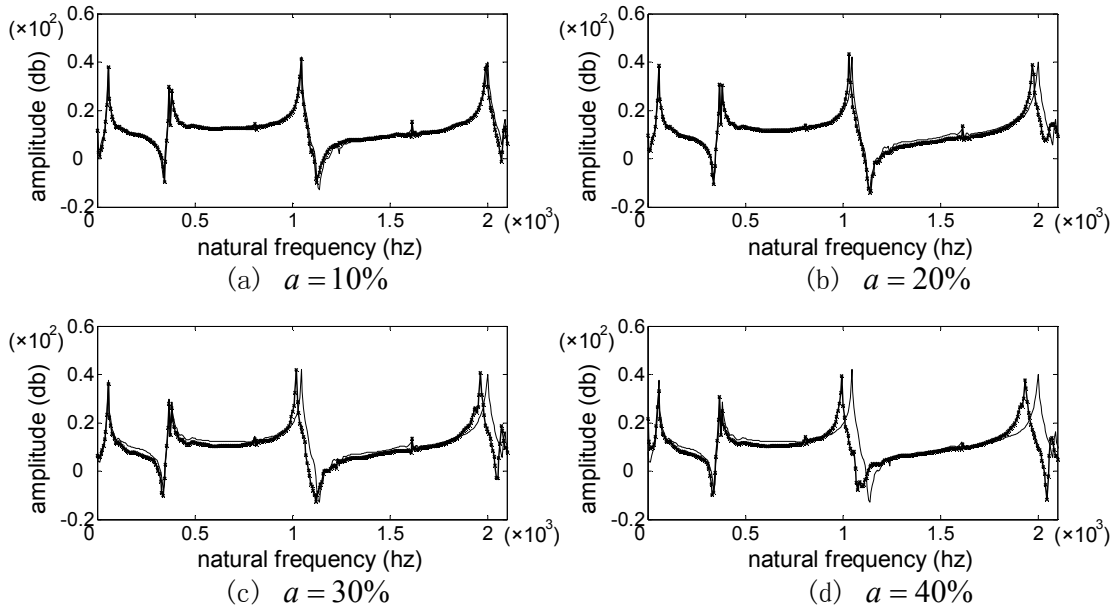


Fig. 6 Variations of of FRFs due to crack extension for crack case of $\beta_2 = 26.7\%$.

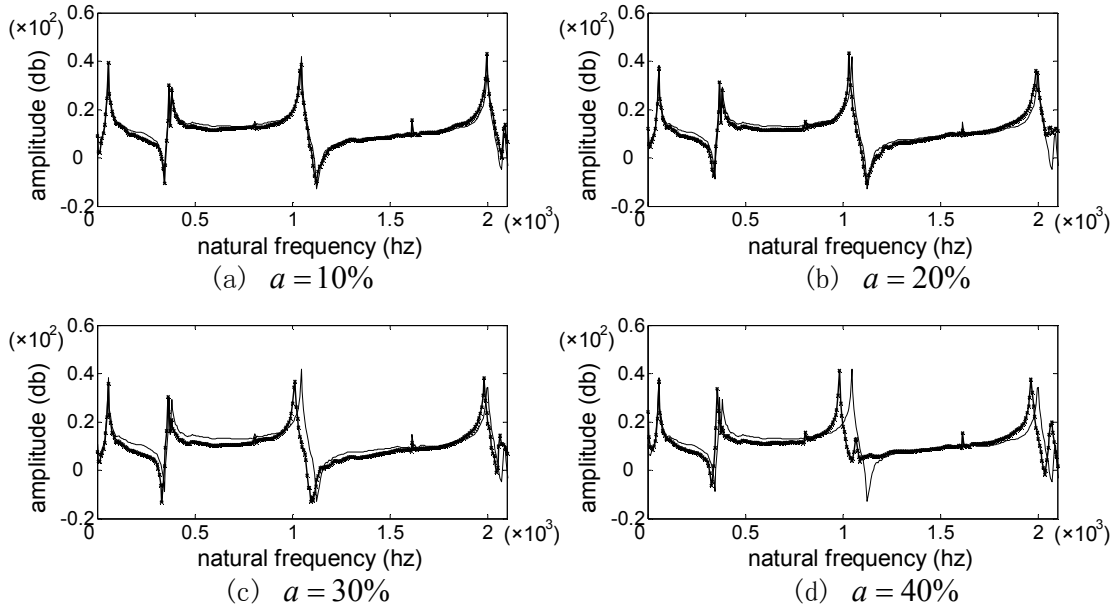


Fig. 7 Variations of of FRFs due to crack extension for crack case of $\beta_3 = 66.7\%$

The FRF change ratios corresponding to three representative crack locations are listed in Table 1. For a crack location, the FRF change ratios increase with the augment of the depth of a crack, demonstrating the effectiveness of FRF change ratio in quantitative characterizing the extension of a crack. In addition, for a crack extent, the FRF change ratios decrease with the distance of crack locations approaching the free end, implying the FRF change ratio is affected by the crack location.

For comparison, the change ratio of natural frequencies is given as

$$D_i^k(\omega) = \frac{1}{N}\sum_{\omega=0}^{\bar{\omega}} \left| \frac{\omega_i - \omega_i^{d,k}}{\omega_i} \right|, \ i = 1-4 .$$

Fig. 8 provides the performance comparison of two change ratio indices with respect to FRF and natural frequency. Obversably, the former outperforms the latter significantly.

Table1 FRF Change Ratios (%) for various damage cases

Cases	$a=10\%$	$a=20\%$	$a=30\%$	$a=40\%$
β_1=3.3%	28.7	51.5	72.7	118.6
β_2=26.7%	29.7	44.7	63. 1	109.4
β_3=66.7%	26.3	43.2	53.3	75.5

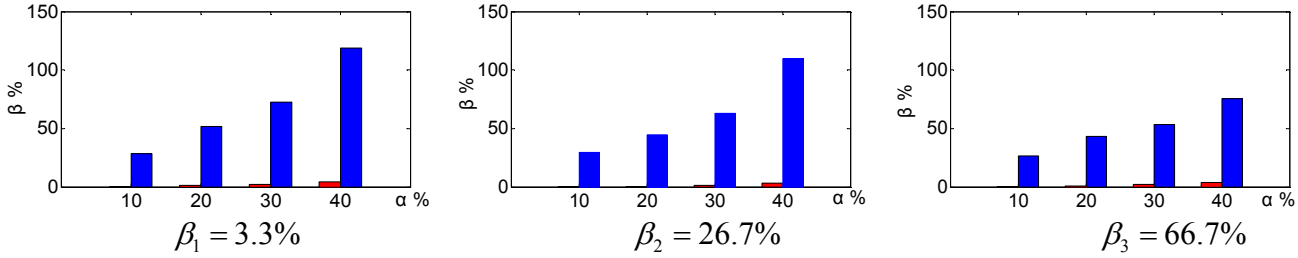


Fig. 8 Comparison of change rates for FRF and natural

4. Conclusions

In this study, a new damage index, FRF change ratio, is established and its characteristics are examined in the experiments on detecting crack extension. The results show that the FRF change ratio outperforms conventional natural frequency in characterizing crack. The FRF change ratio holds promise for practical applications in damage detection for beam-like structures.

Ackmowledgements

The authors are grateful for National Natural Science Foundation of China (Grant no.50978084 and no. 50539030).

References

[1] Ma Hongwei,Yang Guitong: Methods and Anvances of Structural Damage Detection. *Anvances in Mechanics*.1999, 29 (4), p.513-527.

[2] ZhengDongliang, Li Zhongfu, HuaHongxing: A Summary Review of Structural Initial Damage Identification Methods . *Journal of Vibrationand Shock.* 2002, 21 (2) , p.1-10.

[3] Zong Zhouhong,Ren Weixin,Ruan Yi: Recent Advances in Research on Damage Diagnosis For Civil Engineering Structures. *Civil Engineering Journal.* 2003, 36 (5), p.105 - 110.

[4] Shen Wenhao, Zhang Senwen: Study on Single Damage Detection Based on Experimental Modal Analysis . *Journal of Jinan University(Natural Science).* 2006, 27(3) ,p.420-424.

[5] Adams R D , Cawley P , Pye C J: Vibrational technique for non destructively assessing the integrity of structure. *Journal of sound and vibration*. 1983,(91), p.251-262

[6] Joshi A,Madhusudhan BS:A unified approach to free vibration oflocally damaged beams having various homogeneous boundary conditions.*Journal of Sound andVibration.*1991,147,p.475-488.

[7] Salawu O S: Detection of structural damage through changes in frequency : *A review. Eng. Struct.* 1997 ,19 (9) , p.718-723

[8] Schulz M, Naser A S: Locating structural damage using frequency response functions . *Journal of Intelligent Material Systems and Structures*. 1998, (9) , p. 899-905.

[9] Park G, Cudney HH, Inman DJ: An integratiod health monitoring technique using structural impedance sensors . *Journal of Intelligent Material Systems and Structures.* 2000, 11,p. 448-455.

[10] Wang Zhisong , Guo Huiyong , Li Zhengliang: Identification Methods for Different Structural Damage . *Engineering Mechanics*.2008,25(6), p.6-13

doi:10.4028/www.scientific.net/AMM.50-51.880

A Level Set Method for Image Restoration

Waibin Huang[a], Weiran Lin[b]

School of Computer and Computing Science, Zhejiang University City College, Hangzhou 310015, China

[a] waibin@163.com, [b] (Corresponding author) Linweiran1982@163.com

Keywords: Level Set, Geodesic Curve, Occlusion, Image Restoration

Abstract. This paper presents a level set method for image restoration. In the light of the loss of image quality caused by occlusion or part stain, it adopts the method of picking-up image level sets and filling in level sets of occlusion to reconstruct the image. In the process of linking level lines, besides the traditional geodesic curves, it makes use of the Meaningful beeline detection technique. The experiment results show this method is of great importance in the preservation of images.

1 Introduction

In the process of image production and transmission, we usually can't get the expected image quality due to the imperfect equipments and physical limitation. For example, the relative movement between photography equipment and the objective as well as the influence of air current during aviation shooting can both lead to the obscurity of image. And the long period in the management of files and preservation of photos can also cause loss of image quality. This kind of phenomena often occurs in some fields of applied science and engineering. Therefore restoration of the original image is extremely important.

In image restoration, it is necessary for us to find the reason of image damage so as to restore the original image. Generally speaking, image damage is caused by noise, and in this case, the restoration is mainly to denoise. Many mature linear filter wave techniques have been used in denoising and restoring image. However, the pure linear techniques can't get satisfactory results. Though the filter can effectively get rid of noise, the border will be blurred. As a way of improvement, the non-linear techniques have done a good job. The Wavelet and Partial Differential Equation are the commonly-used methods.

In another commonly-seen circumstance, the image is partly occluded by other objects, and at that time, restoration means disocclusion, which is frequently applied in file management and old picture storage [1]. Part stain may occur due to the long period of storage. However, the research in this aspect has just begun. AMLE [2] (Absolute Minimal Lipschitz Extension) proposed by V. Caselles has achieved better results in image restoration.

This paper introduces an image disocclusion method based on level set. The level set method [3, 4] proposed by S. Osher has good geometry interpolation. To transform gray interpolation to curve interpolation, the method has great influence in image processing. And it has achieved good results in image restoration and denoising, edge detection, and image recognition. This paper analyzes the level set of the occluded image and then it adopts the Meaningful beeline detection technique [5] and geodesic curve [6] to connect the broken level lines so as to restore the image. The experiment results prove this method has good effect in image restoration and is applicable in file management and picture preservation.

2 Image Model and Level Set

2.1 Occluded Image Model

Assume $u_0(x)$ is the original image without occlusion, Ω is the occluded part in the image, $u(x)$ is the occluded image, and $u_\Omega(x)$ is the occluding object, then

$$\begin{cases} u(x) = u_{\Omega}(x) & if \quad x \in \Omega \\ u(x) = u_0(x) & otherwise \end{cases}. \tag{1}$$

As shown in Fig. 1.

(a) (b)

Fig. 1 image model (a) original image($u_0(x)$) (b) occluded image($u(x)$)

What it should be done now is to start from $u(x)$ and try to restore the image so as to make the final image resemble the original image $u_0(x)$.

2.2 Limitation of Ω

In this paper, Ω is given because in the file preservation the occluded part has distinct features, which can be spotted by some chemical and physical methods in file preservation. We hope the occluded part shouldn't be too complicated, such as the finitely perimeter, otherwise, we can select a set $E \supset \Omega$, which meets the condition: $\forall \delta > 0, d_H(E, \Omega) < \delta$, and d_H is hausdorff distance. Then, we can select E instead of Ω. Therefore, we can assume that Ω is open, bounded and simply connected and we denote by $\partial\Omega$ its boundary which we assume to be a Jordan curve. There exists a simple rectifiable curve Γ arbitrarily close to Ω such that the one-dimensional restriction $\tilde{u}$ of u to Γ has bounded variation and such that the level lines of u are transverse to Γ, namely $\int_{\Gamma_h} |\tilde{u}| dH^1 < \infty$, and H^1 is the hausdorff measure. Let x be a boundary point in Γ, we can define an average direction

$$v_\lambda(x) := v_{\lambda,B} = \frac{1}{H^1(\partial\chi_\lambda \cap B)} \int_{\partial\chi_\lambda \cap B} v\chi_\lambda dH^1. \tag{2}$$

Where $B = B(x, r_0), r_0$ is such that $d(\Gamma, \Omega) > r_0$ and $v\chi_\lambda$ denotes the outer normal at every point of the reduced boundary.

2.3 Level Set and Disocclusion

As to image $u(x)$, $x \in R^2$, we define its λ-level set as a set, $\chi_\lambda u = \{x : u(x) \geq \lambda\}$ whose boundary is level line. All the level lines of a function completely describe its information, which can be illustrated by $u(x) = \sup\{\lambda : x \in \chi_\lambda u\}$.

When the original image is partly occluded by other objects, its level lines will break up in the boundary of occluded area Ω. Disocclusion is to connect the broken level lines in Ω. The level lines in this area can be got according to the level lines in the boundary. What we will do now is to connect the level lines of the same value without intersection. That is to fill in the occluded area with level lines.

3 Energy Formula

There are several different solutions in connecting the level lines. In order to get the unique solution which resembles the original image most, namely the optimum solution, we set the minimum energy on boundary curves $\partial\Omega$ as a standard. M. Nitzberg [7] puts forward an energy formula, $E(u)=\int|Du|(1+|curv(u)|^p)$, $p\geq 1$. We adjust E and make it accord with the situation in this paper. We need a disocclusion which resembles the function of the original image at most. In fact, according to disintersectant set of boundary interpolation level lines, we can easily reconstruct a single function u, whose level lines coincide with the boundary almost everywhere.

In order to solve the disocclusion problem we need to distinguish the cases when $p=1$ and $p>1$. If $p>1$, the curvature of γ at t is given by

$$curv(\gamma(t))=\frac{\det(\gamma'(t),\gamma''(t))}{|\gamma'(t)|^3}. \tag{3}$$

and its absolute value can be written with respect to arc-length as

$$|curv(\gamma(s))|=|\gamma''(s)|. \tag{4}$$

In order to deal with those curves of null length whose endpoints coincide on Γ, namely $\gamma(0)=\gamma(1)\in\Gamma, \left|\frac{d\gamma}{dt}\right|=\left|\frac{d^2\gamma}{dt^2}\right|=0$ everywhere in (0, 1), we can define the energy

$$E_p(\gamma)=\begin{cases}\int_0^{l(\gamma)}(1+|\gamma''(s)|^p)ds & if\ \ l(\gamma)>0\\ \langle\tau(0),\gamma'(0)\rangle+\langle\tau(1),\gamma'(1)\rangle & otherwise\end{cases}. \tag{5}$$

where $\tau(0)=\nu_\lambda(\gamma(0))^\perp, \tau(1)=\nu_\lambda(\gamma(1))^\perp$, $l(\gamma)$ is the length of curve γ.

The case $p=1$ must be treated apart since the space $W^{1,1}$ lacks of fair properties of the space $W^{1,p}$. In order to make up these shortcoming, we can adopt the following method: $\forall x\in\partial\Omega$, we define the line of level λ arriving at x as the segment $S(x,\lambda)$ of length $\alpha<<1$ making the angle $\tau(0)$ with Γ at x. We associate the curve $\tilde{\gamma}:[0,l(\gamma)+2\alpha]\to R^2$ with respect to arc-length such that

$$l(\gamma)>0\ \begin{cases}\tilde{\gamma}=S(\gamma(0),\lambda) & [0,\alpha]\\ \tilde{\gamma}=\gamma & [\alpha,l(\gamma)+\alpha]\\ \tilde{\gamma}=S(\gamma(1),\lambda) & [l(\gamma)+\alpha,l(\gamma)+2\alpha]\end{cases}. \tag{6}$$

$$l(\gamma)=0\ \begin{cases}\tilde{\gamma}=S(\gamma(0),\lambda) & [0,\alpha]\\ \tilde{\gamma}=S(\gamma(1),\lambda) & [\alpha,2\alpha]\end{cases}. \tag{7}$$

Finally we define the energy of γ as

$$E_1(\gamma)=\int_0^{l(\gamma)+2\alpha}(1+|\tilde{\gamma}''(s)|)ds. \tag{8}$$

4 Image Interpolation

What we should do is to connect the level lines with the same value on $\partial\Omega$, and make the lines cross the occluded area. After connecting all the level lines, the image information will be restored according to the level lines. One simple way to connect level lines unintersectantly is to use beelines

or pine curves, but this way will destroy the image content and can't achieve the purpose of restoring image. Therefore we employ geodesic curves to connect the broken level lines [8]. If there is something unusual on the level lines of the boundary, such as the level lines are the vertical lines of the boundary. As shown in Fig. 2:

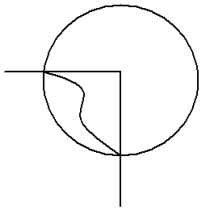


Fig. 2 unusual level lines

The circle area is the occluded area. The two beelines out of the circle are the level lines with the same value of the occluded image, and the curve in the circle is the geodesic curve connecting the level lines. Obviously, in this case, the level lines of the original image should be the two beelines in the circle. If we connect them with geodesic curve, there will be a big windage. According to this situation, we can use the meaningful method to detect the beeline information of the image. If there are no beelines, we can connect level lines with geodesic curve, otherwise we can extend the beelines. This method will have good effect when the occluded area includes sharp gray degrees.

From the above, we can get λ-level set of Ω, and denote it as χ_λ. Then we define the reconstructed function,

$$\forall x \in \Omega, u_d(x) := \sup\{\lambda : x \in \chi_\lambda\} \tag{9}$$

Finally we define the reconstructed function related to u as

$$u_r(x) = \begin{cases} u(x) & if \quad x \in R^2 \setminus \overline{\Omega} \\ u_d(x) & if \quad x \in \Omega \end{cases}. \tag{10}$$

We call p-disocclusion of u with respect to Ω any reconstructed function $u_d(x)$ obtained like above and such that the associated set of curves, denoted by D has finite total energy

$$E_p(D) = \int_{\Re} \sum_{x \in \partial\Omega} E_p(\gamma(x,\lambda)) d\lambda. \tag{11}$$

For simplicity we shall also call D a disocclusion. According to the knowledge of function, we can draw the conclusion as follows: Let $\Omega \subset R^2$ be an occlusion and $u \in BV(R^2 \setminus \overline{\Omega}), |u| < M$, then there exists a 1-disocclusion of u with minimal energy and the reconstructed $u_r \in BV(R^2)$, if $p > 1$, and there is a finite number of level lines on $\Gamma = \partial\Omega$, then there exists a p-disocclusion of u with minimal energy and the reconstructed function $u_r \in BV(R^2)$.

5 Experimental Results

(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 3 experiment results (a) original image (b) occluded image (c) use median filter (d) use our algorithm (e) original image (f) occluded image (g) use median filter (h) use our algorithm

Our approach is evaluated on two testing images, as shown in Fig. 3 (a) and Fig. 3 (e). Fig. 3 (b) and Fig. 3 (f) are the corresponding occluded images which have some occluded areas. The results of using median filter are shown in Fig. 3 (c) and Fig. 3 (g). Fig. 3 (d) and Fig. 3 (h) are the results of our algorithm. It can be observed that our algorithm has better effect.

6 Conclusion

This paper puts forward a method using level set to interpolate the occluded area of an image to restore the image. Together with the experiments with other occlusion and other images, the method mentioned in this paper proves to be of better effect, and is applicable in file management and photo preservation. Meanwhile, compared with the result using middle filter algorithm, the method in this paper has better restoring effects.

References

[1] Y.L. Zhou and J. Liu: Research on Image Preservation Program of National Key Files, Proceedings of International Symposium on Image Protection (2009).

[2] V. Caselles, J.M. Morel and C. Sbert: An axiomatic approach to image interpolation, IEEE Trans. on Image Processing, Vol. 7-3 (1998), p. 376-386.

[3] S. Osher and J.A. Sethian: Fronts propagating with curvature dependent speed: algorithms based on the Hamilton-Jacobi formulation, J. Computational Physics, Vol. 79 (1988), p. 12-49.

[4] R. M. Willett and R.D. Nowak: Minimax optimal level-set estimation, IEEE Trans. on Image Processing, Vol. 16-12 (2007), p. 2965-2979.

[5] A. Desolnerux, L. Moisan and J.M. Morel: From Gestalt theory to image analysis: a probabilistic approach, Springer-Verlag, New York (2008).

[6] V.E. Eric and H. Matthias: The geometric minimum action method for computing minimum energy paths, J. Chemical Physics, Vol. 128-6 (2008), p. 061103-4.

[7] M. Nitzberg, D. Mumford and T. Shiota: Filtering, segmentation and depth, Lecture Notes on Computer Science, Vol. 662, Springer-Verlag, Berlin Heidelberg (1993).

[8] D.M. Morera, P.C. Carvalho and L. Velho: Geodesic Bezier Curves: a Tool for Modeling on Triangulations, Proceedings of 20th Brazilian Symposium. on Computer Graphics and Image Processing (2007), p. 71-78.

©(2014) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.885

A Crime Decision-making Model Based on AHP

Feixue Yan[a], Jing Xia[b], Guanqun Shen[c] and Xusheng Kang[*,d]

School of Computer & Computing Science,Zhejiang University City College,Hangzhou, P.R.China

[a]shirelyyan@gmail.com, [b]xiajing616@gmail.com, [c]Prophet_shen@hotmail.com, [d]kangxs@zucc.edu.cn

Keywords: AHP; geographical profile; key feature; Location Prediction

Abstract. As time goes by, hazard rate of the society would increase if crime prediction was not implemented. Based on objective factors of offenders and victims characteristics, AHP method can be established to get a quantitative and qualitative analysis on crime prediction. Crime prediction is a strategic and tactical measure for crime prevention. According to AHP analysis, two prediction models of the optimal predictive crime locations are put forward. Standard Deviational Ellipses Model and Key Feature adjusted Spatial Choice Model were formulated to account for the anticipated position with various elements from AHP method. These models could be applied in a computer simulation of situation tests of the series murders. Besides, applying those models in certain real case demonstrates how the models work. Through models comparison, the results are summarized that Key Feature adjusted Spatial Choice Model is more conducive in confirming the guilty place. In conclusion, the suggested models, including detailed criminal map, are easy to implement.

Introduction

During the past few decades, interest in environmental criminology, spatial crime analysis and the investigation of offender movement patterns searching “geographical profile” has grown. The break can be defined as the use of the geographic imagination in concert with the sociological imagination to describe, understand and control criminal events. According to Brantingham, P.J., and Brantingham, P.L, this change is characterized by at least three shifts [1] in perspective.

First, a significant shift from the tendency of academics to keep their research contained within their own specific discipline, as environmental criminologists and environmental psychologists introduced techniques and knowledge from different disciplines like geography and geographical information science into their research. Second, a shift from the traditional search for causes of criminal motivation, in which it was simply assumed that people are criminally motivated. The focus is now on the criminal event, to find and explain patterns in where, when and how crimes occur. Third, a move is from the sociological imagination of crime to the geographical imagination. In summary, the field of environmental criminology includes studies of the spatial patterning of crime at different levels of aggregation, the processes by which potential offenders recognize potential crime sites and specific opportunities, and the creation and maintenance of areas of criminal residence [2].

In reality, searching for criminals often have to depend on "profiling", the psychological level can tell who is the offender and the descriptions of the geographical profile aim to delineate the geographic ,that is where is the place to the offender commit a crime next. Then search will have a hypothetical mathematical model and search the spatial distribution problem deeply.

Pre-modeling

The Data Processing

At first, it has to focus on methodological aspects and apply a trivariate Weibull survival model using competing risks concept to predict 3 types of crimes – gender, violence and others. Analytic Hierarchy Process is a decision-making method through combining qualitative and quantitative analysis. The offender and victim , separated as two properties and the features, are checked (refer with Fig.1).

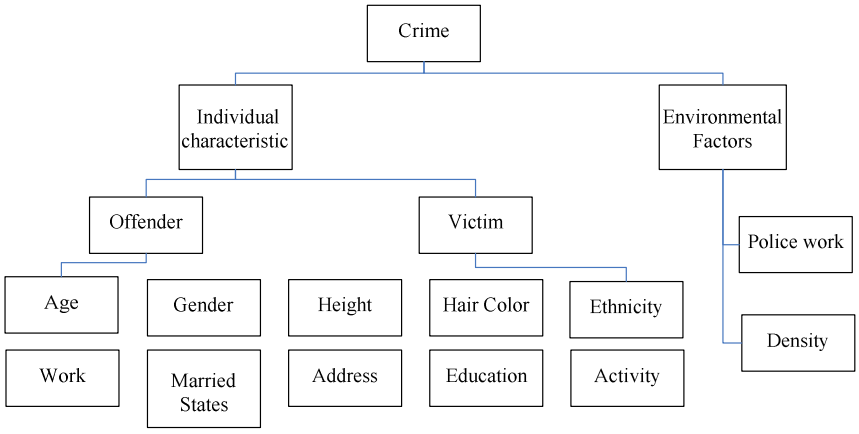


Figure 1. The Impact Of Model Parameters On Geographical Profile

Features Analysis

The entire target range of factors are determined, as well as within the scope of the relation various elements. From Fig.1, the most important observable features of offenders are gender, ethnicity, age, height, build, hair color, hair, length, and facial hair. Combinations of these features were used to construct offender profiles. Offense and victim features were compared with offender features. There was much significant regularity. Offense profiles based on location and site are compared with offender profiles based on address, age, gender, and ethnicity. Address–age–gender victim profiles were compared with address–age–gender–ethnicity offender profiles.

Try to construct and calculate a comparison matrix [3].Then obtain a relative weight value by the correlation between two factors.

The improved AHP decision-making methods are used for analyzing and assessing the place of crime prediction associated parameters. The next step is to make use of the result to answer this question and give a brief overview of the growing significance of location-based approaches in criminology according to the each parameters relative weight and sort.

The following model will demonstrate the performance improvement of adding dimensions of preferences discovered through feature space analysis[4],then future crime locations based on the features of past crime data about three cases of Peter Sutcliffe, ‘green river killer’ Gary Ridgway and ‘Co-ed Butcher’ Edmund Emil Kemper respectively is predicted.

The Location Prediction Model

The Hypothesis and Development of Model

Newton’s circle-based geoforensic analysis technique [5] is a practical method for measuring the range of the crime scenes. The dim scope of the crime scenes was confirmed by applying circular model [5]. It could be found that major crime scenes were contained by a circle.

However, circle-based geoforensic analysis isn’t suitable for a non-uniform arrangement of a series of criminal activities, because the shape of those crimes’ geographical profile is irregular. Considering this factor, the conformation of the original model is needed to change. As a result, an ellipses model [6] is put forward.

This model calculates both the location and dispersion of linked crime scenes committed by a localized serial offender. The approach was developed on the premise of that an area encircled the geographic centre of a crime distribution. To provide a search parameter for this effect, the algorithm generates a circular search area with a radius defined by the extent of the two furthest crime scenes in the sequence. As each new crime scene is added, the size of the search area is reduced.

Develop the standard deviational ellipses model:

Confirm the origin of the ellipse:

Along the x-axis and y-axis, the ellipse’s major and minor axes are determined by measuring the distribution’s data so that axises are orthogonal to each other: $x' = (x - \bar{x})\ and\ y' = (y - \bar{y})$ (1)

where the mean, $\bar{x}$ and $\bar{y}$, are subtracted from each of the original x and y coordinates to give the transposed coordinates denoted by x' and y'.

Determine the angle of rotation along the x-axis and y-axis so that the sum of squared differences is minimized:

$$\theta = ARCTAN\left\{\frac{(\sum x'^2 - \sum y'^2) + \sqrt{(\sum x'^2 - \sum y'^2)^2 + 4(\sum x'y')^2}}{2(\sum x'y')}\right\} \tag{2}$$

where θ is the angle of rotation observed in the distribution.

Calculate the data along the transposed axes

$$s_x = \sqrt{2\frac{\cos^2\theta\sum x'^2 - 2(\sin\theta\cos\theta\sum x'y') + \sin^2\theta\sum y'^2}{n-2}} \quad s_y = \sqrt{2\frac{\sin^2\theta\sum x'^2 - 2(\sin\theta\cos\theta\sum x'y') + \cos^2\theta\sum y'^2}{n-2}} \tag{3}$$

where s_x and s_y are the standard deviations parallel to the x- and y-axis of the ellipse, and n is the number of points in the given sequence.

Finally, the elliptical search area is reduced in accordance with Newton's hypothesis:

$$s'_x = \frac{s_x}{n-1} \text{ and } s'_y = \frac{s_y}{n-1} \tag{4}$$

where s'_x and s'_y represent the final standard deviations on the major and minor axes, respectively, and where n is the number of incidents in the sequence.

Key Feature Adjusted Spatial Choice Model

Spatial choice is a procedure that individuals use to decide a specific site in space to meet their objectives. Criminals choose potential sites in space to commit crimes. Their choices show certain patterns over a geographic region. The geographical sites form a universal spatial alternative set A. Individuals will make a choice from the optional position. The spatial alternatives in choice set A have some specific characteristics, which make the spatial choice process unlike from other optional processes. First, the spatial choice process is a discrete procedure. The number of alternatives, N, will be restricted. Second, the spatial alternatives have relatively stable positions during the choice process.

Spatial choice problems can be stated as a sub-optimal or locally optimal challenge. Concerning individuals' hierarchical information processing, individuals make spatial choices from alternatives with features that they have evaluated $A_d \subseteq A$. The choices individual d making will most possibly have the highest utility among all alternatives in choice set A_d. Some methods are proposed to identify or estimate the probability that an alternative a_i is considered by individual d, $P(a_i \mid M, d)$.

During the identification of an individual's choice set, two factors are considered in his spatial choice procedure: 1) the utility of option to individual d, and 2) the possibility that an option is accessible or considered by individual d. Since the number of alternatives in a spatial choice problem is large, it is possible that some choices can have higher utility values but are never considered. If the two factors are equally significant for the individuals' selection, then the combination of $P(a_i \in A_d)$ and the utility of alternative to individual d, U_{id}, can give a better estimate of the likelihood of alternatives. The probability that individual chooses alternative from choice set can be stated as $P(U_{id} > U_{jd} + \ln P(a_j \in A_d), all\ a_j \in A)P(a_j \in A_d)$. Under the extreme distribution assumption, the spatial site selection model, which is based on the factors weight from AHP Method using same derivation as McFadden [7] and Fortheringham [8], are got.

$$P(a_i\ A_d, d) = \frac{\exp(V(d, s_i)) \cdot P(a_i \in A_d)}{(\sum_{a_j \in A} \exp(V(d, s_j) \cdot P(a_j \in A_d)))} \tag{5}$$

This model is a multinomial logit model where each alternative's observable utility is weighted by the probability that the alternative is evaluated.

The Solution and Checking of Model

1) *Method 1: The standard deviational ellipses model*

The data used in this analysis were selected from a database of Peter Sutcliffe which consists 13 solved serial crimes that occurred in Baltimore County, MD, between 1994 and 1997.

Calculate the relative outcomes from the database:

The position of elliptical search area: $s_x'' = 0.0126839$ *and* $s_y'' = 0.0352301$ (6)

The probability P of elliptical model is the proportional of scene spots in the elliptical to the total scene spots: P=0.72 (7)

It could be found that the probability shows the accuracy of this model.

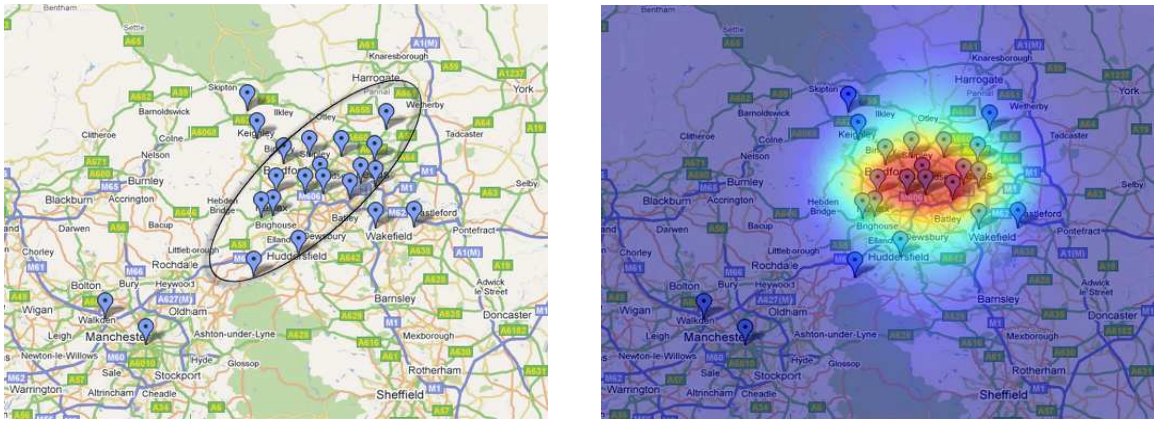
2) *Method 2: Key Feature Adjusted Spatial Choice Model*

In order to classify the criminal details in gender, age, preference, income, marriage and religion, Analytic Hierarchy Process Method put the above parameters into Key Feature Adjusted Model. Based on all above parameters, and the location data set of "Co-ed Butcher Edmund Emil Kemper" was used, a prediction Map is generated in MATLAB.

According to the figure, the probability of the 29th geographic profile was obviously obtained, where the locations in relative high possibility are Bethlehem, PA, United States and Pluckemin, NJ, United States. Police searching for the case should focus on the above two regions, in those places the criminal is most likely to be active.

3) *Model Comparisons of Two Methods*

Using the testing data set of the case of Peter Sutcliffe between September, 1969 and November 17, 1980, the estimation of the three models based on the factor weight from AHP Method is obtained. From the estimation results, by comparing the prediction data with practical data, predictions of future crime locations are made. Another testing data set of the case of "the green river killer Gary Ridgway" contains incidents between August 15, 1982 and June 1, 1986. The testing incidents and data validation for Standard Deviational Ellipses Mode, Spatial Choice Models and Key Feature Adjusted Spatial Choice Model are displayed in a and b in Fig.2.



a b

Figure 2. Verification of Standard Deviational Ellipses Model and Key Feature Adjusted Model with crime incidents from Sept, 1969 to Nov 17, 1980

To compare diverse models, the predictions of the three models (Standard Deviational Ellipses Model, Spatial Choice Models, Key Feature Adjusted Model) are standardized. Under the hypothesis of all future crime incidents, the proposed model, Key Feature Adjusted Model, will statistically outperform the comparison model.

Model Comparison Results:

Table 1. Testing Data Set (The case of Peter Sutcliffe 7 / 5 / 1975 – 11 / 17 / 1980)

	Mean	*Std. Dev*	*t-Statistic*	*p-Value*
Spatial Choice Models vs. Standard Deviational Ellipses Model	0.295455	0.406828	-2.408667	0.010238
Key Feature Adjusted Model vs. Standard Deviational Ellipses Model	0.136364	0.375494	-1.20446	0.117578
Key Feature Adjusted Model vs. Spatial Choice Models	0.159091	0.314185	1.679409	0.050249

The comparison results above indicate that the two spatial choice models dramatically outperform the Standard Deviational Ellipses Model. Furthermore, the Key Feature Adjusted Model improves the prediction results over spatial choice model significantly. On the ground of previous studies [4], the key feature space density estimation gives a better prediction than space density estimation does. In accordance with the specification of priors in feature space over all alternatives in choice set, a more efficient and accurate method for predictions of criminal's future spatial choices are provided. This result has clear implications.

Conclusion

Police detect series crimes usually begin with collecting the proof from the crime scenes and then statistics offenders and victims' characteristics and environmental factors. The features mainly include gender, ethnicity, age, height, build, hair color, and facial hair, work, martial statue.

Police should proceed with the factors. Within the scope of the relation, various elements regarding as parameters are important to our two models for "geographical profile". Then a computerized offender profiling system can be recommended, based on criminological theories and empirical data about statistical regularities linking the characteristics of offenders, offenses and victims.

The simulated result by Standard Deviational Ellipses Model and Key Feature Adjusted Model shows that the latter one is a more conducive method to predict geography. Owing to the crime analysts use real data with methodologies like aggregate crime rate analysis, space point process modeling to analyze and predict the spatial patterns of crimes. It can narrow the scope of crime location. Predicting future crime locations based on the features of past crime data about three cases of Peter Sutcliffe, 'green river killer' Gary Ridgway and 'Co-ed Butcher' Edmund Emil Kemper, what is verified the correctness of the model closely.

Determining the probability of the offender residing in various areas and displaying those results through the use of choropleth or isopleth maps can assist police efforts to apprehend criminals. This union information allows police departments to focus their investigative activities to reduce the crime occurrence .According to the similarity of crime features, police should give a brief overview of the growing significance of location-based approaches in criminology, and then take actions to warn and protect the interrelated victims from attacking.

References

[1] Brantingham, P.J. and Brantingham, P.L: *Environmental Criminology*. Waveland Press, Illinois, (1981).

[2] Johan G.J. van Schaaik and Jasper J. van der Kemp, in: Real Crimes on Virtual Maps: The Application of Geography and GIS in Criminology, pp. 217-237. (2009)

[3] Yan Lei, and Hu Ping, in: Stratification analysis based on the safety assessment of parameters of model", Computer Engineering and Applications, pp. 126-129. (2006)

[4] D. Brown, H. Liu, and Y. Xue: *Mining preferences from spatial-temporaldata*, in Proc. SIAM Conf. (2001)

[5] Newton, M. B: Geographical discovery of the residence of an unknown dispersing localized serial murder. Baton Rouge, LA: Louisiana State University (1988)

[6] Levine, N, CrimeStat: A spatial statistics program for the analysis of crime incident locations, vol 3.1, Houston, TX: Ned Levine & Associates; and Washington, DC: The National Institute of Justice (2007)

[7] Brent E. Turvey Criminal Profiling: An Introduction to Behavioral Evidence Analysis, No.3, (April 2008)

[8] Bernasco W: Them Again? Same-offender Involvement in Repeat and Near Repeat Burglaries, European Journal of Criminology, 5.4, forthcoming (2008)

*Corresponding author.

©(2014) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.890

Simulation & Prediction of Pacific Plastic Pollution

Cao Li [a], Yu Xu [b], YiShi Xu [c], XuSheng Kang [*,d]

School of Computer & Computing Science, Zhejiang University City College, Hangzhou, P.R. China
[a]110342219@qq.com, [b]jiaowo19871230@sina.com, [c]zhejiangxuyishi@hotmail.com,
[d]kangxs@zucc.edu.cn

Keywords: Plastic, Great Pacific Garbage Patch, Marine Pollution，differential equation

Abstract. The Great Pacific Garbage Patch which characterized by exceptionally high concentrations of suspended plastic, chemical sludge, and other debris, has a variety of severe influence on the marine life and human beings. Therefore, cleaning up the patch immediately is a matter of great urgency. To solve the problem, three models which are based on different differential equations to predict the amount of waste plastic poured into the ocean are established. According to the basic model, in about next 20 years, the Pacific Ocean will be unable to hold more garbage. Based on the fact that salvaging the waste plastic in the ocean contributes to controlling marine pollution, in accordance with the better model, the Pacific Ocean still has the ability to accommodate garbage before 2060. By improving the better model, a series of results are obtained. Around 2060, the area of the Pacific Ocean Garbage Patch will reach relatively steady state; meanwhile, the annual emissions of the plastic will reach 32,397,273.47 tons. To keep a relatively stable state the area which should be cleaned up will reach about 3,601,672 square kilometers per year. To compare the effect of government increasing taxation and plastic salvaging, a conclusion is reached that the huge economic losses caused by taxes are far more than the amount of money spent in salvaging the plastic garbage.

Introduction

The Great Pacific Garbage Patch is a gyre of marine litter in the central North Pacific Ocean located roughly between 135° to 155°W and 35° to 42°N and estimated to be twice the size of Texas [1].

The patch is characterized by exceptionally high concentrations of suspended plastic, chemical sludge, and other debris that have been trapped by the currents of the North Pacific Gyre [2]. Despite its size and density, the patch is not visible from satellite photography because it consists of very small pieces, almost invisible to the naked eye [3] and most of its contents are suspended beneath the surface of the ocean [4].

The source of Pacific garbage patch

Every piece of plastic is likely to flow into the ocean. The plastic litter in the sea is mainly from three sources: first, the buried plastic garbage on land washed into the sea by storm; second, people who lack of environmental awareness will pour the plastic garbage into the sea; the third is plastic products inadvertently left by the containers on the deck of a cargo ships due to various marine accidents [5].

Some Indicators of plastic and Plastic Pollution

The production of plastic. The paper uses the sum of production which comes from the world top five plastics productive countries. According to the capacity of plastic-producing in top five countries in year 2001 from Internet and some records of the plastic production in China from a statistical website, the total production of the five countries is estimated, since 2001 to 2008 (see Table 1).

TABLE 1. THE ESTIMATED PRODUCTION OF PLASTIC PRODUCT IN THE TOP 5 COUNTRIES

The estimated production of plastic product in the Top 5 countries	
	units 10,000 tons
time	accumulative total
Dec-2001	9506.4268
Dec-2002	11218.0851
Dec-2003	13205.6505
Dec-2004	14754.4139
Dec-2005	17544.3492
Dec-2006	22330.9039
Dec-2007	26286.6264
Dec-2008	29524.869

TABLE 2. THE ESTIMATED NUMBER OF WASTE PLASTIC IN THE TOP 5 COUNTRIES

The estimated number of waste plastic in the Top 5 countries	
	units 10,000 tons
time	accumulative total
Dec-2001	4477.527023
Dec-2002	5283.718082
Dec-2003	6219.861386
Dec-2004	6949.328947
Dec-2005	8263.388473
Dec-2006	10517.85574
Dec-2007	12381.00103
Dec-2008	13906.2133

TABLE 3. THE ESTIMATED NUMBER OF WASTE PLASTIC POUR INTO THE OCEAN

The estimated number of waste plastic pour into the ocean	
	units 10,000 tons
time	accumulative total
Dec-2001	447.7527023
Dec-2002	528.3718082
Dec-2003	621.9861386
Dec-2004	694.9328947
Dec-2005	826.3388473
Dec-2006	1051.785574
Dec-2007	1238.100103
Dec-2008	1390.62133

The waste and the recycling number of plastic. Waste plastic is an important factor leading to the pollution. An article [6] involves the production and the amount of waste plastic of the United States in the same year. According to this, the rate is estimated and the annual amount of waste plastic in top5 countries (see Table 2) is calculated. The range of recycling rate is about 20% to 45% according to the Internet.

The amount of plastic waste poured into the ocean and the floating plastics. Studies have shown that 80% of marine debris is plastics. It is estimated that over 13000 pieces of plastic litter are floating on every sq km of ocean, and that 10% of the world's plastic waste, perhaps as much as 100 million tons, finds its way to oceans [7]. So, the annual amount of waste plastics poured into the sea can be calculated. (see Table 3).

According to the incomplete statistics, the garbage poured into the ocean is 640 million tons each year. About 70% of them are thrown into the sea floor, and about 15% of them floating at sea for long. Also, there are about 15% of them washed back to the beach. Sinking garbage is usually broken fishing nets, fishing gear, cans, glass bottles and so on. Floating objects at sea are mostly fast-food containers and some plastic materials.[8]

The basic model for plastic pollution

Symbols Agreement.

Table 4. Variables used in the basic model

Variables used in the basic model	
C_I	The yearly amount of the discarded plastic which is poured into the ocean
I	The pacific-rim nations' Gross Domestic Product (GDP)
R	The amount of plastic which can be recycled in C_I every year
P_I	The political influence on plastic recycling

The Development of Models and solution of the model. In order to predict C_I, assume that there is a linear relationship between the change rate of C_I and the change rate of the region's GDP. Use the top five countries' GDP instead of the Pacific Rim countries' GDP. So define an equation as follows:

$$\frac{dC_I}{dt} \Big/ \frac{dI}{dt} = k \cdot \frac{C_I}{I}. \tag{1}$$

k is a coefficient which represents the political influence. Referring to a document [9], let

$$k = (\frac{R}{C_I} - x)^{P_I}. \quad (2)$$

In equation (2), $\frac{R}{C_I}$ is a coefficient which represents the recycling rate of the waste plastic, x is the monitoring standard .Through(1)(2), acquire:

$$\frac{dC_I}{dt} = \left(\frac{R}{C_I} - x\right)^{P_I} \cdot \frac{C_I}{I} \cdot \frac{dI}{dt}. \quad (3)$$

Then, make conclusions as follows:

Assume that in the next 50 years, GDP will continue to grow. So $\frac{dI}{dt} > 0$ and obviously $\frac{C_I}{I} > 0$. When $\frac{R}{C_I} < x$, government should inhibit the consumption of plastics. How to simulate the case of government's inhibiting consumption of plastics? Use P_I to represent the political influence. Through adjusting P_I, $\frac{dC_I}{dt} < 0$ can be acquired. In other words, the consumption of plastics will decrease. When $\frac{R}{C_I} > x$, the consumption of plastics increases and to some extent, the effect of government is limited.

Working from GDP data, fit a linear regression of GDP for The Pacific Rim countries to time:

$$I = 13.880t + 159.202 \quad (R^2 = 0.994). \quad (4)$$

$$\frac{dI}{dt} = 13.880. \quad (5)$$

By estimation, let: $\frac{R}{C_I} = 0.3$; Obviously, C_I has been increasing year by year, so $\frac{dC_I}{dt} > 0$. Because x is the monitoring standard, an appropriate value to express the government's policy can be set. In order to match the fact, let $x = 0.25$. So $(\frac{R}{C_I} - x) > 0$ can be guaranteed.

For the purpose of seeking an appropriate value for P_I, fit an equation based on data from Table 3:

$$C_I = 374.78e^{0.166t} \quad (R^2 = 0.944). \quad (6)$$

$$\frac{dC_I}{dt} = 62.2135e^{0.166t}. \quad (7)$$

According to data from Table 3, let: $C_I = 1390.62133, I = 272.209, t = 8$.

Acquire $P_I = -0.3996$ and compared to the real data, a conclusion that it is reasonable to do some simulation of our model can be drawn.

The better model

Symbols Agreement.

TABLE 5. THE NEW VARIABLES USED IN THE BETTER MODEL

The new variables used in the better model	
Q	The area of the Pacific Ocean Garbage Patch
A	The area of the Pacific Ocean
P_C	The political influence about salvaging plastic

The Development of Models and solution of the model. In order to improve the basic model, consider the government may hire a lot of people to salvage plastics. So establish a more complete model:

$$\frac{dC_I}{dt} = (\frac{R}{C_I} - x)^{P_I} \cdot \frac{C_I}{I} \cdot \frac{dI}{dt} - (\frac{Q}{A} - y)^{P_C} \cdot \frac{dQ}{dt} \cdot \frac{C_I}{Q}. \quad (8)$$

Where y is the monitoring standard.

According to the reference data [10], acquire:

$$Q = 154.275e^{0.1t}. \quad (9)$$

As to the fitting process, acquire:

$$\frac{dQ}{dt} = 15.4275e^{0.1t}. \tag{10}$$

By searching the Internet [11], the area of the "Great Pacific Ocean Garbage Patch.", is about $3,430,000km^2$.So $\frac{Q}{A} = 0.0189$ can be acquired, and let $y = 0.01$. Then, acquire:

$$\frac{dC_I}{dt} = (0.3 - x)^{P_I} \cdot \frac{13.880C_I}{13.880t + 159.202} - (\frac{154.275e^{0.1t}}{18134.4} - y)^{P_C} \cdot \frac{C_I}{10}. \tag{11}$$

To solve the differential equation, the result about ocean plastic emission in the future can be acquired.

Further Research

The deficiency of the better model is that it can not be used to predict in the long term. The reason is that the model ignores the effect of plastic salvaging. However, ignoring the effect of plastic salvaging will lead to the area of pacific garbage patch decreased break the law of exponential growth. On the other hand, default $\frac{dQ}{dt} > 0$ in the better model.

To predict the plastic emission in the long term, the decrease of $\frac{dQ}{dt}$ should be taken into account. Under certain conditions, Q changed according to the exponential function, so if the area of pacific garbage patch is decreased by salvaging, the growth rate will slow down. Here, use a coefficient of b to control the situation. In conclusion, $\frac{dQ}{dt} = abe^{bt}$ can be acquired, where $a = 154.225, b = 0.1$,

Define:

$$\frac{dQ}{dt} = abe^{(b - \frac{Q}{A} \cdot \frac{t}{C}) \cdot t}. \tag{12}$$

C in the model is a buffer, that is to say, after C year(s); the Q's growth rate will slow down. Of course, for this purpose government needs to pay a higher cost.

Combine (8) and (12) together, acquire:

$$\frac{dC_I}{dt} = (\frac{R}{C_I} - x)^{P_I} \cdot \frac{C_I}{I} \cdot \frac{dI}{dt} - abe^{(b - \frac{Q}{A} \frac{t}{C})t} \cdot (\frac{Q}{A} - y)^{P_C} \cdot \frac{C_I}{Q}. \tag{13}$$

(13) has not been solved, but another access to analyze the problem has been found and the result can be obtained by it.

Results

Simulation & Prediction. According to the information from the Internet [11], the area of the "Great Pacific Ocean Garbage Patch." is about 3,430,000km^2. So, estimate that there will be 6,182,600,000 tons [12] of plastics in the Pacific Ocean .Define $\mu = 6{,}182{,}600{,}000$ tons, if there is μ plastics in the Pacific Ocean, the Pacific Ocean will be severely polluted. Assume that the maximum capacity of litter in the Pacific Ocean is π.

If $\pi = 0.3\mu$, a conclusion that after 24 years, there will be 2,033,500,000 tons of plastics in the Pacific Ocean can be drawn

If $\pi = 0.2\mu$, the conclusion is that after 20 years the amount of plastics in the Pacific Ocean will be over π.

For the better model, use the ODE45 numerical integrator in MATLAB on (11), and find the results as the following Fig. 1 demonstrates:

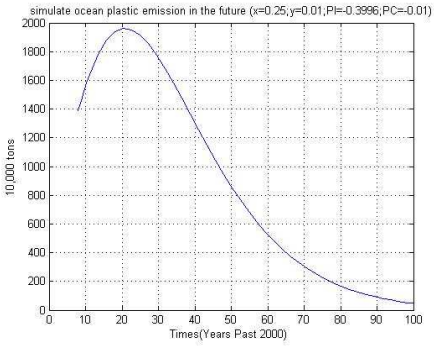


Fig. 1. Simulation of ocean plastic emission in the future

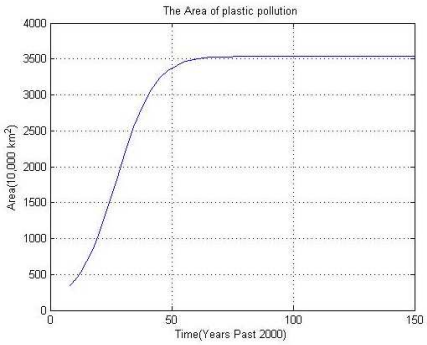


Fig. 2. The Area of plastic pollution

After calculation, the following conclusions can be drawn: if let $\pi = 0.3\mu$, in the next 50 years, the total amount of the plastics is far less π.

To a large extent, salvaging the plastics will delay the marine plastic pollution.

For the further research, change the value of C, simulate for (12). Finally let C=100, and a result as Fig. 2 as shows acquire:

According to the figures presented above, after 50 years, Q will be stable at $3535.1\times10^4 km^2$. Then, after 31 years, the waste plastic area in the Pacific Ocean will come to Q. So, that can be calculated from the year 2031 to 2032, the area of Pacific Garbage Patch will increase360.1672×10^4 km^2.This shows that under a steady state, effective measures should be taken to make the Pacific Ocean clear and the area that should be cleared is $360.1672\times10^4 km^2$ annually at least. Whereas, the move will take a lot of manpower and material resources, so suggest that reducing the emissions of plastics be used to achieve this goal instead of clearing the ocean. Under such circumstances, human beings can produce waste plastics 68784.02×10^4 tons, 32397.27347×10^4 tons will be discarded, 9719.182×10^4 tons will be recycled, 3239.727347×10^4 tons will be poured into the ocean, and 485.9591×10^4tons will be likely to float on the sea.

The Economic Costs and Benefits about the Management of Marine Litter. According to these models , estimate that the amount of plastics poured into oceans will come up to 15,400,000 tons in 2011.The amount of the plastics which float on the sea will come up to 2,300,000 tons .Assume that with the government's control, the emission of the plastics reduces by 1% every year. Then, the plastics floating on the sea will reduce 23,000 tons. Also, estimate that in order to recycle 10,000 tons of plastics in the ocean, 13,000,000 dollars will be spent. So under the government's control, 29,900,000 dollars per year can be saving. But the benefit from clearing the ocean can not be calculated with figures.

Take the U.S. for example, in order to control the plastic pollution in the ocean, government increase the tax. Then the unemployment rate will increase .Use the following diagrams to explain the case:

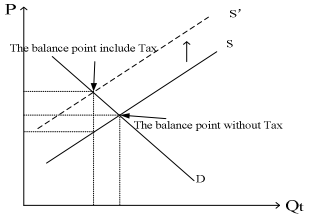


Fig. 3. Tax to merchant

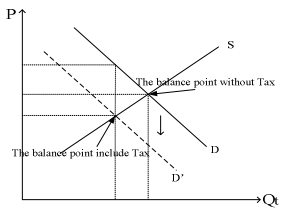


Fig. 4. Tax to consumers

Where: P is the mean price of the plastic products, Q_t is the output of the plastic products, D is the demand curve, S is the supply curve.

According to Fig. 3, when government raises tax, the enterprises' income will decrease, so they will reduce the production of plastics, so that the supply curve will move up. Call the new supply curve S'. Because of the reduction of the production, many employees will loss their job, so the unemployment rate will go up. The same result in Fig. 4 also can be drawn. Finally it will be found that the output of plastic product will go down. So, the unemployment rate will go up.

It can be estimated that there will be one hundredth workers who are employed by the plastic manufacturers losing their job. Take the U.S. for example, according to the data from Bureau of Labor Statistics [13], it can be estimated that there will be 10,000 workers' losing their job (see from the per capita GDP [13]). In 2008, the GDP of U.S. per capita is 43,250 dollars. Thus it can be got that if 10,000 workers lose their job, the economic loss will amount to 432,500,000 dollars.

All in all, it can be drawn that a conclusion that in order to clear the ocean billons of dollars will be cost. But, the benefit from clearing the water can not be calculated. For example, people can save a lot of fish and make the environment more beautiful.

References

[1] Bradshaw and Kate: *The Great Garbage Swirl*, Maui Time Weekly (Maui: Linear Publishing), January 29, 2009

[2] Information on http:// en.wikipedia.org/wiki/Great_Pacific_Garbage_Patch

[3] Cecil Adams: *Why don't we ever see pictures of the floating island of garbage?,* Straight Dope / Chicago Reader, September 8, 2009

[4] Steve Gorman: *Scientists study huge plastic patch in Pacific*, Reuters, August 4, 2009

[5] Information on http://baike.baidu.com/view/ 2525639.htm?fr=ala0_1_1

[6] Information on http://www.zz91.com/cn/trade485.html

[7] P. Agamuthu: *Ocean (Land) fill for plastic wastes?*, Waste Management & Research, 2009:27:551–552

[8] Bentai Wan: *Strengthen the Prevention of marine litter's pollution, further advance the protection work for marine environment*, Environmental Protection, 2008:19:59-61

[9] Steven Krumholz, Frances Haugen and Daniel Lindquist: *Preventing the Hydrocalypse: A Model for Predicting and Managing Worldwide Water Resources*, UMAP J 26 no2 Summ 2005:145-162

[10] Information on http://baike.baidu. com/view/1635061.html

[11] Information on http://news.iqilu.com/keji/20080630/31303.html

[12] C. J. Moore, S. L.Moore, M. K. Leecaster and S. B. Weisberg, A Comparison of Plastic and Plankton in the North Pacific Central Gyre, Marine Pollution Bulletin, Vol.42, No.12, pp.1297-1300, 2001

[13] Information on http://www. bls.gov/oes/oes_dl.htm

*Corresponding author.

Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.896

Numerical Simulation of Small Cold Storage

Xiaohui Zhong[1,a], Yuling Zhai[1,b] and Yujun Gou[2,c]

[1]College of Metallurgy and Energy, Hebei Polytechnic University, Tang Shan City, China

[2]Department of Civil Engineering, Tangshan University, Tang Shan City, China

[a]zhong_xiaohui@163.com, [b]zhaiyuling00@126.com, [c]hbtsgyj2008@163.com

Keywords: cold storage; pre-cooling; defrosting; switch door; numerical simulation.

Abstract. The airflow change in cold storage were simulated by FLUENT under following three conditions, pre-cooling, stop to defrost and switch cold storage door. The result shows that the pre-cooling wind speed should be chose reasonably while guarantees the pre-cooling effort, hot air defrost had a favorable defrosting performance, and it is essential to seek optimal defrosting time according to the frost thickness and energy consume. The temperature field near the switch door was analyzed, and the temperature gradient near the switch door is considerable, but the outdoor air can be obstructed by increase air curtain.

0 Introduction

With the improvement of people's living standard and higher and higher requirement of food's freshness quality, the cold storage industry developed rapidly. But there are still many problems in cold storage, such as unreasonable distribution in storehouse, hot and cold gas permeability in cold storage gate, high energy consumption for defrosting and so on, so the reasonable air current design in cold storage is a difficult problem. Although the traditional experimental method can obtain quite detailed flow field information, it also has the very big limitation in the actual utilization, which cost time and efforts and is high cost.

In recent years, numerical simulation technology has been widely used in all areas' actual project, and achieved quite satisfactory effort, however, due to airflow and heat and mass transfer in cold storage are complicated and diversified, the application of numerical calculation in cold storage is basically staying in design and verification stage of a particular model[1-7], and is not comprehensive enough. In this paper, based on the former researches the airflow change in cold storage were simulated by FLUENT under following three conditions, pre-cooling, stop to defrost and switch cold storage door, and the research can provide theory basis for optimize the design of cold storage.

1 Model establishment

1.1 Physical model

Select a small cold storage for research object as shown in Figure1, the internal size of length × width × height is 4.5m × 3.3m × 3.0m, and single finned tube evaporator cooling is in the ceiling, because the shape of the evaporator has little influence on velocity and temperature field, the evaporator is simplified to be cuboid, the length × width × height of evaporator is 0.7m × 0.3m × 0.5m, the length × width × height of air outlet is 0.5m × 0.3m × 0.5m, the length × height of door is 1.2m×2.0m.

1.2 Mathematical model

The airflow in the cold storage under three kinds of conditions are chaotic, belongs to large space forced convection heat transfer problem. It can be solved by k-ε model, which including the continuity equation, momentum equation, energy equation, k and ε equations. Wall-function method is employed to simulate near-wall. To simplify the model, the goods were ignored and the cold storage is considered to be empty.

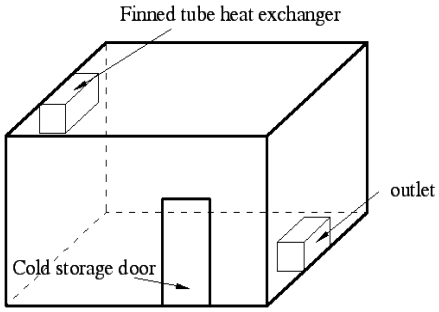


Fig.1 Schematic diagram of cold storage

1.3 The boundary and initial conditions

1.3.1 Start pre-cooling condition

The most important factor of cooling rate is wind speed. Therefore, different inlet velocity was used to simulate the flow field and temperature field. The condition was set as follows:

(1) Inlet boundary: use speed boundary with different wind speed separately: V_1=1.5m/s，t=-15℃；V_2=2.5m/s，t=-15℃；V_3=3.5m/s， t=-15℃.

(2) Outlet boundary: use the outflow boundary, the wind speed is calculated voluntarily by the software obtains.

(3) Turbulence intensity was set as 5% [6], and the hydraulic diameter of the inlet and outlet are calculated to be 0.42m and 0.5m.

(4) Wall: use the first boundary condition, and take the initial temperature as wall surfaces temperature whose value is 30 ℃.

1.3.2 Stop defrosting conditions

With the running of refrigerant unit, the frost will grow on evaporator surface and get more and more thick, this cause the refrigeration efficiency reduce. At this point stop the refrigeration unit, exert a constant wind to the evaporator surface to defrost, the wind speed is 0.5 m/s and the temperature is 50 ℃. In this paper, when evaporator frost thickness reaches 5mm, the surface of evaporator star to defrost, defrosting time is about 10 minutes by calculating. Set the monitor window in the fluent software, which monitoring the temperature and flow field changes. The condition is set as follows:

(1) Inlet boundary: use the speed boundary: V=0.5m/s, t=50℃.

(2) Outlet boundary: use the outflow boundary, the wind speed is calculated voluntarily by the software obtains.

(3) Wall: the first boundary condition, its value is the temperature when achieve steady state, the value is t =- 10 ℃.

1.3.3 Switch door condition

As the wind pressure and the barometric pressure combined action, the fluid nearby cold storage door will have a strong heat and mass transfer, especially when there are goods in and out of cold storage, the cold storage door takes a long time to open to undermine the internal cold air flow field.

The flow fluid nearby cold storage door were simulated when time is 10s, 20s, 30s. The condition is set as follows:

(1) Inlet boundary: use evaporator inlet velocity boundary, V_2=2.5m/s, t=-15℃. When open the door, use inlet velocity, V2=0.2m/s, t=30℃. When close the door, use wall conditions.

(2) Evaporator outlet boundary: the wind speed is calculated voluntarily by the software.

(3) Wall: use the first boundary condition, and the temperature is 30 ℃.

1.4 Mesh

Use gambit to pre-process the software modeling and establishes the three-dimensional cold storage space, the evaporator and the air outlet in gambit2.2.30, quad meshes with the structured grid and to local region, such as air inlet, outlet and front door place encryption.

2 Simulation and Analysis

In this paper, a few representative surfaces were selected to analysis the three dimensional computed result

2.1 Start pre-cooling condition

Three pre-cooling wind speed were set and the representative section y = 1.6m (including inlet and outlet section) were selected. Figure 2 shows the flow and temperature fields that achieved steady state conditions when wind speed is1.5m/s, 2.5m/s and 3.5m/s.

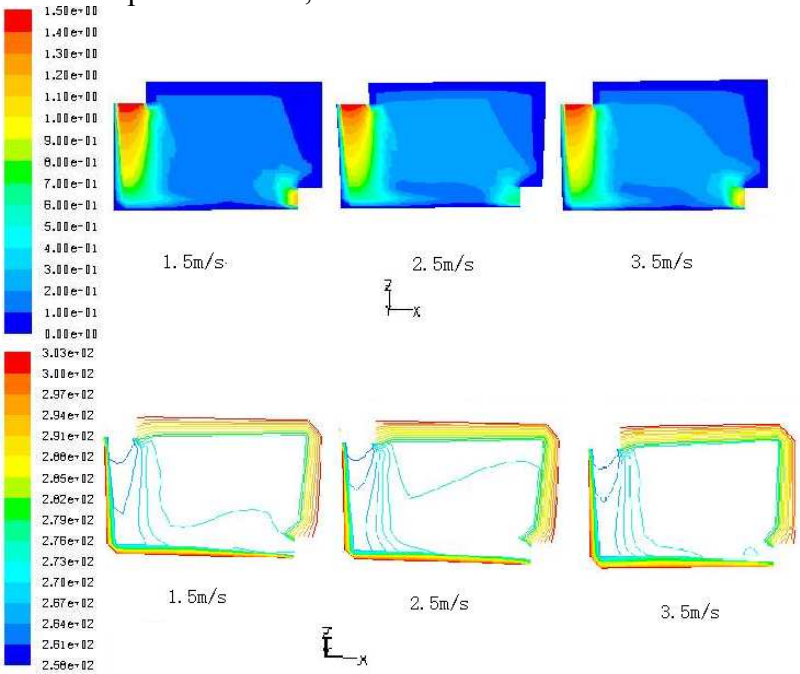


Fig.2 The flow and temperature fields with different wind speed

Figure2 shows, with the increase of evaporator blowing speed, it can obviously improve the wall pressing weir flow phenomena. The low temperature region is biggest when speed rate at 3.5m/s. Contrast 2.5m/s condition and 3.5m/s condition, there is the little difference between the two low-temperature regions. It shows that we should not infinitely increase the blowing speed while the pre-cooling effect is guaranteed. Therefore, it should be comprehensively consider all factors to find the best blowing pre-cooling speed.

2.2 Stop defrosting condition

A representative section y =1.6m (including inlet and outlet section) was selected too. From Figure 3 we can see that the hot air between inlet and outlet form a rectangular wind zone, the jet flow wind is the highest speed and temperature of the all region. The gradient of velocity and temperature is large at the beginning in cold storage, which shows that the heat transfer of entire flow field is fierce.

The curve of monitor window shows, the curvature of average cure is large at the beginning then become flat gradually. It can also be reflected in the temperature field, the time is longer the temperature gradient is flatter. It indicated that the temperature in the cold storage can reach hot air temperature in short time as long as the refrigeration unit stop and defrost begin. The time of defrosting is longer, the quality of defrosting is better, but the cold storage's temperature is higher, so the time of recovery refrigeration operation condition is longer. Therefore, it must measure the relationships between the quality of defrosting and the energy consumption of refrigeration, as a result to find an optimum defrosting time.

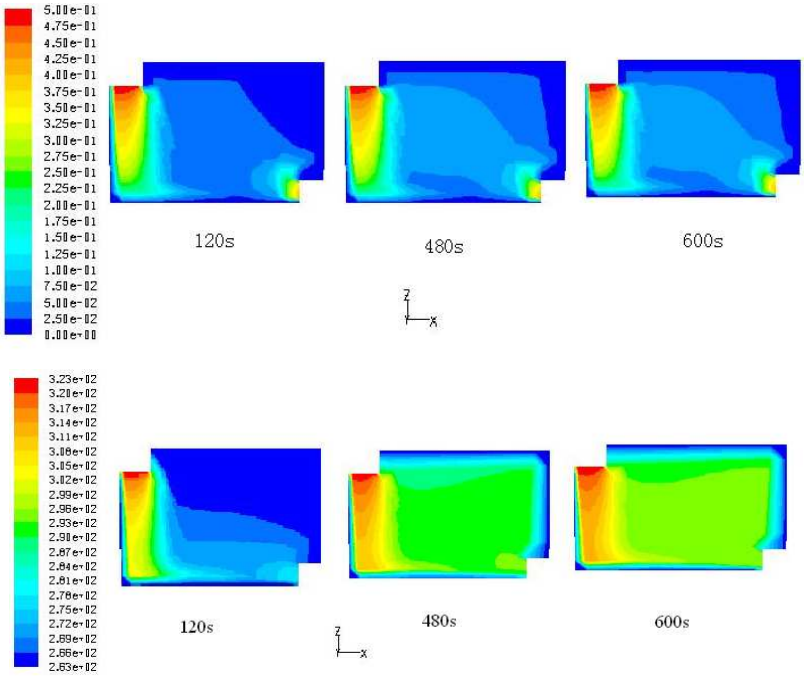


Fig.3 The velocity filed and temperature filed of various time

2.3 Switch door condition

A representative section x = 2.25m (including cold storage door) was selected. From figure 4 we can see, when open cold storage door in very short time the change for temperature field in x direction is very big, so there has intense heat and mass transfer in this direction. When the outdoor air enters into cold storage, because the density of cold air is larger than hot air, the outdoor air rises. It affects the upper flow field. With time advancement gradual, it can influence the entire flow field, but as a result of evaporator's function, the temperature field change in deep is not big.

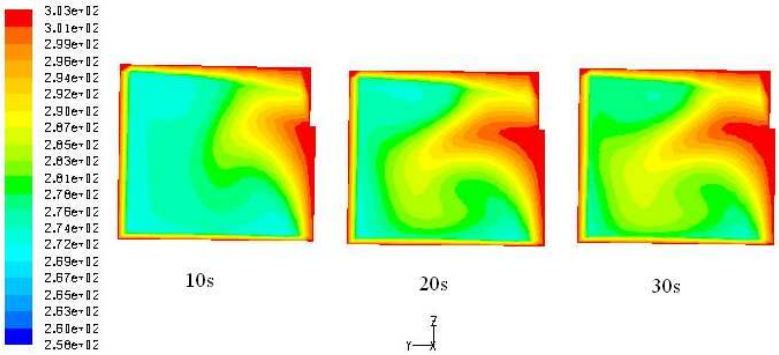


Fig.4 The temperature field when open the door 10s, 20s and 30s

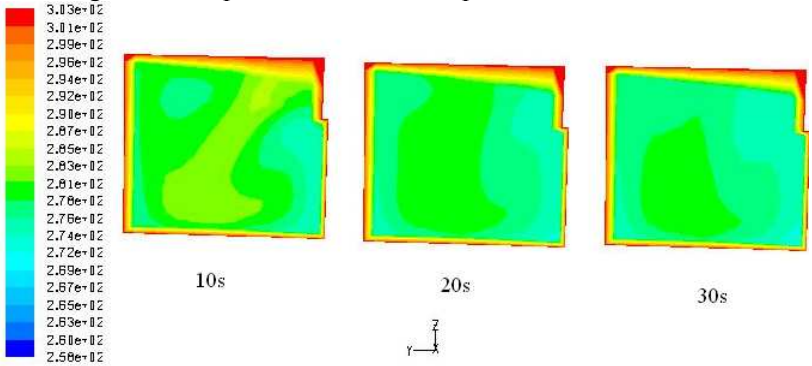


Fig.5 The temperature field when close the door 10s, 20s and 30s

Figure 5 is the temperature field after closing cold storage door 10s, 20s, 30s respectively. From the figure we can see that from the closing moments of the temperature drop rate to be very quick, because the evaporator is the outdoor wind speed of 10 times. As the evaporator is located in the upper

cold storage, after closing the upper part of the temperature field recovery faster, as time goes on inside cold storage temperature gradually returned to steady state condition.

3 Conclusion

The airflow change in cold storage were simulated by FLUENT under following three conditions, pre-cooling, stop to defrost and switch cold storage door. The conclusion is as follows:

(1) Under the same inlet temperature, the cooling rate increase with the evaporator inlet velocity increase and the time achieving steady state conditions decrease. The pre-cooling operating condition should be mainly considered that the two factors between pre-cooling effect and the unit energy consume. At the basis of ensuring the pre-cooling effect, suitable pre-cooling cold wind should be identified to save energy.

(2) The influence of hot air defrosting on flow and temperature fields were simulated. The results show that blowing hot air is a good method to defrost, but the temperature of cold storage rises quickly, and only a small number of endothermic heat uses in the frost level heat absorption. Hence, it should be considered two factors between defrosting performance and energy saving to find the best defrosting time.

(3) The temperature field is violent disturbance when opens cold storage door, the influence is more seriously especially near the door. In order to ensure the steady temperature field near the door and to avoid ice phenomena of the door, the air curtain should be added to obstruct the outdoor air.

References

[1] J. Xie, X. H. Qu, S. Q. Xu. Cold storage of gas flow field numerical simulation and verification. *Transactions of the CSAE*. Vol. 2 (2005), p. 11-16

[2] J. Xie, T. Wu. Numerical simulation on temperature field in the door way of a minitype cold store. *Journal of Shanghai Fisheries University.* Vol. 5(2006), p. 333-339

[3] L. Yang, X. S. Wang. Numerical simulation on temperature filed of a mini type cold store. *Journal of Northwest A&F University.* Vol.9(2008), p. 219-223

[4] A. M. Ji. Numerical simulation and experimental verification of temperature variation during controlled atomosphere cold storage of fruits and vegetables. *Transactions of the CSAE*. Vol. 5(2006), p. 24-27

[5] Z. H. Hu, D. W. Sun. CFD simulation of heat and moisture transfer for predicting cooling rate and weight loss of cooked ham during air-blast chilling process. *Journal of Food Engineering.* Vol. 3 (2000), p. 189-197

[6] G. Cortella, M. manzan, and G. Comini. CFD simulation of refrigerated disp lay cabinets. *International Journal of Refrigeration*. Vol. 3 (2001), p. 250-260

[7] S. Guyot, N. Marney and P. Sanoner et al. Direct thiolysis on crude apple materials for high-performance liquid chromatography characterization and quantification of polyphenols in cider apple tissues and juices. *Methods Enzymol.* Vol. 335 (2001), p. 57-70

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.901

Effect on Galena Oxidation Process of Temperature and Oxide Concentration

Bing.Cao [1.a] Nan.Cui [1.b] Yang.Wu[1.c]

[1]Department of agroforestry engineering,Tangshan Vocational Technical Institute, Tangshan City,Hebei

[a]Caobing0918@sina.com,[b]ambercui@126.com,[c]wuyangabu0808@163.com

Keywords: Galena; Oxidation Kinetics; Reaction Rate

Abstract：In this thesis,a series of oxidation experiments on galena are conducted under different condition such as pH is 2,temperature is 50℃ and 25℃.After the experiment is completed,The rate constant,reaction order，reaction reaction rate and energy of activation are also analyzed and computed.The effects of temperature and oxidized medium on lead ore oxidation speed were studied. Under the condition of 50℃, The rate of oxidation reaction is increase when concentration is 100mg/l.The higher the concentration the more sulfuric acid ion,the easier to produce sediment of lead sulfate , so the reaction will be impeded .the reaction will be extremely slow or even can't carry on at 25℃.

1 Row materials and the method of experiment

1.1 Row materials, experimental facilities and reagents

Using galena as row material and crashing the sample, then picking it out by hand under the imcroscope, the sample whose purity reachs to 99% can be gained. Then sieving the sample, divide it into two groups according their sizes range at 60~80 and 80~100 meshs. Before the experiment, putting these particles into a desiccator and drying them for the next process.

1.2 The method of experiment

The experiment uses $FeCl_3$ as oxidant. The pH of regulation reacting solution is 2. The experiment adopts water bath method. Put the sample of galena and reacting solution into a triangle bottle.Covering the bottle jack with a piece of pledget, it could protect the solution from pollution and keep the pressures inside and outside of the solution same. Each group of experiment consists of four bottles of galena sample (0.1g) within solution at different concentration. Every period of time, extract 5mls of solution from the bottles, and analyse the Pb^{2+} content of the solution with atom absorption spectrophotometer. After the reaction, separate the solution from the remainder galena sample. Analyse the Fe^{3+} content of the separated solution with EDTA volumetric method. According to the vary of the content of Pb^{2+} and ultimately concentration of Fe^{3+}, the oxidation reaction rate constant, order and activation energy would be fixed on. The equation of oxidation reaction kinetics would be set up.

The equation

$$PbS + 8Fe^{3+} + 4H_2O = 8H^+ + SO_4^{2-} + Pb^{2+} + 8Fe^{2+} \qquad (1)$$

Eight experiments divided into two groups have been performed. The concrete conditions of those experiments are:

1) 50℃, pH=2, galena oxidation reaction with different $FeCl_3$ concentration.

2) 25℃, pH=2, galena oxidation reaction with different $FeCl_3$ concentration.

2 Outcomes of experiment and analysis

2.1 Effect on oxidation rate of oxidation medium concentration

(1) The condition is 50℃, pH=2. The respective concentrations of $FeCl_3$ are 100mg/l, 200mg/l, 400mg/l, 600mg/l. To see the results of those experiments in fig 1, G1-1, G1-2, G1-3, G1-4 as four experiments of group 1.

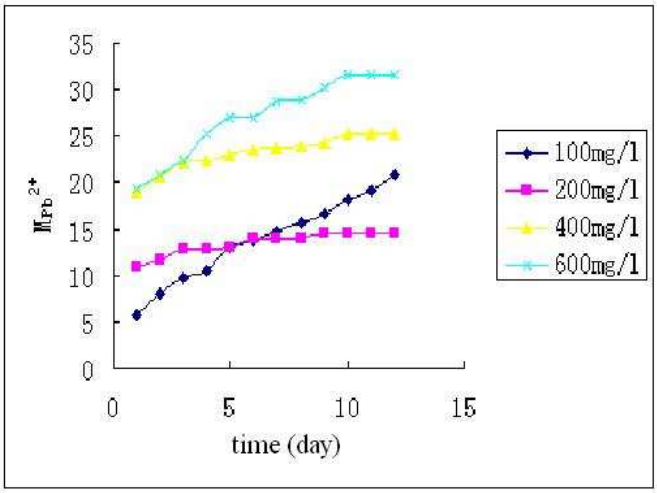


Fig.1 the results of the experiments with different $FeCl_3$ concentrations, 50°C, pH=2

Form.1 the reaction computational results of G1

Experiment number	G1-1	G1-2	G1-3	G1-4
reaction rate (mol/s)	4.9572×10^{-10}	5.7870×10^{-10}	8.0783×10^{-10}	1.0383×10^{-9}

(2) The condition is 25℃, pH=2. The respective concentrations of $FeCl_3$ are 100mg/l, 200mg/l, 400mg/l, 600mg/l. To see the results of those experiments in fig 2, G2-1, G2-2, G2-3, G2-4 as four experiments of group 2.

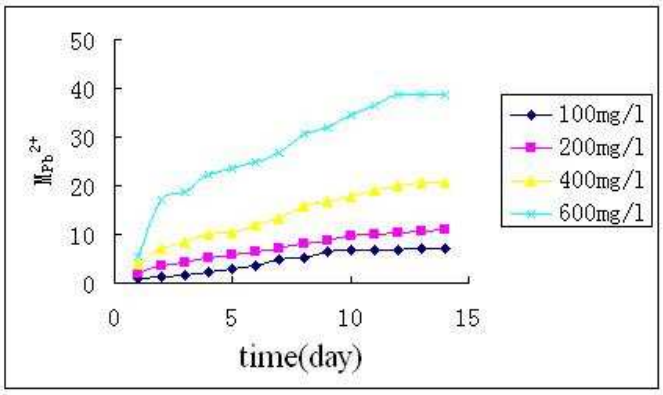


Fig.2 the results of the experiments with different $FeCl_3$ concentrations, 25°C, pH=2

Form.2 the reaction computational results of G2

Experiment number	G2-1	G2-2	G2-3	G2-4
reaction rate (mol/s)	2.4766×10^{-10}	3.3747×10^{-10}	7.3227×10^{-10}	1.7610×10^{-9}

The fig1 and fig2 show that with a higher concentration of $FeCl_3$ the rate of reaction goes faster under the same temperature and pH. When the concentration of the medium is at 600mg/l, the oxidation effect on galena goes best.

2.2 The effect on the rate of oxidation reaction of temperature

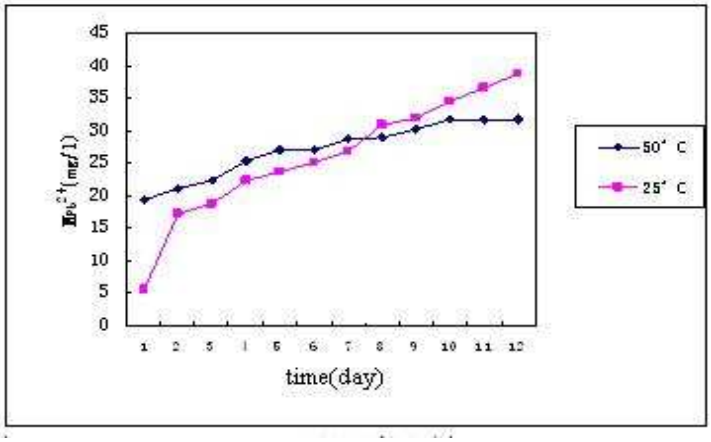


Fig.3 the results of the experiments with different temperatures, $FeCl_3$ 600mg/l

The fig 3 shows that when $FeCl_3$ is used as oxidant, the temperature effect plays more on the rate of reaction. When the temperature is 50℃（G1-4）, the reaction goes fast and the concentration of Pb^{2+} reaches to 19.357 mg/l at the first day. But the final concentration of Pb^{2+} is 5.533mg/l only. At the same time, when the temperature is 25℃（G2-4）, the reaction does not go fast and the concentration of Pb^{2+} reaches to 5.533 mg/l at the first day only. But the final concentration of Pb^{2+} is up to 38.667mg/l. The reason is that there are two reactions in the solution. The main reaction is the oxidation of galena. The side reaction is $Fe^{3+}+OH^-=Fe(OH)_3\downarrow$.When the temperature is higher, the two reactions go faster. Therefore, the concentration of Pb^{2+} of G1-4 is higher than that of G2-4 at the first stage. Then as time goes by, the effect of the side reaction of G1-4 becomes stronger and generates a lot of brown $Fe(OH)_3$ which deposits on the surface of galena. That slows the oxidation rate of galena. So the final concentration of Pb^{2+} of experiment G1-4 is lower than that of experiment G2-4.

2.3 The foundation of the oxidation kinetics equation

The primary equation of experiment kinetics is:

$$R = (S/V) \times k \times [m_{Fe^{3+}}]^n \qquad (2)$$

In this equation, R is the rate of reaction. S is the superficial area of galena.V is the volume of the solution. K is the rate constant of the reaction. $[m_{Fe^{3+}}]$ is the concentration of Fe^{3+}. n is the reaction order.

The key to the foundation of the equation is to find out the rate constant and the reaction order.

(1) Caculating the rate constant and the reaction order of experiment G1

Take the logarithm from the both sides of the equation 2. Take the logarithm of the concentration of Fe^{3+} as abscissa, and take the logarithm of the rate of reaction as ordinate. There would be a straight line. The slope of the line is the reaction order. The intercept is lg(s*k/v). From it the rate constant of the reaction can be gained. See fig 4.

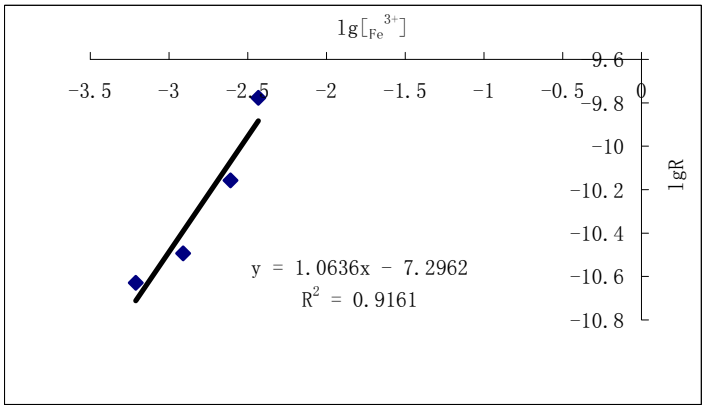


Fig.4 the reaction order of experiment G1

Annotation: the concentration of Fe^{3+} is the final concentration in the solution. The respective concentrations are 17.6858mg/l, 46.0633 mg/l, 88.2436 mg/l and 87.4825 mg/l.

The slope of the line is 1.0636. That is, n=1.0636, lg(s*k/v) =-7.2962. It can be worked out that $k=4.8026*10^{-9}$ and the correlation coefficient R= 0.9571.

（2）Caculating the rate constant and the reaction order of experiment G2

Take the logarithm from the both sides of the equation 2. Take the logarithm of the concentration of Fe^{3+} as abscissa, and take the logarithm of the rate of reaction as ordinate. There would be a straight line. The slope of the line is the reaction order. The intercept is lg(s*k/v). From it the rate constant of the reaction can be gained. See fig 5.

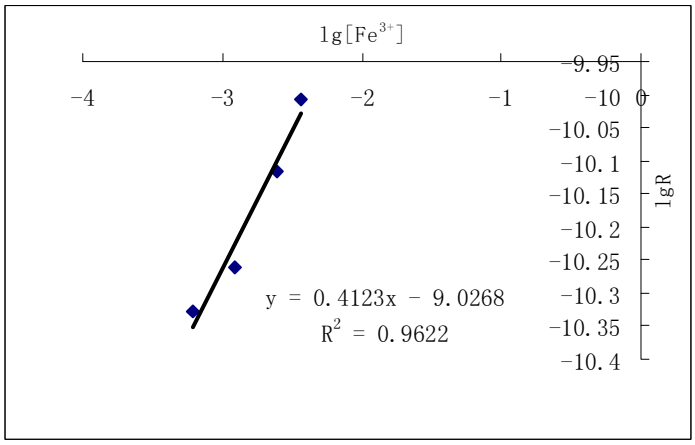


Fig.5 the reaction order of experiment G2

Annotation: the concentration of Fe^{3+} is the final concentration in the solution. The respective concentrations are 7.6794 mg/l, 35.323 mg/l, 78.692 mg/l, and 146.606 mg/l.

The slope of the line is 0.4123. That is, n=0.4123, lg(s*k/v) =-9.0268. It can be worked out that $k=8.9305*10^{-11}$ and the correlation coefficient R= 0.9809.

(3) The activation energy of reaction

$$E = \frac{\lg k_2 - \lg k_1}{T_2 - T_1} \times 2.303 \times 8.314 \times T_1 \times T_2 = 269.806KJ \qquad (3)$$

(4) The foundation of the kinetics equation

According to the experiment reaction order and the reaction rate constant above, the kinetics equation can be founded as below:

The kinetics equation of experiment G1

$$R = (S/V) \times 4.8026 \times 10^{-9} \times [m_{Fe^{3+}}]^{1.0636} \quad (4)$$

The kinetics equation of experiment G2

$$R = (S/V) \times 8.9305 \times 10^{-11} \times [m_{Fe^{3+}}]^{0.4123} \quad (5)$$

In the two equations:

S is the superficial area of galena (cm_2);

V is the volume of the reaction solution (ml).

With the galena oxidation kinetics experiment, the rate of reaction, the rate constant of reaction, the reaction order and the activation energy of reaction have been achieved. The oxidation kinetics equation has been founded. It provides theoretical basis for the research of the oxidation kinetics reaction of sulfide.

3. Conclusion

In this thesis, Fe^{3+} has been used as oxidant. Under the conditions of different temperatures and same pH, experiments of galena oxidation have been done in solution with different concentrations. By analysing the the effect of temperature and concentration on the oxidation rate of galena, the oxidation kinetics parameters of galena have been gained. By caculating out the oxidation rate of galena, the rate constant of oxidation reaction, the order of reaction and reaction activation energy, the equation of galena oxidation kinetics has been founded. Analysing the results of those experiments, some primary conclusion can be drawn:

(1)When pH=2 and the temperature is 25℃ or 50℃,the oxidation reaction goes faster with the concentration of Fe^{3+}, the oxidant, going higher. When the concentration is 600mg/l and the temperature is 25 ℃, the effect of oxidation is better. While the temperature is 50℃, the side reaction generates a lot of brown $Fe(OH)_3$ which deposits on the surface of galena and slows the oxidation reaction.

(2)The equation of galena oxidation kinetics has been founded and would provide theoretical basis for the further research of such experiments.

References

[1]Lu L, Wang R C, Xue J, Chen J, Lu J J. The Surface Reaction of Sulfide Minerals and Its Application to the Study of Mine Environment [J]. Acta Petrologica Et Mineralogica,2001,20 (4):387-394

[2]Tan K X, Zhang Z R, Wang Z G. The chemical oscillation and chaotic attractor of the dissolving reaction of copper pyrites in the solution of NaCl[J].Chinese Science Bulletin,1996,41 (8):720-722

[3]Chen T H, Feng J H, Xu X C. The research on the simulation experiment of efflorescence and oxidation of the sulfide in tailings. Acta Petrologica ET Mineralogica,2002,21 (3):298-302

[4]Lan Y Q, Zhou G, Liu Z H, Huang X. Pyrite oxidation under different conditions. journal of nanjing agricultural university, 2000,1 (23):81-84

doi:10.4028/www.scientific.net/AMM.50-51.906

Produce and Design of the Teaching Website -With an Example of Education Website of Mineral Processing Department

Limei Bai[1, a], Yingchun Wang[1, b], Shujuan Dai[1, c], Fusheng Niu[1, d]

[1]College of Environment and Recourses, Hebei United university, Tangshan, China

[a]limeibai @126.com, bYingchun0601@163.com, [c]shujuandai@163.com, [d]Luckyniu @126.com

Keywords: construction of teaching Website, mineral processing, wibsite design

Abstract. With the development of Network technology and multimedia, applying network to teaching and developing education and research services on the network has become the trend of deportment in teaching. Department of Mineral Processing according to its specialized characteristic designed and made the teaching website. The paper introduced the overall design and produce details of Education Web site of the mineral processing Department. In the paper, it was mainly mention who to design and plan in the pre-development, collect data and make Web page beautify in medium-term, and debug, publish, update and maintenance in the final, and so on.

Introduction

Mineral Processing is an engineering and traditional profession. In order to help students better understand the mineral processing equipment and field practical, and enter professional areas in advance, Department of Mineral Processing devote their efforts to make Teaching Website. Students can know the field and equipment of mineral processing by the columns of teaching and training, technology exchange and scientific research etc. in the Education Web Site of Mineral Processing Department. Students could have intercourse with specialists and technicists experienced who were engaged as Consultants of "College Classroom" columns. The gap between production sites and research institutes with colleges and universities in scientific research was effectively narrowed by the teaching website platform, and on which the combination of production-learning–research was promoted. Students' subjective initiative and innovation ability would be cultured and the learning enthusiasm of college major would be developed by the combination of production - learning – research. The design of Teaching Website is divided into structural design, data collection, graphic design, etc by nature of the work, and in accordance with the design period, can be divided into pre-planning, mid-production, distribution and later maintenance. Therefore, the division of work can be put into practice in the process of making website.[1-3]

The overall site planning and design of the webside

In brief, web page is the basic element of the site. Many web page make up the site, while many website make up the modern Internet Network .The construction of Education Website of Mineral Processing Department is inseparable from the design of web page. Making web pages depend on software. Choosing the suitable tool and compile the right web page which is suit the Teaching Website is the basic task of site construction. [4-6]

Location of Site Visitors. The purpose of Education Website of Mineral Processing Department is to provide a platform for students and teachers of Mineral Processing Engineering to exchange knowledge and experience with other. Main Clients are students of Mineral Processing or interested in it. In order to attract engineers and researchers to participate in this platform, the location of the visitors is important particularly, which not only meet the public taste, but also have the unique of teaching website. However, in colleges and universities, the style of schools and colleges should be taken into consideration. So as the website is not with schools, colleges out of touch. After the formal operation of the site, Site technical staff should leave their E-mail and QQ number to know the views

and suggestions of visitors, so that can adjust and revise the website timely, which could make the website more suitable for Visitors' taste and achieve better results.

election of Site Theme. Rich sites always have the same weak point that is hard to search the information. Lots of site only mind exhaustive which make content red tape, which is hard for e-friend to find information. Therefore from design view, a simple page including the index of the web is used as the homepage of the Education Website of Mineral Processing Department, which is convenient for students and other visitors to find out needed information. In this way, the site can be distinctive and theme-clear.

Design of Site' Overall Style. People' first imagine of site come from visual impact, so it is important to determine site' Standard colors. Different colors with different effects, which may affect visitors' mood. Education Website of Mineral Processing Department mainly use white with tone, which gives white, bright, innocence and clear feels. Text use sky-blue, which can give visitors sense of elegance, and it is helpful to attract visitor and increase the rate of second glance. In addition, style has always been an abstract concept should be combined with the whole site, and different people have different aesthetics. Education Website of Mineral Processing Department is wide-ranging. The site includes Brief Introduction, Teaching and Training, Finding Job, Scientific Research, Student Section, The place of answer questions and so on. The site should have the same style. Since single hue and background will give visitors the sense of monotonous, different columns should have different element. Let the lively animation section and serious knowledge exchange have different styles can make people comfortable.

Details of Website making [5,7]

Browser's Version and Resolution. Different browsers have different page-show. The browser, which is beautiful in Internet Explorer, may perform poor in Netscape Navigator. In the IE Age, we should also consider minority users which can be your potential visitors. Each webpage was put into two browsers to test and solve the question timely. It is better to consider old edition's compatibility, because some page may not fit the old edition. Nowadays, the resolution of 1024×768 is most popular, but lots of user use 800×600, it is needed to consider more visitors requirement, which are all in the top priority.

Design about CI and Banner. Logo of site likes trademark. Logo reflects the characteristic and content of your site. We can recall your site when we see the logo. Logo, which can be Chinese, English, mark, pattern, sign, tint, font and poster, is the key of a site set up the image of CI. Precisely, it is the face of site. When you design and accomplish the steps, your site will have a complete change and increase a lot in the whole image. In order to increase students' enthusiasm and creative ability, Department of Mineral Processing held a Site Logo Design Contest. We invite some specialized teachers and art expert as jury to select the logo of the Education Website of Mineral Processing Department. In order to let the logo meet international standards, we make a lot of changes. In the final, the logo contents fashion elements, traditional charm and the element of university and college. Logo's English is eye-catching and clear and Chinese is easy-to-understand and well arranged which make it easy-remembered.

Idea of whole site should perform the design of Banner. Certainly hue is also important. In the whole consideration, the Education Website of Mineral Processing Department considers the idea fully. Banner designed by FLASH belongs to dynamic effects which give people the sense of mobile and gains more visitors.

Structure of Mineral Processing Department of Education Website. We spent one year perfecting *the structure of* Education Website of Mineral Processing Department. At first, the structure of site was confirmed by the research of our students. Since this teaching website not only attracts students who are in this professional, we continuous improve it under visitors' message and suggestion. In final, we definite the structure like table 1.

Exploit of Website. At present, the Education Website of Mineral Processing Department mainly used Dreamweaver and Firework to complete the page which omit the trouble of the compiling of HTML language by hand and get what you see, and speed of page-making was improved. Providing interactive access capability and realizing information exchange is the key of development work. In addition, Macromedia professional software didn't always conform to the requirement of development of all in the site. In order to meet the improving requirement of users, and with the improving condition of our department, the secondary development of the site usually must be.

Development tools of Education Website of Mineral Processing Department usually are Dreamweaver, Photoshop, Fireworks, and Flash etc. In the page, we use JavaScript language to realize interactive function. We use ASP.net technique in forum development and exchange to realize the function of access database. Delphi used in where need high efficiency and was complicated. The development of mineral processing site will be accompanied by the existence of site, development of Department of Mineral Processing, updating of computer technology.

Details and Skills in Designment. Standard font is Special font which is used in symbol, title and main menu. In order to reflect the different and special style of the site, the Education Website of Mineral Processing Department selects some special font according to special requirement. For example, in order to reflect professional, we used Crude italics, in order to reflect beautifully designed, we used Ad form, in order to reflect affability, we used handwritten form and so on. There were hundreds of English Fonts and some Special English Art Font Download, so it is easy to find a satisfactory font. We need to emphasize that there are not special fonts in visitor's PC, so Non-default fonts should use picture form. When design the charts of the page, don't put the whole page in one chart, it is better to put it into different charts. The structure of single chart should be tidy and the chart's Nested hierarchy few as far as possible. Because it is too slowly to download when the Layer of chart is tedious, the size of the picture should not be too big. We should control the size around 6KB.

Table 1Structure of Education Website of Mineral Processing Department

<table>
<tr><td rowspan="20">MINERAL PROCESSING DEPARTMENT of EDUCATION WEBSITE</td><td>Introduction</td><td>Overall</td><td rowspan="20">MINERAL PROCESSING DEPARTMENT of EDUCATION WEBSITE</td><td rowspan="3">Ranks of Teachers</td><td>Professors</td></tr>
<tr><td rowspan="4">Teaching and Training</td><td>E-teaching Plan</td><td>Instructors</td></tr>
<tr><td>Discipline Construction</td><td>Engineers</td></tr>
<tr><td>Online Classroom</td><td rowspan="4">Student Corner</td><td>Professional and Class</td></tr>
<tr><td>Course Arrangement</td><td>Innovation in Science and Technology</td></tr>
<tr><td rowspan="4">Scientific Research</td><td>Direction of Research</td><td>Precious Times of Past</td></tr>
<tr><td>Projects under Research</td><td>Alumni Style</td></tr>
<tr><td>Laboratory</td><td rowspan="4">Technical Exchange</td><td>Rudiments</td></tr>
<tr><td>Achievements In Scientific Research</td><td>Academic Frontier</td></tr>
<tr><td rowspan="5">Apply For Job and Graduate</td><td>Science Engineering Graduate</td><td>Science and Technology Service</td></tr>
<tr><td>Unit recruitment</td><td>Online Exchange</td></tr>
<tr><td>Student job Search</td><td rowspan="3">Question Place</td><td>Scientific Research</td></tr>
<tr><td>Advice from Others</td><td>Teaching Research</td></tr>
<tr><td>Graduate Style</td><td>Contact and Exchange</td></tr>
<tr><td>Tendency in Department</td><td>Latest Tendency</td><td>Landing</td><td>Mineral Forum</td></tr>
</table>

Safety and Maintenance of Site

The completion of Construction of Education Website of Mineral Processing Department didn't mean once and for all. The maintenance workjavascript: showjdsw ('jd_t','j_') must be done usually. Not only the content of site but also the style of homepage must update regularly. Gain new content and delete the hyperlink and information which are already invalidation. Only in this way the Education Website of Mineral Processing Department could have exuberant vitality and can't be swallowed by the vast flood of information on Internet.

While the safety of Education Website of Mineral Processing Department can't come up with military affairs and economy, various types of data collected during teaching and research is also important. The data is the blood and sweat of whole teachers and students, is the life of whole Research and teaching. Therefore, it's necessary to make corresponding network security measures. Limit the access to inner Network Server and set up gateways during intranet and extranet are general practice during internal management system of Mineral processing Department and WEB Server. [8]

Conclusion

The Education Website of Mineral Processing Department is a link of the students, teachers and others in Mineral Processing field. It also is a bridge between Mineral Processing Department and outer world. The website meets maximatily the demand of different people in the society and supplies the modern scientific and technological services to teachers and researchers, and improves work of teaching and researching in school, lets students take part in researching discussion fully, cultures Mineral Processing personnel who have high professional quality and competence and innovation. At the same time, it serves to the development of Department of Mineral Processing Subject.

References

[1] Y.P. Cai, "Strengthening the construction site, promoting education information," Jounal of HuNan Business College, vol. 12, no. 6, pp. 120-122, December 2005

[2] G. F. Shi, "Discussion Vocational School Website," Jounal of Shaanxi Railway Institute, vol. 7, no. 2, pp. 23-25, December 2009

[3] H. W. Bai, et al, "Discussion the Construction of Zhangjiajie Aviation Institute Teachers' Space," Guangxi Journal of Light Industry, vol. 11, no. 132, pp. 88-89, December 2009

[4] F. Chen, Web Design Three Musketeers. Shanghai: Shanghai Popular Science Press, 2005,pp.10-30

[5] D. Y. Hu, Web Designer Tutorials. Beijing: Tsinghua University Press, 2001

[6] N. Sun, F. Wei, and J. f. Liu. "The Key Technology and Application of Alumni Website Construction" Electronic Sci. & Tech., vol. 23, no. 6, pp. 94-96, Jun. 2010

[7] X. D. Kang, Enterprise Business Website complete solution. Beijing: Posts & Telecom Press, 2006

[8] L. J. Ding, "Construction of University Website Based on CMS Value Engineering Value Engineering," Value Engineering, vol. 9, no. 11, pp. 159-160, November 2008

doi:10.4028/www.scientific.net/AMM.50-51.910

Control and utilization method for urban rain-flood

Jun YAN[1, a], Changyou JIANG[2, b], Sainan CHEN[1, c], Ju ZHANG[1, d]

[1]North China University of Water Conservancy and Electric Power, Zhengzhou, China

[2]PLA Information and Engineering University, Zhengzhou, China

[a]email: yanjun@ncwu.edu.cn, [b]jiangcy@pla.edu.cn, [c]chensainan@ncwu.edu.cn, [d]zhangju@ncwu.edu.cn

Keywords: rain-flood; mechanism; utilize; countermeasure

Abstract. Urban rain-flood disasters become more ordinary and serious recently under global warming and unreasonable exploitation. Summed up the characters of urban rain-floods, the mechanism of urban rain-flood disasters is analyzed. More ever, the main countermeasures to control and utilize urban rain-flood comprehensively and effectively are put forward.

The concept of "urban rain-flood disaster"

The urban rain-flood disasters became ordinary recently, but here is not a clear definition about rain-flood in "Ci Hai". The definition of "rainwater" and "flood" has multi-term. The explanations of "rainwater" are: 1) falling water such as rain; 2) the water form rain; 3) one of 24 festivals according to Chinese lunar calendar; 4) rain. The explanations of "flood" are: 1) water level rises up high enough and floods the fields; 2) the river level rises suddenly and sharply by reason of heavy rain storms or by melting snow, which often causes disaster; 3) flood.

Base on the definition of "rainwater" and "flood", "rain-flood" can be understood as the flood beyond the storage capacity caused by high intensity rainfall and it can be used as resources and can lead to disaster either. The urban rain-flood disaster refers to the disaster produced by high intensity urban rainfall.

The characters of Urban rain-flood and urbanization

For the repeated and intensive urban rain-flood disasters, many developed countries have developed the research of the urban rain-flood disasters in recent years.

According to its definition, urban rain-flood can be used as resources and can lead to disaster either. Accordingly, the urban rain-flood disaster has the obvious characters.

On the one hand, insufficiency of water resources will be a direct threat to human survival. Therefore, the resource of urban rain-flood will attract more and more attention. The main benefits of urban rain-flood are as follows:

1) Economic benefits. Purified water is used as water resources for living, municipal use and chemical plant, which can save the government and residents water costs; treated rain-flood is used to develop aquaculture, agriculture, stock farm, which can increase people's income and improve economic efficiency; rain-flood is used to develop tourist attractions, water entertainment, which can create comfortable living space as well as wealth at the same time;

2) Social benefits. Rational use of rain-flood can not only alleviate the urban water load, but also decrease the possibility of urban floods caused by rainfall[3], more ever, the employment problem can be partly solved by the industry cycle from collection to use;

3) Environment benefits. Rain-flood can increase the quantity of urban areas, decrease the subsidence of water surface, enhance the water level of shallow ground water and maintain the stability of the shallow ground water, to complement the surface water and groundwater and to maintain the virtuous cycle of the surface vegetation and microorganisms systems [3]. After simply treated, the urban rain can be directly used to spray the road, water plant, dusting down the clean air, which can improve the ecological environment and living environment. The canals of the city,

man-made lake, reservoir and reservoirs can be united and supply water effectively, which can therefore adjust the urban microclimate. The comprehensive system of rain-flood can be developed to generate electricity, culture fish, more ever, even the industry of livestock and agriculture could be significantly increased by the rain resources.

With the persistent decrease of per capita available natural resources, the human and social development will be seriously affected.

On the other hand, if the rain-flood beyond the urban storage capacity caused by high intensity rainfall it may lead to intensive disaster. The main reasons are as fellows:

1) The rain-flood is uneven. The distribution of rain is usually uneven and intermittent on time and space, which will cause high intensity rain-flood beyond the urban storage capacity and cause disaster;

2) The storage of rain-flood collecting system isn't great enough. The jam-packed urban architectures become more and more, which make the urban area lack of enough region to save flood in time.

The mechanism of Urban rain-flood disaster

The mechanism of urban rain-flood disasters can be illustrated as shown in figure 1.If the intensity of rain-flood is higher than the up limitation it will cause flood disaster, contrarily, it will cause draught disaster if the volume of rain-flood is less than the down limitation.

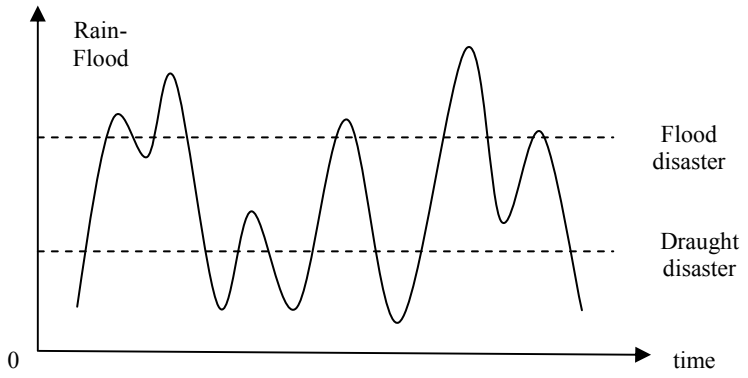


Figure1. The mechanism of urban rain-flood disasters

Analysis utilization of rain-flood shows that the utilization of urban rain-flood requires the whole use of water collection systems, water conveyance systems, water storage systems, water purification systems and to develop laws to resolve the problems concerning with rainwater collection, transport, storage, purification and management:

1) Water collection system. Aim at building a beautiful, clean, ecological integrity and modern city, the urban greening level was increased gradually, therefore the lawn can be used to collect rain-flood, purify water and supply the water resource; rain-flood collection and storage system can be built in the densely populated area of urban area, which can lead the rain-flood on roofs, courtyards, streets, squares and lawns into the rain-flood storage areas;

2) Water conveyance system. the existing drainage system should use reasonably in urban districts and the water pipelines should be equipped dedicatedly according to the actual situation, the different water conveyance systems should be planned based on geological conditions in new district;

3) Water storage system. The lower areas should be chosen to construct reservoirs, man-made lakes and canals to accumulate rainwater, especially the existing lake should be put into the city developing plan to lower the project's cost;

4) Water purification system. The clean water plants, sewage treatment plants should be build in the city center and suburbs respectively, therefore, the rainwater can be accumulated, purified and made full use of;

5) The relevant laws. The relatively sound laws on rain-flood utilization should be established, the intelligent control system should be developed to overall control and partially grasp, or manual

operating system, the scope of water-use from the chemical plant, agriculture, stock farm, fisheries, power, entertainment and so on, should be expanded to monitor and manage the rain-flood transport, collection and storage, purification and output systems.

The problems of urban rain-flood utilization and control

Although it can bring great benefits to economy, society and environment, etc., urban rain-flood is restricted by time, space, man-made factors and system faultiness, etc. The main problems of urban rain-flood utilization and control are as follows:

1) The rain-flood is not durative. The distribution of rain is usually uneven and intermittent on time and space, which will cause the exploitation rain-flood become seasonal and in local area scope, therefore, rain-flood can hardy be used as constant water supply;

2) Water collecting systems are normally deficient. Since most of the existing urban drainage systems are not unified, they can hardly gather the rain effectively and be used comprehensively in large-scale during the rain-flood collecting process;

3) The storage of rain-flood collecting system is not great enough. The jam-packed urban architectures become more and more, which make the urban area lack of enough region to save flood. The rain-flood resources of higher value exploitation are un-durative, short-term and paroxysmal, more ever, the insufficient of rain-flood saving region results in the storage capacity is not strong enough and difficult to ensure the urban rain-flood resources be used in large scale universally and chronically;

4) Re-pollute. Re-pollute means during the process of the runoff, the dissolvable pollutants or solid pollutants in large range are resolved by the runoff and enter the flood storage area, which causes water pollution;

5) Limited usage. The uses of urban rain-flood currently are just for flood control or domestic water. These kind of narrow and single usage baffles the development of rain-flood;

6) The legal system is not sound enough. The study about the utilization of urban rain-flood is comparatively late and most of the research achievement are still in the stage of theory and can not arouse widely communities' concern; More seriously, the relevant encouraging and supporting policies are not enough and the relevant legislation are not strong enough, which seriously obstructed the development of the use of urban rain-flood comprehensively.

The main countermeasures

Because of the strong characters of seasonality, territory limitation and intermittence, rain-flood resources can hardly be used directly as a constant water supply. But after adopting related engineering measures, rain-flood can take over other headwaters and be used as assistant headwaters. Aim at solving the main questions of the local rain-flood using and improving urban security, the main countermeasures are shown as figure 2.

1) Improve the rain collecting system. In the urban crowd circumstance, the roofs are huge and can be used to store rainwater. Water pipes on all sides of walls can be established to conduct the rainwater into the underground tube or open channel, aim at leading the rainwater flow into lower reservoir to be used as unified savings. In addition, on the main street and square, on which the harden area are greater, the penetrative pavement can be spread and the collecting pipe can be set under the main street and square to induce the water to the underground collecting pipe into storage pool. The whole system is set up under the city spread areas and doesn't influence the city sight. The grass lawn can obstruct rain water and creates good living environment in the city. Although the grass area is built up higher than urban road normally, the paling should better be heighten and used together with open channel or underground ditch, to lead rainwater into the reservoir, which can not only alleviate the flood control burden of city area in flood season, but also store rain-flood and utilize it synthetically;

2) Enhance the efficiency of rain-flood transport. Along with the speeding of urbanization progress, the urban extension in new city or reformation in old city is carried out currently on large

scale in most domestic cities. In the new extension area and reformation area the conduit can be merged directly into urban layout programming to service the rain-flood utilization. The conduit buried in advance can be divided as rain gathering pipe and sewage pipe according to the function. The lightly polluted rain-flood, domestic and municipal waste water and heavily polluted chemical plant sewage can be transported respectively, which can simplify the coming water processing procedure and improve the water transport efficiency;

3) Enhance rain-flood storage capacity. Select relatively lower urban areas in each urban district and dig smaller pools to store rain-flood, through the underground branch pipes and canals (open channels). Canals should be built according to urban real terrain, geological conditions, building structure, reasonable human and environmental planning. In the relative lower place of city center, larger reservoir (man-made lake) can be dug deeply to connect the underground branch pipes and canals, in order to construct the whole water transport system to store the rain-flood effectively and deliver water smoothly. As to the whole flood control, the larger reservoirs which located in the lower suburb can be constructed to increase the urban storage capacity for rain-flood. Suburbs and the inner city's reservoirs can be connected with the river outside the city to form the channel net which can enhance the flood control capacity and store more rain-flood under the worse condition;

4) Dispose the sewage respectively. The purification treatment plant can be built in the center of the man-made lake to collect and dispose the light polluted rain-flood, which can be used as life and municipal water after purification process treatment. The sewage treatment plants can be constructed in suburban reservoirs to dispose the heavily polluted water which can be used to meet the water demand of peripheral industry and agriculture, stock farm and aquaculture;

5) Utilize the rain-flood comprehensively. Deep man-made lakes can be dug in lower district meanwhile the man-made hill can be heaped up accordingly accompanying with water recreation places to developed the tourist scenic spots. On the hills the pumped storage power stations and man-made waterfalls can be built. In the lower suburbs and medium reservoirs the aquaculture, agriculture and stock farm can be developed spaciously;

6) Perfect the legal system. Rain-flood utilization is the requirement of both flood mitigating and environment improving. The government should fully exert the function and take all kinds of measures to establish the sound legal system, such as to set up special funds and establish favorable policies, mobilize the enthusiasm of urban constructers, developers and related enterprisers for rain-flood utilization. While the enterprisers and individuals who disobey the laws and regulations on rain-flood utilization should be punished. Therefore, the legal system can be established and perfected gradually to ensure the healthy development of rain-flood utilization.

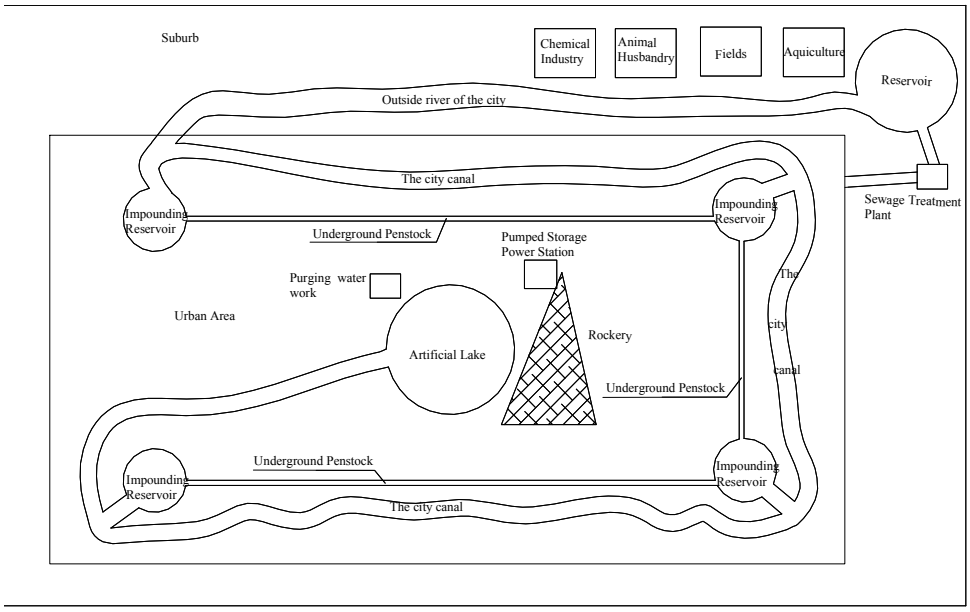


Figure 2. The sketch of urban rain-flood utilization and control

Summary

The urban rain-flood disasters became more ordinary and serious recently under the unreasonable urbanization development and the global warming. If the intensity of rain-flood is out of the limitation it will cause flood or drought disaster. Restricted by time, space, man-made factors and system faultiness, etc., urban rain-flood is not easy to be used and controlled. According to the mechanism of urban rain-flood disaster, the processes of collecting, transporting, storing, purifying and utilizing can be constructed synthetically to form the comprehensive system, by which rain-flood can be used effectively and controlled commendably.

Acknowledgement

Project supported by the National Natural Science Foundation of China (Grant No. : 50509008), the Natural Science Foundation of Henan Province (2007570006 & 2009B570002), Foundation of Technology Innovation Talents in University of Henan Province (2009HASTIT025), Senior Talents Foundation of North China University of Water Conservancy and Electric Power, etc.

References

[1] Shefang YIN. The Discussion of A Number of Issues about Beijing Urban Rain-flood Utilization [J]. Beijing Water, 2009(1):77-78

[2] Gensheng LI, Fengxia LIU. The Discussion of A Number of Issues about Urban Rain-flood Utilization [J]. Water Science and Engineering Technology, 2006(1):30-31

[3] Weiqing LI,Jie ZHANG,Yanli LI. Analysis and Exploration of rainwater utilization in Henan Province [J]. Modern Agricultural Technology, 2009(13):268-269

[4] Tao GUO. Urban Water Conservancy and Urban Hydraulics [J].Chengdu Water Conservancy, 1995(4):24-28

[5] Xiang GAO,Guodong WANG, Ming ZHANG. On The Planned Use of Urban rain-flood[J]. Shanxi Architecture, 2008 (26):180-181

[6] Yuqi ZHONG,Huixian CAO. The Countermeasure of Rainwater Utilization in Urban Beijing [J]. Beijing Water Science and Technology, 1991(4):21-25.

[7] Weilin LIU,Zengchuan CAO, Zhongmin LIANG. Storm and Flood Similarity Analysis and Application Study [J]. China Rural Water and Hydropower, 2007(2):132-135

[8] Shihong LIU, Xiange CUI, Fengtai GUO. The Discussion of The Northern City of Rainwater Run-off Water Quality and Use Patterns [J]. Beijing Water Science and Technology, 2008 (35):15467-15469.

[9] Jianke DONG, Songbo LIU, Yu SUN.The Use Discussion of The Northern City of Rainwater[J]. Haihe River, 2009(5):5-7

[10] Yuheng CHEN. The Concept of Urban Rainwater Utilization [J]. Water Resources Development, 2002(4):32-33

[11] Mingwen LI. The Rainfall Analysis of Urban Rainwater Utilization [J].Ji Lin Water Conservancy, 2009(2):40-42

doi:10.4028/www.scientific.net/AMM.50-51.915

Delay-dependent robust stability of cellular neural networks with time-varying discrete and distributed delay

Weiwei Wang [1, a], Weiwei Su [2, b] Yiming Chen [3,c]

[1] College of Science, Hebei North University, Zhangjiakou Hebei, China

[2]Tianjin Broadcom genius Technology Co., Ltd. Tianjin China

[3] College of Science, Yanshan University, Qinhuangdao, Hebei, China

[a]email：wangweiweibetter@163.com, [b]jashonbetter@163.com, [c]chenym@ysu.edu.cn

Keywords: cellular neural networks; delay-dependent robust stability; time-varying discrete and distributed delay.

Abstract. Delay-dependent robust stability of neural networks with discrete and distributed delays is considered in this paper. Stability criteria are derived in LMIs avoiding bounding certain cross terms and the restriction of derivative of time-varying delay is removed. Numerical examples are given to indicate significant improvements over some existing results.

Introduction

Neural networks with time-delays are often used to describe dynamic systems due to its practical importance and wide applications in many areas such as industry, biology, economic and so on. So in the past few years, the stability problem of neural networks with time-delays has been attracted a large amount of research interest. On the other hand, the uncertainties in various neural networks are unavoidable thanks to modeling errors, measurement errors and so on.

Different classes of neural networks such as Hopfield neural networks, cellular neural networks, Lotka-Volterra neural networks, Cohen-Grossberg neural networks, and bi-directional associative memory neural networks have been extensively studied in [1-3]. In our paper, we discuss the cellular neural networks with time-varying delays. The cellular neural networks (CNN) are introduced in [1], and a CNN with a delay (DCNN) is first reported in [2]. It's well known that time-delay can easily cause instability and oscillations in system. Existing criteria for asymptotic stability of time delay neural networks can be classified into two types: that is, delay-independent stability [4-6].

In our paper, we consider delay-dependent robust stability for cellular neural networks with time-varying discrete and distributed delays. First we give the delay-dependent stability criterion for nominal cellular neural networks. Delay-dependent stability criterion is derived as LMI without using The effectiveness of the proposed stability criteria and their improvement over some existing results are illustrated in numerical examples.

System description

Consider the following uncertain delayed neural networks with n neurons:

$$\dot{x}(t) = -C(t)x(t) + A(t)f(x(t)) + B(t)f(x(t-d(t))) + D(t)\int_{t-\tau(t)}^{t} f(x(s))ds + I \tag{1}$$

where $x(t) = [x_1(t), \ldots, x_n(t)]^T \in R^n$, $f(x) = [f_1(x_1(t)), \ldots f_n(x_n(t))]^T$ is the neuron activation function, $f(x(t-d(t))) = [f_1(x(t-d(t))), \ldots f_n(x(t-d(t)))]^T$, $C = diag(c_i)$ with $c_i > 0$, A is the feedback matrix, B and D represent the discretely delayed connection weight matrix and the distributive delayed connection weight matrix, respectively. while $I = [I_1, \ldots, I_n]$ are the constant

inputs. The delay considered is time-varying, which satisfies $0 \le d(t) \le d$ and $\dot{d}(t) \le \mu$, $\tau(t)$ is the distributed time delay satisfying $0 \le \tau(t) \le \tau$.

Lemma 1. Let E, F and H be real matrices of appropriate dimensions with F satisfying $F^T F \le I$. Then for any scalar $\varepsilon > 0$

$$EFH + H^T F^T E^T \le \varepsilon^{-1} EE^T + \varepsilon H^T H$$

Stability criteria for delayed neural networks

We consider the stability of the nominal system

$$\dot{z}(t) = -Cz(t) + Ag(z(t)) + Bg(z(t - d(t))) + D\int_{t-\tau(t)}^{t} g(z(s))ds \tag{1}$$

Theorem1. The system (1) is stabile if there exist $P > 0, R > 0, W > 0, Q_1 > 0, Q_2 > 0, \varepsilon > 0$, $L = diag\{l_1, ..., l_n\} \ge 0$, $S = diag\{s_1, ..., s_n\} \ge 0$, $M = diag\{m_1, ..., m_n\} \ge 0$, and any appropriate dimensional matrices $Y_i\ (i = 1,2,3,4,5), T_i\ (i = 1,2,3,4,5)$ such that

$$\Xi_1 = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} & \Phi_{14} & \Phi_{15} & -dY_1 & \tau T_1 D - \varepsilon F_1^T F_4 & T_1 E \\ * & \Phi_{22} & \Phi_{23} & \Phi_{24} & \Phi_{25} & -dY_2 & \tau T_2 D & T_2 E \\ * & * & \Phi_{33} & \Phi_{34} & \Phi_{35} & -dY_3 & \tau T_3 D + \varepsilon F_2^T F_4 & T_3 E \\ * & * & * & \Phi_{44} & \Phi_{45} & -dY_4 & \tau T_4 D + \varepsilon F_3^T F_4 & T_4 E \\ * & * & * & * & \Phi_{55} & -dY_5 & \tau T_5 D & T_5 E \\ * & * & * & * & * & -dR & 0 & 0 \\ * & * & * & * & * & * & -\tau W + \varepsilon F_4^T F_4 & 0 \\ & * & * & * & * & * & * & -\varepsilon I \end{bmatrix} < 0 \tag{2}$$

$$\Phi_{11} = Q_1 + Y_1 + Y_1^T - T_1 C - C^T T_1^T + \varepsilon F_1^T F_1 \quad \Phi_{12} = -C^T T_2^T - Y_1 + Y_2^T$$

$$\Phi_{13} = T_1 A - C^T T_3^T + Y_3^T + HS - \varepsilon F_1^T F_2 \quad \Phi_{14} = T_1 B - C^T T_4^T + Y_4^T - \varepsilon F_1^T F_3$$

$$\Phi_{15} = P - T_1 - C^T T_5^T + Y_5^T \quad \Phi_{22} = -(1-\mu)Q_1 - Y_2 - Y_2^T \quad \Phi_{23} = T_2 A - Y_3^T$$

$$\Phi_{24} = T_2 B - Y_4^T + HM \quad \Phi_{25} = -T_2 - Y_5^T \quad \Phi_{33} = Q_2 + T_3 A + A^T T_3^T - 2S + \tau W + \varepsilon F_2^T F_2$$

$$\Phi_{34} = T_3 B + A^T T_4^T + \varepsilon F_2^T F_3 \quad \Phi_{35} = L - T_3 + A^T T_5^T$$

$$\Phi_{44} = -(1-\mu)Q_2 + T_4 B + B^T T_4^T - 2M + \varepsilon F_3^T F_3 \quad \Phi_{45} = -T_4 + B^T T_5^T \quad \Phi_{55} = dR - T_5 - T_5^T$$

Proof. Replacing C, A and B in

$$\Xi = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} & \Phi_{14} & \Phi_{15} & -dY_1 & \tau T_1 D \\ * & \Phi_{22} & \Phi_{23} & \Phi_{24} & \Phi_{25} & -dY_2 & \tau T_2 D \\ * & * & \Phi_{33} & \Phi_{34} & \Phi_{35} & -dY_3 & \tau T_3 D \\ * & * & * & \Phi_{44} & \Phi_{45} & -dY_4 & \tau T_4 D \\ * & * & * & * & \Phi_{55} & -dY_5 & \tau T_5 D \\ * & * & * & * & * & -dR & 0 \\ * & * & * & * & * & * & -\tau W \end{bmatrix} < 0 \tag{3}$$

$$\Phi_{11} = Q_1 + Y_1 + Y_1^T - T_1 C - C^T T_1^T \quad \Phi_{12} = -C^T T_2^T - Y_1 + Y_2^T \quad \Phi_{13} = T_1 A - C^T T_3^T + Y_3^T + HS$$

$$\Phi_{14} = T_1 B - C^T T_4^T + Y_4^T \quad \Phi_{15} = P - T_1 - C^T T_5^T + Y_5^T \quad \Phi_{22} = -(1-\mu)Q_1 - Y_2 - Y_2^T$$

$$\Phi_{23} = T_2 A - Y_3^T \quad \Phi_{24} = T_2 B - Y_4^T + HM \quad \Phi_{25} = -T_2 - Y_5^T$$

$$\Phi_{33} = Q_2 + T_3 A + A^T T_3^T - 2S + \tau W \quad \Phi_{34} = T_3 B + A^T T_4^T \quad \Phi_{35} = L - T_3 + A^T T_5^T$$

$$\Phi_{44} = -(1-\mu)Q_2 + T_4 B + B^T T_4^T - 2M \quad \Phi_{45} = -T_4 + B^T T_5^T \quad \Phi_{55} = dR - T_5 - T_5^T$$

with $C+\Delta C(t)$, $A+\Delta A(t)$ and $B+\Delta B(t)$, $\Delta C(t)$, $\Delta A(t)$ and $\Delta B(t)$ are described, we get

$$\begin{aligned}&\Xi+\begin{bmatrix}-F_1 & 0 & F_2 & F_3 & 0 & 0 & F_4\end{bmatrix}^T H^T(t)\begin{bmatrix}E^TT_1^T & E^TT_2^T & E^TT_3^T & E^TT_4^T & E^TT_5^T & 0 & 0\end{bmatrix}\\&+\begin{bmatrix}E^TT_1^T & E^TT_2^T & E^TT_3^T & E^TT_4^T & E^TT_5^T & 0 & 0\end{bmatrix}^T H(t)\begin{bmatrix}-F_1 & 0 & F_2 & F_3 & 0 & 0 & F_4\end{bmatrix}<0\end{aligned} \tag{4}$$

From Lemma and Schur complement, condition (4) and (3) are equivalent.

We consider the stability of the nominal system

$$\dot{z}(t)=-Cz(t)+Ag(z(t))+Bg(z(t-d(t)))+D\int_{t-\tau(t)}^{t}g(z(s))ds \tag{5}$$

Theorem1. The system (5) is stabile if there exist $P>0, R>0, W>0, Q_1>0, Q_2>0, \varepsilon>0$, $L=diag\{l_1,...,l_n\}\geq 0$, $S=diag\{s_1,...,s_n\}\geq 0$, $M=diag\{m_1,...,m_n\}\geq 0$, and any appropriate dimensional matrices $Y_i\ (i=1,2,3,4,5)$, $T_i\ (i=1,2,3,4,5)$ such that

$$\Xi_1=\begin{bmatrix}\Phi_{11} & \Phi_{12} & \Phi_{13} & \Phi_{14} & \Phi_{15} & -dY_1 & \tau T_1D-\varepsilon F_1^TF_4 & T_1E\\ * & \Phi_{22} & \Phi_{23} & \Phi_{24} & \Phi_{25} & -dY_2 & \tau T_2D & T_2E\\ * & * & \Phi_{33} & \Phi_{34} & \Phi_{35} & -dY_3 & \tau T_3D+\varepsilon F_2^TF_4 & T_3E\\ * & * & * & \Phi_{44} & \Phi_{45} & -dY_4 & \tau T_4D+\varepsilon F_3^TF_4 & T_4E\\ * & * & * & * & \Phi_{55} & -dY_5 & \tau T_5D & T_5E\\ * & * & * & * & * & -dR & 0 & 0\\ * & * & * & * & * & * & -\tau W+\varepsilon F_4^TF_4 & 0\\ & * & * & * & * & * & * & -\varepsilon I\end{bmatrix}<0 \tag{6}$$

$$\Phi_{11}=Q_1+Y_1+Y_1^T-T_1C-C^TT_1^T+\varepsilon F_1^TF_1\quad \Phi_{12}=-C^TT_2^T-Y_1+Y_2^T$$
$$\Phi_{13}=T_1A-C^TT_3^T+Y_3^T+HS-\varepsilon F_1^TF_2\quad \Phi_{14}=T_1B-C^TT_4^T+Y_4^T-\varepsilon F_1^TF_3$$
$$\Phi_{15}=P-T_1-C^TT_5^T+Y_5^T\quad \Phi_{22}=-(1-\mu)Q_1-Y_2-Y_2^T\quad \Phi_{23}=T_2A-Y_3^T$$
$$\Phi_{24}=T_2B-Y_4^T+HM\quad \Phi_{25}=-T_2-Y_5^T\quad \Phi_{33}=Q_2+T_3A+A^TT_3^T-2S+\tau W+\varepsilon F_2^TF_2$$
$$\Phi_{34}=T_3B+A^TT_4^T+\varepsilon F_2^TF_3\quad \Phi_{35}=L-T_3+A^TT_5^T$$
$$\Phi_{44}=-(1-\mu)Q_2+T_4B+B^TT_4^T-2M+\varepsilon F_3^TF_3\quad \Phi_{45}=-T_4+B^TT_5^T\quad \Phi_{55}=dR-T_5-T_5^T$$

Proof. Replacing C, A and B in

$$\Xi=\begin{bmatrix}\Phi_{11} & \Phi_{12} & \Phi_{13} & \Phi_{14} & \Phi_{15} & -dY_1 & \tau T_1D\\ * & \Phi_{22} & \Phi_{23} & \Phi_{24} & \Phi_{25} & -dY_2 & \tau T_2D\\ * & * & \Phi_{33} & \Phi_{34} & \Phi_{35} & -dY_3 & \tau T_3D\\ * & * & * & \Phi_{44} & \Phi_{45} & -dY_4 & \tau T_4D\\ * & * & * & * & \Phi_{55} & -dY_5 & \tau T_5D\\ * & * & * & * & * & -dR & 0\\ * & * & * & * & * & * & -\tau W\end{bmatrix}<0 \tag{7}$$

$$\Phi_{11}=Q_1+Y_1+Y_1^T-T_1C-C^TT_1^T\quad \Phi_{12}=-C^TT_2^T-Y_1+Y_2^T\quad \Phi_{13}=T_1A-C^TT_3^T+Y_3^T+HS$$
$$\Phi_{14}=T_1B-C^TT_4^T+Y_4^T\quad \Phi_{15}=P-T_1-C^TT_5^T+Y_5^T\quad \Phi_{22}=-(1-\mu)Q_1-Y_2-Y_2^T$$
$$\Phi_{23}=T_2A-Y_3^T\quad \Phi_{24}=T_2B-Y_4^T+HM\quad \Phi_{25}=-T_2-Y_5^T$$
$$\Phi_{33}=Q_2+T_3A+A^TT_3^T-2S+\tau W\quad \Phi_{34}=T_3B+A^TT_4^T\quad \Phi_{35}=L-T_3+A^TT_5^T$$
$$\Phi_{44}=-(1-\mu)Q_2+T_4B+B^TT_4^T-2M\quad \Phi_{45}=-T_4+B^TT_5^T\quad \Phi_{55}=dR-T_5-T_5^T$$

with $C+\Delta C(t)$, $A+\Delta A(t)$ and $B+\Delta B(t)$, $\Delta C(t)$, $\Delta A(t)$ and $\Delta B(t)$ are described, we get

$$\Xi+\begin{bmatrix}-F_1 & 0 & F_2 & F_3 & 0 & 0 & F_4\end{bmatrix}^T H^T(t)\begin{bmatrix}E^TT_1^T & E^TT_2^T & E^TT_3^T & E^TT_4^T & E^TT_5^T & 0 & 0\end{bmatrix}$$
$$+\begin{bmatrix}E^TT_1^T & E^TT_2^T & E^TT_3^T & E^TT_4^T & E^TT_5^T & 0 & 0\end{bmatrix}^T H(t)\begin{bmatrix}-F_1 & 0 & F_2 & F_3 & 0 & 0 & F_4\end{bmatrix}<0 \quad (8)$$

Example

Consider the delayed neural networks (1) with the following parameters:

$$C=\begin{bmatrix}2.3 & 0 & 0\\ 0 & 3.4 & 0\\ 0 & 0 & 2.5\end{bmatrix},\ A=\begin{bmatrix}0.9 & -1.5 & 0.1\\ -1.2 & 1 & 0.2\\ 0.2 & 0.3 & 0.8\end{bmatrix},\ B=\begin{bmatrix}0.8 & 0.6 & 0.2\\ 0.5 & 0.7 & 0.1\\ 0.2 & 0.1 & 0.5\end{bmatrix},\ D=\begin{bmatrix}0.3 & 0.2 & 0.1\\ 0.1 & 0.2 & 0.1\\ 0.1 & 0.1 & 0.2\end{bmatrix},$$

$$h_1=0.2,\ h_2=0.2,\ h_3=0.2,\ h_4=0.2 \quad (9)$$

If we consider the system (2) and the following parameters:

$$E=\begin{bmatrix}1 & 1 & 1\end{bmatrix}',\quad F_1=\begin{bmatrix}0.1 & 0.2 & 0.1\end{bmatrix},\quad F_2=\begin{bmatrix}0.2 & 0 & 0.3\end{bmatrix},\quad F_3=\begin{bmatrix}0.1 & 0.1 & 0.2\end{bmatrix},$$
$$F_4=\begin{bmatrix}0 & 0.1 & 0.1\end{bmatrix} \quad (10)$$

By Theorem 1, when $\mu=0$, the upper bound of time delays to guarantee the global asymptotic stability of uncertain DCNN (1) with (9) and (10) is $h=\tau=22.5474$. Some comparisons are made.

Summary

In this paper, we give the delay-dependent robust stability criteria for cellular neural networks with time-varying discrete and distributed delays. The previously used bounding technique has shown to be unnecessary in deriving the delay-dependent stability results. we discuss the stability of nominal system. Then the robust stability for system with norm-bounded uncertainties is considered. Our results are more general than some previous ones since they can be used even if the time-derivative of time-delay is not less than one. numerical examples are given to demonstrate the effectiveness of the proposed stability criteria.

References

[1] L.O.Chua, L.Yang. Cellular neural networks: theory. *IEEE trans circuits sys* Vol. 35 (1988), p. 1257

[2] T.Roska, T.Boros, P.Thiran, L.O.Chua. Detecting simple motion using cellular neural networks. Proc.IEEE Int. *Workshop on Cellular neural networks and their applications* Vol. 1 (1990), p. 127

[3] Tianping Chen, Libin Rong. Robust global exponential stability of Coben-Grossberg neural networks with time delays. IEEE transactions on neural networks, Vol.15(2004) p. 203

[4] Arik S. Golbal asymptotic stability of a larger class of neural networks with constant time-delays. *Phys Letters A*. Vol. 311 (2003), p. 504

[5] Vimal Singh. A generalized LMI-based approach to the global asymptotic stability of delayed cellular neural networks. *IEEE trancactions on neural networks*, Vol.15(2004) p. 223

[6] V.Singh. Robust stability of cellular neural networks with delay:linear matrix inequality approach. *IEEE Proc.-Control Theory Appl.*, Vol.151(2004) p. 125

doi:10.4028/www.scientific.net/AMM.50-51.919

To Optimize a BP network system

Tong Wei, Qin Liping

Northwestern Polytechnical University, Xi'an, China

aiqinliping@sina.com

Keywords: credit assessment; risk control; BP network

Abstract. The neural network has been introduced into the studies of credit risk assessment. However, the ratio of the dataset for training and testing is difficult to determine, so the neural network is not robust enough to give the judgment. Therefore, using the 2000 instances of personal consumer credit data set for approval of credit applications of a provincial-level China Construction Bank, for the BP neural network model, the study focused on the ratio of the dataset for training and testing. The results show that, when the ratio of the dataset for training and testing is 800:1200, the neural network model 2 for credit risk assessment has better performance. And it can achieve the desired accuracy and computational efficiency, so the BP network system for credit risk assessment is optimized.

Introduction

Credit risk analysis is indispensable for financial risk management, and it has been very concerned in the financial and banking areas. Credit assessment is often used to predict the potential risk of the credit portfolio, so we can use different ways and models to judge the risk of a loan.

In theory, the credit assessment can be divided into two different types[1][2]. The first one is applying assessment. This is to classify the credit category to "good" or "bad". And the data used in the modeling generally include the financial information and personal information of applicants. The second is behavioral assessment, that is, to judge by experiences, through the relevant information, such as payment history information. This differs from the first one, because this should consider the model of payment of a loan. In this paper we will focus on the first one.

Neural network has been used actively in credit risk assessment. Abdou, Pointon and Elmasry (2008) used the credit scoring models to assess the Egyptian banks[3], and through the assessment they did comparative studies on the probabilistic neural network and a multilayer feed forward network, discriminant analysis, probabilistic element analysis, logistic regression methods. And they got the conclusion that the average correct classification rate of the neural network model is higher than that of other methods. But it is worth to emphasize that, in the "yes / no" of a credit approval, the neural network is not enough to make sound judgments, and this will influence the accuracy of the assessment. This actually stems from the ratio of the training and the testing data set used in the trial is not better. Therefore, this paper will focus on solving the problem how to set a better ratio, and optimizing the BP neural network, according to the better ratio.

To construct the data set for the credit rating

For the use of the credit evaluation system given below, we use the data set of personal consumer credit and approval of business, which was provided by a provincial-level China Construction Bank. The data set contains 2,000 cases of loan applications. In the original data, there are 20 attributes, including categorical and numerical ones (They are Status of current deposit, term of the loan (in months), credit history, Purpose of loan, average monthly earnings, holding bonds or not, occupation, Installment in the percentage of disposable income, gender, guarantor, apartment, Status of fixed deposit, age, Other loan, house and Bank can give the forehead). There are a variety of financial and personal information of the applicants. Here, we change the type attributes into numerical attributes in

the data set. The samples are marked as type-1 (good, 1400 cases) and type-2 (bad, 600 cases), and used to test the accuracy of the system.

To design the credit risk assessment system

When designing the neural network model for credit risk assessment, taking other issues into account, we focused on identifying the ratio of the training and the testing data set.With the addition of more hidden layer, the computational cost becomes higher, and the running time becomes longer. So we use the single neuron with a single hidden layer and output layer, in order to simplify the computation process. So we can maintain its simplicity, in the design of BP neural network learning algorithm, and thus greatly reduce the computational cost.

If the ratio of the training and testing data set is too high, the learning will be ineffective [4]. Therefore, we used a personal consumer credit data set (including 2000 real applications - making cases) for approval of credit applications of a provincial-level China Construction Bank to train 3 neural network models, under the conditions of 9 different ratios of training and testing data set (see Table 1). These three neural network models have different types of topology, and the numbers of their hidden layer neurons, their learning rate and momentum rate are not the same. We compared the performances of the neural network, under the conditions of all ratios, to get the ideal network and radio. Under the conditions of 9 different ratios, we adjusted the learning rate and momentum coefficient of the 3 kinds of neural network constantly, in order to converge to the minimum error of 0.008 (see table 1). Table 2 lists the final parameter values of the trained neural network models.

There are 2 stages for the credit assessment system based on neural network, that's the stage of data processing and the stage of neural network identification. The first stage is to standardize each value of the attributes within the data set one by one, and the output of this stage provides the standardized value for the next stage. In the second stage, we use the standardized value provided by the previous stage to assess the attributes of applicants, and use the neural network to decide whether to accept an application. In addition, in order to identify the best ratio of the training and testing data set, we have designed a certain proportion to be used.

Table 1 the final training parameters of the neural network models

Neural network models	*artificial neural network -1*	*artificial neural network -2*	*artificial neural network -3*
The number of input layer nodes	20	20	20
The number of hidden nodes	14	19	23
The number of output layer nodes	1	1	1
Learning coefficient	0.0080	0.0092	0.0073
Momentum rate	0.71	0.70	0.80
Random initial weight range	-0.3 to +0.3	-0.3 to +0.3	-0.3 to +0.3
The required minimum error	0.008	0.008	0.008
The actual error	0.007970	0.007998	0.008530
The maximum allowable number of iterations	25000	25000	25000
The actual number of iterations	19226	18639	25000
The optimal training - testing ratios	1000:1000	800:1200	1200:800

To design the ratios of the dataset for training and testing. To determine which ratio of the dataset for training and testing is batter, we have designed 9 ratios R_i （$R_i, i = 1,\cdots,9$） (see table 3). For example, R_1 means the ratio is 200:1800, that's we use the first 200 credit application samples to train the network, and the other 1,800 samples are used to test the classification of the neural network after the completion of training.

Data processing. This stage is to prepare the data for the next stages of the neural network, which are the stages of training and discrimination. The data provided to the input layer of the neural

network is generally between "0" and "1". In credit assessment, the differences between the values (input value) on behalf of the property of the credit applicant are very small. So when setting a simple standard treatments to the entire data, such as dividing each data by the recorded maximum value, most of the input data will be close to the value of "0". And that can not represent the property, which will lead to the failure of neural network training.

We standardized the input data separately, through the natural logarithm, and map them to the value between "0" and "1". So the neural network can do full study, and present the attributes of the applicants appropriately. And we have to ensure that after the standardization, the input values, representing 20 attributes, are valid to the neural network.

Once the standardization is completed, the standardized value will be used as the input data of the neural network.

Neural network identification. In the stage of neural network identification, we used the supervised BP neural network, because its application is simple and can provide adequate training and testing data sets.

Based on the BP algorithm proposed by Rumelhart, etc., according to the number of the input values of the applicant's property, the input layer of the neural network has 20 neurons ($x_1, x_2, \cdots, x_{20}$), and each input neuron receives a standardized value. Here, the only hidden layer contains h neurons. During the experiment, the number h of hidden neurons was continuously adjusted from 1 to 50, and finally we got the number of hidden neurons, which was closed to the given standard (to ensure the effective training, under the minimum cost of time). We use three kinds of neural networks to do this study, that's: the artificial neural network 1, when $h = 14$; the artificial neural network 2, when $h = 19$; the artificial neural network 3, when $h = 23$.

There is only one neuron y in output layer, which consists of the binary output data, that's, "0" for applications, "1" means to refuse the application (We use a simple threshold conversion method to convert the single output neurons of the neural network, so we can divided the feature space into two categories.). We use the threshold of 0.5 to distinguish the credit categories of "good" or "bad." If the output value of the neural network is greater or equal to 0.5, the sample is classified into Class 1 (good, acceptable); otherwise, it's classified into Class 2 (bad, rejected). Thus we know: The applicant $j \in$ good credit rating，if ：$ANN_{out}(j) \geq 0.5$; The applicant $j \in$ bad credit rating，if ：$ANN_{out}(j) \prec 0.5$

Here, $ANN_{out}(j)$ represents the original output value of the neural network, when introducing the j-th attribute of the samples (applicants). And then we use the threshold transformation, to map it to Class 1 or 2. This generally reflects the output credit decision-making process of the applicant j (see Figure 1).

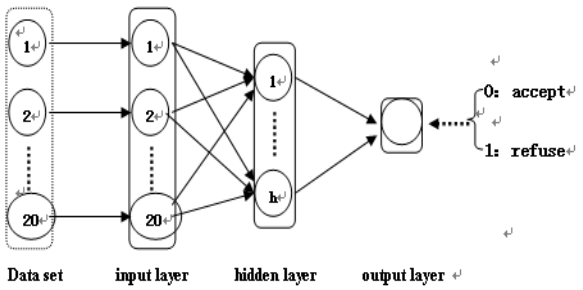


Fig. 1. The general topology of neural network models used here

The experiment and result analysis

Experimental purposes and conditions. The main purpose of this experiment was to identify an ideal ratio of the training and testing data set, which fits a neural network model. So we trained the 3 given neural networks under the conditions of $R_i(R_i, i=1,\cdots,9)$. We can get the experimental results, through the implementation of the algorithms on PC. Experimental conditions Requirement: 2.8 GHz, 2GB of RAM, Windows XP operating system, compiler of Borland C + +.

To set the evaluation criteria. In order to compare the performance of the 3 neural network models, under the conditions of $R_i(R_i, i=1,\cdots,9)$, the performance evaluation criteria must be set in advance, including: (1) limit the number of iterations for training to 25,000. Because it's generally believed that 25,000 iterations are effective, and more iterations will increase the computational cost. (2) Set the error value converges to 0.008. Once reaching these standards, the training of the neural model will come to an end, and we can write down time and the rate of correct evaluations (see table 2). In Table 2, the results of the 3 these three neural models is listed, and the results was got after training under the conditions of $R_i(R_i, i=1,\cdots,9)$.

Table 2. Experimental results.

	R	*model*	*error*	*Tt*	*Ot*	*AT*	*At*
R_4	800:1200	1	0.012780	408.09	5.36	99.03	71.59
		2	0.008000	367.76	10.36	99.59	73.31
		3	0.269983	570.12	26.12	75.60	68.11
R_5	1000:1000	1	0.007962	386.18	12.5	99.35	71.29
		2	0.272016	596.89	19.89	74.1	67.30
		3	0.012150	688.95	31.26	99.02	67.71
R_6	1200:800	1	0.015176	587.79	16.11	98.78	70.60
		2	0.295700	717.78	23.62	72.3	69.30
		3	0.008516	818.38	15.52	99.26	66.11

Note: here, "R" represents the ratio of training and testing data set, "Tt" represents the training time （seconds）; "Ot" represents the operating time （seconds$\times 10^{-5}$）, "AT" represents the accuracy of the training data set （%）, and "At" represents the accuracy of the testing data set（%）.

Comparative Analysis. Based on the results of Table 3, we found that only 3 systems meet or nearly meet our evaluation criteria, namely: System (1) – in the condition of R_4, the application of artificial neural model 2 meet two criteria, which requires at least 18,639 iterations to ensure convergence of the error value of 0.008. System (2) - in the condition of R_5, artificial neural model 1 also satisfies the two criteria, which requires 19,226 iterations. System (3) - in the condition of R_6, artificial neural model 3 can almost meet the two criteria, but it requires 25,000 iterations to converge on the error of 0.008516.

Comparative Analysis of the computational cost. Taking the training time into consideration, system (1) requires about 368 seconds, less than the other two. Taking the running time into account, system (1) requires only 10.36×10^{-5} seconds to judge a credit application, and this is still better than the other. Thus we can see that the combination application of artificial neural model 2 and R_4 shows the highest efficiency of cost calculations.

Comparative Analysis of the assess accuracy. The accuracy of a credit rating system is the most important factor to judge its merit. The overall accuracy of the training set and testing set can not represent the accuracy of a credit rating system, so we used a more scientific method, that is, to compare the accuracies of the training set and testing set in the system separately. This is because the accuracy of the testing set is received, through introducing visible inputs or samples into the neural model, and certainly it reflects the robustness of the training model. The accuracy of the training set is also very important, because the data set of consumer credit application is unstable and difficult to be handled through intelligent systems. Based on these considerations, we only did the precision of these two types of accuracies separately, but not of the overall accuracy. Accordingly to the experimental results, in these 3 systems, when using the artificial neural model 2, under the condition of R_4, we can

get the highest accuracy. And the accuracy of training set is 99.59%, while the accuracy of testing set is 73.31%.

In other words, when using the artificial neural network model 2, trained under the condition of R_4, through the best configuration (the ratio of the training data and the testing data of to be estimated data was adjusted to 40%:60%), the credit risk assessment system can do assessment to the individuals consumer credit risk, successfully and effectively.

Conclusion

In this paper, we studied the application of 3 supervised neural network models for the credit risk assessment, in the conditions of different ratios of training-testing data sets. And experimental results obtained: the neural network model of credit risk assessment 2 is in the best performance in the condition of R_4, and this program uses a training set of 800 samples and 1200 samples for testing set. The accuracy of the training data set and the test data set can respectively reach 99.59% and 73.31%. The training of this neural model can be completed in about 368 seconds, and the decision-making time of the trained neural network model is very short, that is 10.36×10^{-5} second. Thus, the nervous system can be effectively applied to automatic data processing for credit approval. In addition, the results also confirmed that a more optimal ratio exists, and it's just like the one (50%: 50%) used by Sakprasat & Sinclair (2007)[5] or lower. Based on this, the future credit risk assessment could focus on the design, training and implementation of a better multiple-output the nervous system (for example, a system which can show why a credit application is refused).

References

[1] Laha,A.: Building contextual classifiers by integrating fuzzy rule based classification technique and k-nn method for credit scoring.Advanced Engineering Informatics,281--291 (2007).

[2] Vellido,A.,Lisboa, P.J.G., &Vaughan, J.: Neural networks in business: AA survey of applications (1992–1998).Expert Systems with Applications, 51--70 (1999).

[3] Abdou, H., Pointon, J., & Elmasry, A.: Neural nets versus conventional techniques in credit scoring in Egyptian banking. Expert Systems and Applications, 1275--1292 (2008).

[4] Asuncion,A.,&Newman, D.J..UCI machine learning repository [OL]. http:// www.ics.uci.edu/~mlearn/MLRepository.html. (2007).

[5] Sakprasat,S.,&Sinclair, M.C.: Classification rule mining for automatic credit approval using genetic programming. In Proceedings of the IEEE congress on evolutionary computation[C], Singapore, 548--555 (2007).

doi:10.4028/www.scientific.net/AMM.50-51.924

Study on Dispersion Character of Fine Oolitic Hematite Ore Particle

Fengjiu Li

College of Resources and Environmentof , Hebei Polytechnic University, Tangshan, China

Mining Development and Safety Technology Key Lab of Hebei Province, Tangshan, China

tslifengjiu@126.com

Keywords: Oolitic Hematite Ore, Fine Particle, Dispersion

Abstract. The fine oolitic hematite ore (＜20μm) is easily covered by the ore slime, therefore, it is processed very difficultly with traditional crafts, for example, gravity treatment, magnetic separation, and flotation. The tiny iron ore is unable to recycle effectively, bring about a large of useful minerals running off. It is indicated that the selective flocculation is effective separation craft in many research works. The good dispersion of fine particles is the selective flocculation essential condition, the excessive dispersion will destroy the selective flocculation, at the same time it can be influenced by the water quality, pH, the mixing time, the shear rate and the dispersing agent use level. In this article, to oolitic hematite ore, the chemistry dispersion research is conducted to provide the foundation for further selective flocculation separation.

Introduction

Oolitic hematite can been found very easly in our counctry, which is known by people because of the oolitic structure, and its total account for 1/9 iron ore reserves. In the process of geological action, the gangue minerals such as hematite, quartz, and clay are generally irregular-shape. The multi-layer structure is formed, which is from the center of oolitic particles to outward. The boundary between iron and gangue mineral is not very clear, and takes on transition behaviour, at the same time in such iron ore, the disturbution grit of iron and gangue minerals are smaller size, when the iron minerals single degree of dissociation 85%, mineral particles's diameter is usually -20μm , as a result the separation becomes very difficult by the existing technology of mineral processing equipment[1~3].

With the price of iron ore increasing rapidly and the resource of easy mined iron ore droping day after day, the exploitation and utilization of oolitic hematite has been concerned. In the separation process of fine-grained weak magnetic, the selective sorting flocculating has a strong vitality, and its core technology is that the fine mineral particles is fully distributed under ultrafine grinding, and through the selective flocculation, the fine hematite becomes larger diameter particles in order to reach the scope of separation [4~7]. In this article, the chemistry dispersion research is conducted to provide the foundation for further selective flocculation separation of oolitic hematite ore.

1Raw materials and test methods

Raw materials

Test for the raw materials used Xuan Long-oolitic hematite, XMB-Φ240 × 300-rod mill ground to 100% -500 head. The main raw materials, mineral composition and size of the percentage

in Table 1. Chemical dispersant is the six sodium (SHP), sodium tripolyphosphate (STPP) Fenxi Chun is reagents, sodium silicate (SS) is industrial products, the module is 3, the concentration is 40%; pH value regulator: Fenxi Chun NaOH, using Fenxi Chun match HCl concentration of 1mol/L.

Table 1Main mineral composition and size of the percentage

Mineral's name	hematite	ferrite	quartz	carbonate	chlorite	amphibole	achmatite
Size of percentage	60-65	2-3	20-25	2-4	3-5	1-2	1-2

Test Method

Fine particles (W0) of water mixed with the test, matching 2% the concentration of ore pulp 800ml in 1000ml beaker, using high-speed digital IKA RW20 stirring instrument mixing, stirring after the ore pulp, the settlement 1min, with the next level siphonic law lessons 650ml, the remaining Weighing ore pulp drying filter (W). W-W0 defined as the ratio of the sedimentation rate Es, as a measure of fine particles in the water dispersion medium, that is

$$E_S = \frac{W}{W_0} \times 100\% \quad (1)$$

Sedimentation rate of particles under test ES larger particles that are dispersed, the worse the effect, on the contrary, the particles dispersed, the better the effect.

Test results and discussion

The quality of water and pH effects on fine particle suspension dispersion behavior

Placing the mine in distilled water and tap water mixed media, adjusting pH to 2, 4, 6, 7, 9, 11, 12 , mixing speed at 1000r/min, stirring time is 5min, temperature 25 ℃, the time of settlement is1min, E_S settlement of its rate as shown in Figure 1.

It can be seen that the micro-particle in the water, a wide range of pH in the range of 2 to 14 have a strong mutual condensate, in distilled water at pH 9 after the settlement is significantly lower in the vicinity of pH11 minimum efficiency of the settlement, continue to rise High-efficiency settlement pH change, the quality of water and pH has greater impact in fine particle suspension dispersion behavior , the fine particle dispersion effect of distilled water medium is stronger than tap water; distilled water in the medium strong alkaline conditions, the effect of scattered stronger than in neutral and acidic conditions.

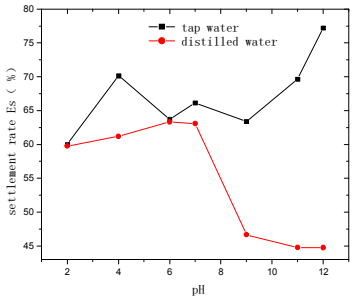


Figure1 The quality of water and pH effects on fine particle suspension dispersion behavior

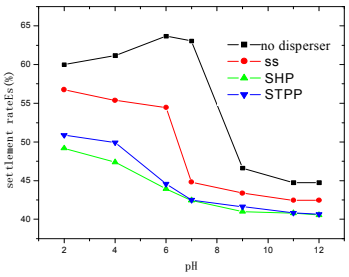


Figure2 Disperser on the fine particle suspension dispersion behavior

Dispersant on the fine particle suspension dispersion behavior

Mixing mine with distilled water, adding dispersant SHP, STPP, SS 5mg/L, then adjust pHto 2, 4, 6, 7, 9, 11, 12 , stirring speed at 1000r/min, stirring time is5min, temperature at 25 ℃, the time of settlement is 1min, the settlement rate of ES shown in Figure 2.

Figure 2 shows three disperser have clear anti-condensate each other, after the pH of distilled water highter than 7 , the settlement rate reduced significantly, when the pH at 11 the settlement efficiency reached the least point , if continued to increase the pH settlement efficiency just had a little change, the result match with no disperser in fine particle suspension with the spread of pH changes in trends in a very good relationship, and the order of ability to spread is SHP> STPP> SS.

Dispersant's use level effects on fine particle suspension dispersion behavior

Mixing mine with distilled water, adding dispersant SHP, at the use level of 0.625mg/L, 1.25mg/L, 2.5mg/L, 5mg/L, 10mg/L, 15mg/L; STPP has the same use level with SHP; SS's use level at 2.5mg/L, 5mg/L, 10mg/L, 20mg/L, 40mg/L, 60mg/L; adjusting ore pulp's pH to 11, stirring speed at 1000r/min, stirring time is 5min, temperature at 25 ℃, the time of settlement is 1min, the result shown as in Figure 3.

It can been seen that, with the disperser gradually increase the use level ,the settlement rate has been gradually reduced, the effect gradually dispersed.If want to fully dispersed micro-fine-grained minerals, SHP's use level is 3mg/L, STPP at 5mg/L, when the use level of SS reaches 20mg/L sedimentation rate decreases no longer, it's effection is slightly weaker than the SHP and STPP.

Stiring speed effects on the fine particle suspension dispersion behavior

Mixing mine with distilled water, adding dispersant SHP, the use level is 5mg/L; STPP's use level is 5mg/L; and SS at the use level of 20mg/L; adjusting ore pulp's pH to 11,setting stirring speed at 600r/min, 800r/min, 1000r/min, 1200r/min, 1400r/min, stirring time is 5min, temperature at25 ℃, the time of settlement is 1min, the settlement rate of ES shown in Figure 4.

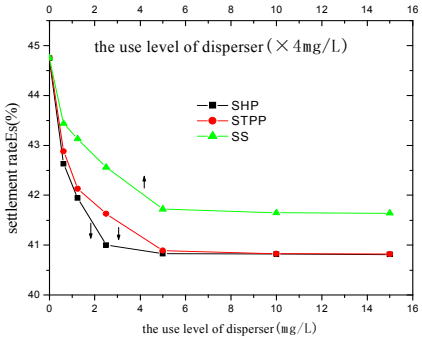


Figure 3 Dispersant's use level effects on fine particle suspension dispersion behavior

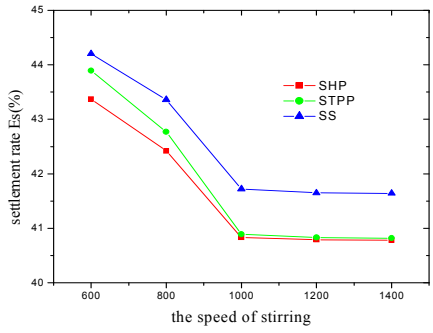


Figure 4 Stirring speed effects on the fine particle suspension dispersion behavior

It is known that from figure 4, with the speed of stirring constantly increase, the settlement rate is lower step by step, but if the speed is above 1000r/min the settlement rate just has a little change, so stirring speed has a great effection on the fine particle suspension dispersion behavior , when the time and the use level of dispersant are certain, the next step is fix the stirring speed for fine particle suspension.

Stirring time effects on fine particle suspension dispersion behavior

Mixing mine with distilled water, adding dispersant SHP, it's use level is 5mg/L ;STPP's use level is 5mg/L; SS's use level is 20mg/L; adjusting ore pulp's pH to 11, stirring speed at 1000r/min, stirring time at1min, 2min, 3min , 4min, 5min, 6min, at temperature of 25℃ , the time of settlement is 1min, the settlement rate of ES shown in Figure 5.

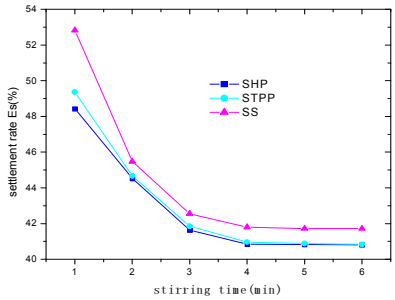


Figure 5 Stirring time effects on the fine particle suspension dispersion behavior

It can be seen that from figure 5, with the icrease of stirring time, the settlement rate has been gradually reduced, when stirring time more than 4min, the settlement rate will remain basically unchanged, so when the use level of dispersant and the speed of stirring are certain, the stirring time has a greater impact on fine particle suspension of decentralization, there is a certain time to let fine particle suspension dispersed to achieve stability.

Scattered Mechanism

From analysising the main mineral composition and the size of percentage content of oolitic hematite, we know that the main composition of mixed ore are hematite and quartz, besides, it contains a small amount of limonite, chlorite, and other easy-mineral mud. According to the DLVO theory, when there is no disperser, the fine particles dispersed suspension stability of the system will depend on the role of the gravitational potential energy of the van der Waals (UA), electrostatic potential energy (Uel) and membrane hydration repulsion potential energy (UHR) and the. UA has always been able to attract, Uel depends on the fine particles of ζ potential value and its symbols, UHR and the degree of surface water are closely related; dispersant exists, it will be extended DLVO theory, the total energy (UT), said: UT = UA + Uel + UHR + Ust + ..., for which the Ust steric energy.

It must add in the alkali and dispersing agent to let ore pulp in dispersed state, the role of the alkali is to enable the particle surface to increases the negative potential, so that between the particles is dispersed by the exclusion of like electricity. Most of dispersants are hydrophilic material, they absorption on the surface of the mine to increase the hydrophilic pills, and promote the the stability of particles. Water glass (SS) can generate rubber in the ore pulp, hydrolysis and dissociat H2SiO3 elements, SiO32-and HSiO3-plasma, these ions, molecules or particles adsorbed on the surface of minerals, have greatly boosted the particle surface Hydrophilic. Sodium tripolyphosphate (STPP) is a multi-charge of a micelle structure of the electrolyte, which charge more than micellar absorption in the mineral grains on the surface to enhance the mineral grains on the surface of the potential negative. Six sodium (SHP) has a unique structure of the molecules, resulting in water ionization (Na4P6O18) 2 - and PO43-, mineral absorption in the grain surface can be enhanced not only tablets ore on the surface of the water, but also have a strong space - Resistance effect.

In the water media, at the broad range of pH (2 to 12), fine particle suspension setting each other strong phenomenon can not be avoided, mainly because of the existence of Ca, Mg ion in the water, Ca, Mg ion would undermine the existence of suspension dispersion's stability. In the high-Ca, Mg ion the existence of strong alkaline medium, fine particle's interaction is not agree the DLVO theory ,at this time, calcium and magnesium on the way of surface sediment absorpt on the surface of minerals, the fine particle precipitation will depend on the surface of the bridge So that

the role of the United Mine occurred tablets non-selective coagulation, and damage the tiny mineral grains in strong alkaline solution of the stability of dispersion. In the distilled water medium in which exist little of Ca, Mg ion, fine particle suspension in pH2 ~ 7 occurred in the scope of each condensate strongly, strong alkaline conditions has improved dispersion effect. This is mainly because quartz's isoelectric point is about pH1.7, and hematite's about pH6.2, the two minerals have contrary electricity on the surface, the role of static electricity to attract each other, the suspension system will inevitably occur each condensate, with the increase of pH value, the surface of the mineral potential of the two symbols the same as mutually exclusive, pH value, the higher the more negative potential ζ, the stronger the role of exclusion. Added in the SHP, STPP, SS three dispersant, the dispersant in the particle surface, on the one hand, an increase of minerals on the surface of the run so that the UHR value increased humidity, On the other hand, three types of dispersant in the mineral surface to produce steric effect Ust, to join the three dispersant can make the role of the potential energy of the total increase, to a very good anti-setting role of each other.

Conclusion

(1) The water quality has a great impact on fine mineral particles dispersed suspension of the stability , the existence of Ca, Mg ion will undermine the stability of dispersion, make the occurrence of non-selective particle coagulation. It is essential to soften of water quality to the selective flocculation of fine-grained hematite.

(2) Scattered three dispersant capability for the strength of the order of SHP> STPP> SS, to dispersed fine oolitic hematite, SHP need to 3mg/L or so, STPP need to 5mg/L or so, SS dispersed the weakest capacity to Dispersion stability of the need to 20mg/L about, three dispersant in strong alkaline medium (pH11 or so) to achieve the best ability to spread.

(3) Stirring speed and mixing time on the fine particle suspension dispersion stability as important, given the conditions, to achieve stability of the fine particle dispersion, mixing speed and mixing time, to determine the existence of a value.

References

[1] Tong Xiong, Li Yinshu.The new technology research of refractory oolitic red iron ore dressing test [J]. *Chinese Engineering Science* 2005 (4): 38 ~ 41.

[2] Zhang Qufei. Kind of selective use of magnetic flocculation process to the United States Tilden iron ore from soil tests conducted research [J]. *China Mining Magazine,* 2001, 10 (3): 56 ~ 58.

[3] Wang Jing, Shang Yanbo. The preliminary study on oolitic hematite's flotation [J]. *Mining and Metallurgical Engineering,* 2004 (6): 38 to 40.

[4] Jiang Linhua, Zhu Shuquan. Cationic polymer flocculant in the application of fine coal in the water [J]. *Coal Science and Technology*, 2008, 36 (5): 97 to 100.

[5] Ji Shoufeng, Li Guichun. Reunion ultra-fine powders Mechanism [J]. *China Mining Magazine* 2006 (15): 54 ~ 56.

[6] Ma Wan-chu. Ultrafine/nano-particles in the water dispersed dielectric behavior of [J]. *Mining and Metallurgical Engineering*, 2003 (15): 43 ~ 46.

[7] Johnson SB, Franks GV, Seales PJ, etal. Surface chemistry of a rheology relationships in concentrated mineral suspensions. *Mineral Proeessing,* 2000, 58 (1): 267 ~ 304.

doi:10.4028/www.scientific.net/AMM.50-51.929

Spallation of Concrete under Dynamic Loading: Mesh Size Effect

L. Yuan[1, a], T. Xu[1, 2, b], Q.Xu[2, c]

[1] Center for Material Failure Modeling Research, Dalian University, Dalian 116622, China

[2]State Key Laboratory of Geo-hazard Prevention and Geo-environment Protection, Chengdu 610059, China

[a]tanzon_04@126.com, [b]neuxutao@126.com, [c] xq@cdut.edu.cn

Key words: concrete; spallation; mesh size effect; numerical simulation; dynamic load

Abstract: Spallation of concrete under dynamic loading has been the hot issue of concern about civil engineering structures and protective engineering. In the present paper, the principle of propagation of stress wave and the induced spallation process along a one-dimensional bar and the RFPA-Dynamics code which considers the heterogeneity of the concrete materials are briefly introduced. In order to numerically investigate the effect of mesh size of numerical model on the computational results, the failure process of concrete with three mesh sizes under dynamic loads was numerically simulated using RFPA-Dynamics code. Numerical simulations show that the failure patterns of concrete with different mesh sizes under dynamic stress waves are different. It is found that for the numerical models with the same mechanical parameters, the smaller mesh size of the numerical model is, the longer propagation of stress wave delay, the smaller the compressive stress is, thus the greater the tensile stress is, and the number of cracks is increasing, in general, rupture is more serious. The accuracy of mesh is little to improve when the number of the grid increases to a certain value.

Introduction

Dynamic failure of materials has been the hot issue of concern about aerospace, civil engineering structures, defense projects, and so on [1, 2]. As a building material widely used in human activities, spallation of the concrete wall on the inner side of the protective structure poses threat to the personnel and equipment inside the structure even though the structure and the wall itself do not suffer general failure. Extensive studies on the mechanical behavior of concrete for protection against short duration dynamic loading effects have been made in the last few decades [3]. Through progress has been achieved, the analysis of concrete structures subjected to dynamic loading is still a very complex issue. In recent years, with the development of computer technology, numerical methods have been the most popular way of studying the problem of dynamic failure of brittle materials such as concrete. Some techniques or numerical tools such as LS-DYNA have been used to study the mechanical response of concrete structures. As we know, the finite element method is mesh-dependent, different results are caused by using different mesh sizes which is important to the analysis of the modeling results. Thus, in order to investigate the effect of mesh size on the propagation of stress wave and the spallation of concrete under dynamic loading, the propagation of stress wave along the concrete rod and the spallation of concrete rod with three sized meshes under dynamic loading are numerically simulated and the failure patterns of concrete rods with different mesh size were compared in this paper.

Theory of Propagation of One-Dimensional Stress Wave in Rods

The stress wave is transmitted from left to right, when the incident compressive stress wave is transmitted into the left, and reflects in the free surface as the unloading wave, it interacts with the end of the unloading waves. When the resultant stress is only slightly higher than the fracture strength of the material, at this time spallation is caused. When the first spallation occurs, new free surface are being created by the flaking off of the spalls. It follows that the shape and amplitude of stress wave has a significant impact on spallation. The maximum transient tensile stress fracture criterion [4] is first proposed in the form of Eq.1:

$$\sigma \geq \sigma_c . \tag{1}$$

According to the geometrical relationships [4], Eq.2 demonstrates the thickness of the first spallation and Eq.3 defines the reflection time for the start time:

$$\delta_1 = \frac{\lambda}{2} \cdot \frac{\sigma_c}{\sigma_m} \tag{2}$$

$$t_1 = \frac{\delta}{C_0} = \frac{\lambda}{2C_0} \cdot \frac{\sigma_c}{\sigma_m} . \tag{3}$$

where σ_c is spalling strength of the materials. When the maximum tensile stress σ is higher than spalling strength σ_c of material, it will occur to spalls. According to the waveform, amplitude and different spall strength of the materials, fracture of the specimen may be caused many times. Based on theory of one-dimensional stress wave in rods, we assume that σ_c is the tensile strength of the material, λ is a wavelength, σ_m is amplitude, C_0 is wave velocity for the triangular stress wave in the free end of a long straight bar, δ_1 denotes a distance from the free face when spalling occurs.

Brief Outline of RFPA-Dynamics

In this paper, the spallation of concrete is numerically simulated by the RFPA-Dynamics Code for mesh effect [5, 6]. The RFPA-Dynamics Code, which has the same basic idea with the basic version of RFPA except for the dynamic finite element algorithm, which are using to simulate. Spallation process of stress waves in inhomogeneous media under the impact loads. In this paper, Weibull's distribution is mainly due to taking into account the heterogeneity of concrete, and a homogeneity index is introduced to represent the heterogeneity of the various units. The stress and the displacement is calculated by finite element method and used maximum tensile stress (or tensile strain) and Mohr-Coulomb criterion as the threshold condition .Describing to damage evolution of the units is based on constitutive relation to elastic damage [5, 6]. In this paper, mesh effect on Numerical Simulation of the spalling of concrete is studied by using RFPA2D-Dynamic program under the impact loads.

Mesh Size Effect

Model Set-up. To validate the numerical model, a homogeneous and isotropic bar sample subjected to a triangle pressure incident wave was used for the simulations, The one-dimensional bar model with 5mm in width and 100mm in length is studied in 2D plane stress problem, as shown in Fig.1(a) and the applied stress waveform acted on the bar is shown in Fig.1(b). The compressive stress is imposed at the left end, acting on a triangular pulse and the right end of it is kept free. According to the young's modulus and density can be calculated as one-dimensional wave velocity 4000m / s,

and therefore wavelength λ = 24mm. No strain rate effect is considered for the mechanical parameters such as Young's modulus and tensile strength of concrete in this paper. The parameters and calculation conditions are listed in Table 1.

(a) (b)

Fig. 1 (a)Numerical model of the propagation of stress wave along a one-dimensional bar (b)Applied incident waveform acted on the concrete model.

Table 1 Material Properties for the model

Young's Modulus		Uniaxial compressive strength		Poison's ration		Density [kg/mm^3]
Homogeneity Index	Mean [MPa]	Homogeneity Index	Mean [MPa]	Homogeneity Index	Mean [MPa]	
200	40000	200	200	100	0.25	2.5e-6

For the sake of better study of mesh effect on numerical simulation of the spalling of concrete, in this paper, we assume that the other mechanical parameters are consistent, and the shape of grid is divided into square, while grid sizes are: 5×100, 10×200, 15×300. Different specific models are shown in Fig.2.

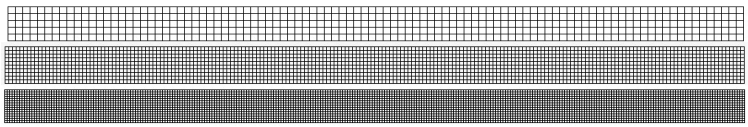
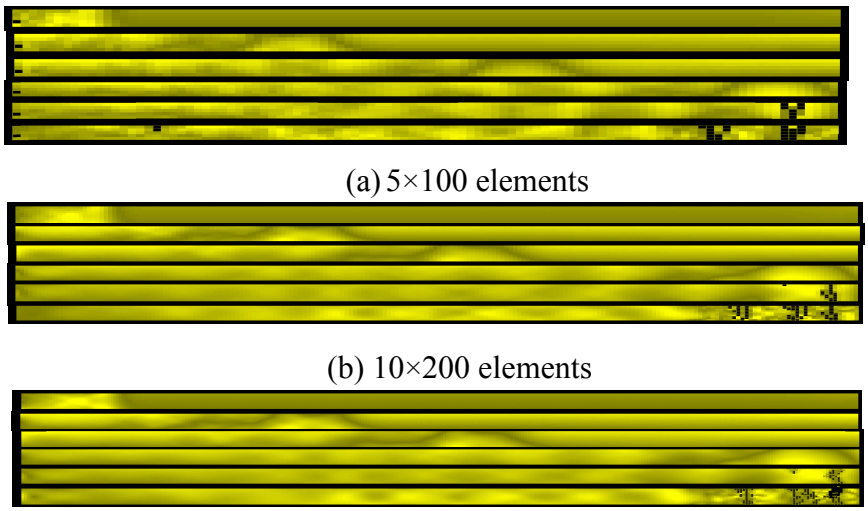


Fig. 2 Numerical model with different mesh sizes (a) 5×100, (b) 10×200 and (c) 15×300.

Modeling Results and Analysis. The following diagrams show propagation of the one-dimensional stress wave in the rods and cracking process with different mesh sizes, where the level of gray are defined as the value of the shear stress. In Fig.3 (a), the individual units have been damaged because the grid of the specimen 1 is large, concrete is correspondingly homogeneous, when the incident wave propagates from left to right. When t = 25μs, the stress wave reaches the end of the specimen, and then the compression stress waves as a tensile wave is reflected with others. When t = 27.3μs, the sample occurs in the first spallation. With the time duration, all of the compression stress waves are reflected as tensile stress waves, which cause a new round of spallations, but the fracture parts are concentrated and not caused the diffuse peeling rupture, which deviate from laboratory experiments. Fig.3 (b) shows that as the incident wave propagate, the individual units aren't damaged like the specimen 1. When t = 25μs, compression stress wave is reflecting in the free face. The first spallation appears at a distance of 4.5mm from the right end of the free surface at t = 27.4μs. With the time duration, the compressive stress waves are reflected as tensile stress waves, and the new spallation are caused, where many cracks appear. It can be seen from Fig.3(c) that the mesh is relatively smaller, so concrete is more heterogeneity. Due to this, the stress waves reflect and refract clearly in the figures. With the action of stress wave, When t = 25μs, compression stress wave is reflecting in the free face. The first spallation appears at a distance of 4mm from the right

end of the free surface at t = 27.4μs. With the time duration, the visible cracks are formed around the multiple spalling. As mesh is smaller, Fracture serious clearly than those of the specimen 1, 2 at the right end, and because of large enough to crack gaps, block of cracks will fly, which is what we often say it “scab”. As can be seen from the above his time of the simulation results is relatively consistent results with the value of theory. As shown in Fig.4, different grid sizes of the models bring different results. Correspondingly, spalling thickness, time and stress strength is calculated with different grid sizes, which is given in Table2.



(a) 5×100 elements

(b) 10×200 elements

(c) 15×300 elements

Fig. 3 Failure processes of concrete with different mesh sizes at t=5, 10, 15, 25, 27.5, 30 us

Table 2 Numerical parameters for the model

No.	Size [mm×mm]	Mesh [mm×mm]	Spallaion thickness [mm]	Spalling time[us]	Tensile stress with spalling[MPa]
1	5×100	5×100	7	27.3	10.594
2	5×100	10×200	4.5	27.4	10.205
3	5×100	15×300	4	27.4	10.181

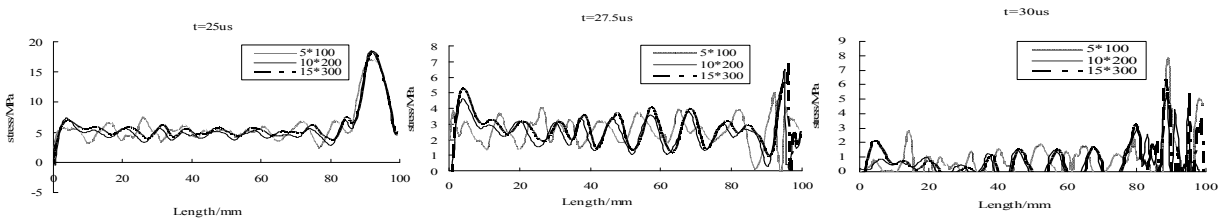


Fig. 4 Compressive stress wave propagation at different time 25us, 27.5us, 30us

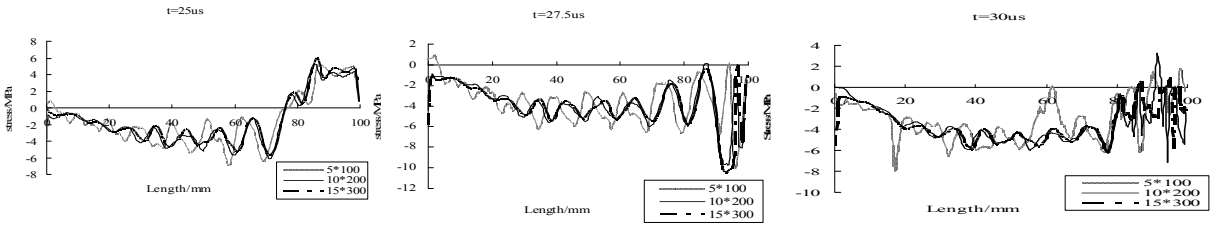


Fig. 5 Tensile stress wave propagation at different time 25us, 27.5us, 30us

It can be seen from Fig.4 and Fig.5 that the smaller grid is divided into, the longer propagation of stress wave delay, the smaller the compressive stress is, the greater the value of tensile stress is, and

the number of cracks is increasing, in general, rupture is more serious for concrete with same parameters. Fig.6 shows clearly that the tensile stress rupture with the smaller mesh, so the value of the tensile stress is smaller, eventually approaching a stable value. With the non-linear analysis, when the compression wave reflect as a tensile stress, which exceeds the tensile strength of concrete, and induce fractures and cracks, if mesh size is large, the value stress are averaged at the crack units, unit stress value is smaller than that of the actual stress fields; but if mesh size is smaller, the unit stress value is close to the actual stress. Therefore, large mesh size is hard to damage; on the contrary, when using smaller grid size, cracking develop better, correspondingly capacity of internal energy release is weaker, which leads to stress smaller value than the actual structure. It follows that RFPA-Dynamic is effective in capturing stress wave propagation and analyzing mesh size effect.

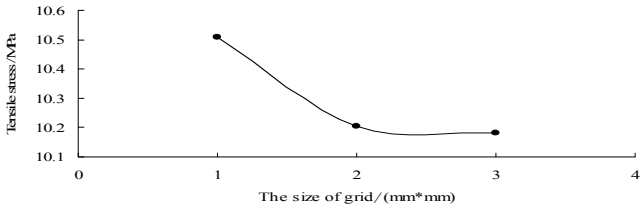


Fig. 6 Relationship between mesh size and tensile stress with spalling

Summary

In this paper, we used a numerical simulator with the advantage of direct simulation of fracture and fragmentation, which considers the heterogeneity of material properties, to investigate mesh size effect on spallation of concrete. Firstly, the numerical model was validated by a concrete bar sample with a free end subjected to a triangle pressure incident wave, and then the grids were divided into 3 different sizes. Secondly, the spallation of concrete subjected to dynamic loads was simulated to examine the propagation of incident compressive stress waves and study the failure pattern and failure mechanism of spallation of concrete. Numerical simulations capture the failure pattern of concrete with different mesh sizes, which is obviously different from the failure pattern of concrete. The approach suggested in this paper also can be used generally for mesh effect on spallation of concrete under dynamic loading. The outcomes showed the grid size is selected, to a large extent, which affected the results. In general, the smaller the mesh is divided into, the higher accuracy is. However, when the number of the grid increases to a value, the accuracy of mesh is little to improve, which lead to wasting the computing time and running space.

References

[1] M.L. Qi, X.Y. Qin, L. Zhang, H.H.Liang and S.L.Yan: Journal of Wuhan University of Technology Vol. 29 (2007), p.72.

[2] M.A. Meyers and C.T. Aimone: Dynamic fracture (spalling) of metals. Progress in Material Science Vol. 28 (1983), p.1.

[3] A. Brara, F. Camborde, J.R. Klepaczko and C. Mariotti: Mechanics of Materials Vol. 33 (2001), p. 33.

[4] L.L. Wang: *Foundation of stress Waves* (National Defense Industry press, Beijing: 2005).

[5] T. Xu, Q. Xu, Y.J. Zuo, L. Yuan and P. Liu, in: *The 5th International Symposium on In-situ Rock Stress*, edited by Furen Xie, CRC Press/Balkema (2010).

[6] W.C. Zhu, S.M. Shang, Z.H. Li, C.H.Wei and L.Zhu: Journal of Building Materials Vol. 11 (2008), p. 710.

doi:10.4028/www.scientific.net/AMM.50-51.934

Method of Concept-Drifting Feature Extracting in Data Streams based on Granular Computing

Chunhua Ju[1, 2, a], Zhaoqian Shuai[1, b]

[1]School of Computer & Information Engineering, Zhejiang Gongshang University, Hangzhou, China

[2]Center for Studies of Modern Business Zhejiang Gongshang University, Hangzhou, China

[a] jch@mail.zjgsu.edu.cn, [b] zhaoqianshuai@gmail.com

Keywords: Granular Computing; FCA; Data Streams; Concept-Drifting; Feature Extracting

Abstract. Business data streams are dynamic and easy to drift, extract concept-drifting feature is one important work of data streams mining. This paper describes the characteristics and the concept drift of data streams, and constructs the formal concept description model of streaming data based on granular computing firstly. Then, the paper proposes the concept lattice pairs' based concept relaxation-matching coincidence degree algorithm; the feature extraction method is also described. Finally, experiment and analysis are presented in order to explain and evaluate the method.

Introduction

With the rapid development of the computer, network and communication technology, many applications have high-speed and dynamic data. These continuous, rapid and over time data is called data streams. The dynamic characteristic of the data streams is usually the target concept, that is, the interest concept often occur a fundamental change. Changes in the hidden context can induce more or less radical changes in the target concept, producing what is generally known as *concept drift* [9]. Concept drifting, that is, the phenomenon which the goal classification model changes with such changes in data characteristic over time [1].

Granular computing (GrC), as a new research field for the study of data mining, is a conceptual framework for many issues [7]. Yao [2] explained the data, knowledge representation and processing based on GrC, discussed the basic tasks of data mining is to find the right level of granularity for the data and knowledge representation, the method applicable to the static data modeling, but does not apply to the dynamic data streams. Formal concept analysis (FCA) is a method based on lattice theory, which deals with the relationship between objects and attributes of formal context in a certain field [4]. So, we can constructs the description of formal concept model of streaming data based on concept lattice, and then extract their features.

Data Streams Modeling based on Granular Computing

A framework is proposed for rule mining based on formal and mathematical modeling in [3]. Within this framework, a concept is defined as a pair consisting of the intension, an expression in a certain language over the set of attributes, and the extension, a subset of the universe, of the concept. Intensions are defined by formulas of a language and extensions are defined by subsets of objects in the universe. GrC based concept analysis mainly has the following steps: (1)concept

hierarchy, establishing relations between concepts, (2)describe the intension and extension of concept, (3)analyze the coincidence on extent, (4)identify and evaluate drifting features.

Data streams can be described by attributes of data streams. Data streams can be represented by multi-dimensional set of attributes, which include time, transaction set, a limited non-empty set of attributes, formal description rules and so on.

Definition 1: Data Streams. DS = {TP, U, A, { | a}, { | a }, {| a },}. TP is a time period, and could be set one second, one minute or one hour, etc. It implies a time granule A time period TP is defined by a time section, written with. Where is the begin time of the time period, and is the end time of the time period. U is a finite nonempty set of objects. A is a finite nonempty set of attributes. is the extension of the concept. is the intension of the concept. is the mapping from object to attribute, called membership function. is corresponding the concept of connotation to the average membership of each attribute , which reflects the degree of each attribute in the concept. shows each object to average level of the average membership of each attribute in the concept of connotation , which reflects divergent degree of this concept.

Definition 2: Fuzzy Formal Context. A fuzzy formal context is expressed as K= (U, A, I), which U is set of objects, A is set of attributes and mapping I is membership function, and I: U×A→[0, 1].

Definition 3: Window. For every attributes in fuzzy formal context, set two thresholds θ_d and ψ_d, which are respectively the lower bound and upper bound of window and meet $0\theta_d < \psi_d 1$.

Definition 4: Mapping f and g. In the K =(U, A, I), OP(U)， DP(A) (P is the symbol of power set), mapping f and g can been defined between P(U) and P(A). f (O) = { d| oO， θ_dI (o, d) ψ_d } ;g (D) = { o| dD， θ_dI (o, d)ψ_d }.

Definition 5: Fuzzy Parameter σ. For object set O P(U) and attribute set DP(A), D = f (O), oO, d D, | O| and | D| are respectively the cardinality of O and D. If | O| ≠0 and | D| ≠0, then (1) , (2)
Σ is a symbol of fuzzy set. When you need to specify the specific object set and attribute set, (O, D), and σ are respectively let (O, D) and σ (O, D).

Definition 6: Fuzzy Parameter λ. For object set OP(U) and attribute set DP(A), D = f (O), oO, dD, | O| and | D| are respectively the cardinality of O and D. If | O| ≠0 and | D| ≠0, then= (3), = (4)

Definition 7: Fuzzy Concept. If OP(U) and DP(A) satisfy O = g (D) an d D = f (O), then C =(O, D, σ, λ) is called a fuzzy concept of fuzzy formal context K. O is the extension of C, and D is the intension of C, which are respectively calculated based on (2) and (4).

Drift Feature Extracting Based on GrC

A. Concept Coincidence

This section will combine the intension and extension of fuzzy concept, and the relationship of concepts, uses the method of formal concept analysis to analyze the indicators and characteristics of data streams and then extract drifting features. The relationships between concepts are studied based on their corresponding extensions [3].A concept is defined as a binary pair consisting of the intension and the extension, which can be expressed in the form, . The former element is the sets of satisfiable transactions that are the extension of the concept, the later element is a description of, and that is the intension of the concept.

As the time goes on, DS is continuously updating, the elements of a specific time period of the DS are also changing: If the transaction data has been added (deleted), the extension of a concept may be expanded or hold (reduced or hold). If the adding and deleting happened in the same time period, the extension of a concept may be expanded, reduced or hold. In a word, the extension of a concept has been changed. If the concept is at , it will change to at . The current time period is , while is the history time period.

Definition 8: Coincidence on Extent. The coincidence on extent of c1 and c2 is defined as =, and are the extensions of concept, is cardinality on its extension.

Definition 9: Coincidence on Intent. The coincidence on intent of c1 and c2 is defined as = , here, is the σ deviation on a property in their respective concept of intersection of intension sets, and are the intensions.

Definition 10: Coincidence on Concept (CoC). CoC, is respectively the level of the concept CL_1 and CL_2, and b is modification in consideration of the influence of concept depth.

In general, a concept of data streams,, will drift as time goes on. The concept drifting in data streams is actually decided by the changes upon the extension of the concept, that is, the meaning set of the concept is drifting [10]. In concept lattices, the concepts of deeper level have more intensions but fewer extensions [6]. By observing the changes of object and property sets in the concept lattices, we can detect drifting.

Definition 11: Pride, concept lattice pairs' based relaxation-matching [8] coincidence degree.

The coincidence algorithm and the similar concept query algorithm are shown as follows:

Pride. Input : CL_1 and CL_2,concept stack S, similar concepts set *Set*1, dissimilar concepts set *Set*2, similar concepts sub-lattices set *Set*3.; Output : relaxation-matching coincidence degree(pride), *Set*1, *Set*2, *Set*3; Initialize : *Lcur = 1 , Set1 = Set2 = Set3 = , Ccur = null , C* = null , CLcur = null*

Step 1. **if** Lcur has any concept not been queried, C is the concept has not been queried **then** Ccur = C(U, A ,σ,λ) **else** *Lcur = Lcur + 1*

if *Lcur > Lmax* **then goto** *Step 4*

Step 2. *C* = FindSimilarConcept (Ccur , CL_1 , CL_2)*

if *C* = null* **then** *Set2 = Set2Ccur* , **goto** *Step 3* **else** *Set1 = Set1{ (Ccur , C*) } , CLcur = CLcur{ (Ccur , C*) } , C* = null*

if *child(Ccur) = null* **then goto** *Step 3* **else** *push(Ccur , S) , Ccur = child(Ccur)* , **repeat** *Step 2*

Step 3. **while** *(top(S)!= null)*

if *has child never been queried* **then** *Ccur = child(top(S))* , **goto** *Step 2* **else pop**(*S)*

if *CLcur != null* **then** *Set3 = Set3CLcur , CLcur = null*

goto *Step 1*

Step 4. Set n as the cardinality of similar concepts sub-lattices set Set3, every similar concepts sub-lattice contains more than one concept node; set m as the sum of all concepts in the similar concepts sub-lattice. Concept lattice pairs' based concept relaxation-matching coincidence degree, that is, Pride(CL1, CL2) = .

FindSimilarConcept(Ccur , CL_1 , CL_2). Input : concept lattices CL1 and CL2, current concept C; Output : similar concepts

Step 1. set CL1 as the basis concept lattice, compute coincidences between concept C and all concepts in the |A| layer of CL2, taking one of the largest degree of the concept of coupling, and denote it by C '.

Step 2. Compute coincidences between C and the child or father concept of c', taking the largest degree of the concept of coupling, and denote it by C*.

Step 3. If the coincidence on C and C* satisfies coincidence(C, C*) , which is the threshold of concept coincidence, then, C* is similar to C and query successful. Otherwise, query failed.

B. Drift Feature Extracting

Data streams change as time goes on constantly, concept lattices are also updating continually. Analyze two periods of data streams information sets by concept lattices, and call the algorithm. We can get the newest similar concept set Set1, dissimilar concept set Set2, similar concept sub-lattices set Set3 and relaxation-matching coincidence degree. And then the key similar concept lattices can be fixed, which are the feature attributes set. The key similar concept lattices is related to Set1 and Set3, and it is concept set that elements have the higher concept coincidence degree in more than one key similar concept lattices. CL_1 is the concept lattice of DS at, while CL_2 is the concept lattice at. By analyzing the "Pride" to detect the possibility of the concept drift occurs, then get the key similar concepts sub-lattices and dissimilar concepts set in continuous time series using the algorithm of "Pride". After that, the drifting features can be detected by analyzing the evolution of key similar concepts sub-lattices and the extension changes of every concept in the dissimilar concepts set. The process is shown in Figure 1.

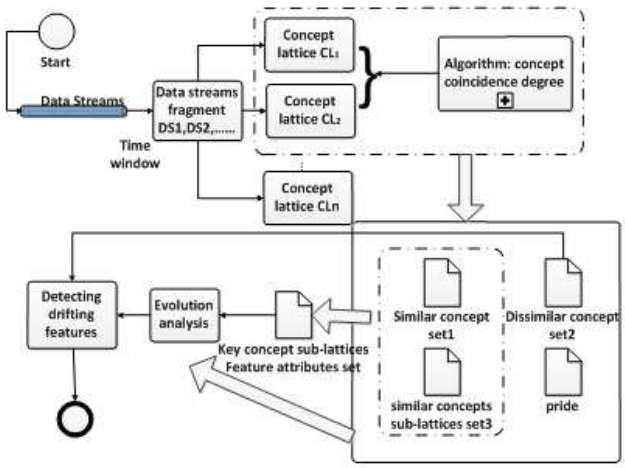


Figure1. Extract drifting feature.

Actually the concept drift of DS is determined by the change of concept extension, that is to say, the meaning collection of concept drift [10]. In GrC, we can say the range of the object the concept can depict changesyooo. When the extension of the concept extends (shrinks), the extension of the concept becomes stronger (weaker), this is to say, the coincidence of the extension increases. Both concepts extend and shrink can cause concept drift. Based on the concept collection of extension change from the analysis of differing concept collection set2, these extension changes are the key factor to concept drift. Extension change can be described as the numbers of attributes the object includes change, or, the distribution of attributes in the object changes.

Experiment and Analysis

Now we'll analyze concept drifting feature extracting through the following example. Table 1 and figure 2 respectively show formal context and concept context CL_1 of one time quantum table 2 and figure 3 respectively show the formal context and concept context CL_2 of the following time quantum table 3 shows the mapping relationship between the two lattices.

(1) Table 1 Formal context 1

U	a	b	c	d	e
1	0	1	0	1	0
2	1	0	1	0	1
3	1	1	0	0	1
4	0	1	1	1	0
5	1	0	0	0	1

(2) Table2 Formal context 2

U	a	b	c	d	e
1	0	1	0	1	0
2	1	0	1	0	1
3	1	1	0	0	1
4	0	1	1	0	0
5	1	0	0	0	1

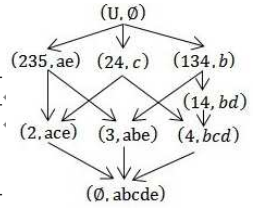


Figure 2. CL_1.

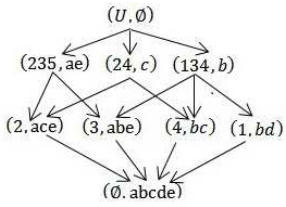


Figure 3. CL_2.

Table 3 Concept hierarchy and mapping

Hierarchy	Concept Lattice	Mapping	Concept Lattice	Mapping
1	Mapping in CL_1	A→	Mapping in CL_2	A'→
2		B→, C→,D→		B'→, C'→,D'→
3		E→		F'→, '→, H'→,E'→
4		F→, G→,H→		I'→
5		I→		

Table 4 Result of coincidence analysis

Output	Value
Set1	{{A,A'},{B,B'},{C,C'},{D,D'}, {F,F'},{G,G'},{H,H'},{I,I'}}
Set2	{E,E'}
Set3	{{A,A'},{B,B'},{C,C'},{D,D'}, {F,F'},{G,G'},{H,H'},{I,I'}}
Pride	0.65

Suppose the threshold of CoC is 0.6, we can get set1, set2, set3, and "Pride" listed in the following table through the pride algorithm to CL_1 and CL_2. From the table we can know that in current time, concept lattice pairs CL_1 and CL_2, the key concept sub-lattice contains {A, B, C, D, F, G, H, I}. It's the concept E:and F: that cause the change, of which the coincidence of E and E' are low, being the characteristic concept causing concept drift. The result of coincidence analysis is shown in table 4.

Conclusions

In this article, we have put forward the modeling way of implied concept drift of data streams and feature extraction, and also analyzed the changes of intension, extension and concepts based on coincidence. This method shows that GrC is a formal way for mining data streams. Besides, we found that there are several aspects to further study as follows: fast algorithms for building concept lattice; fast and effective method for searching similar concept; considering the characteristics of specific business data streams in modeling and algorithm designing.

Acknowledgment

This article is supported by the National Natural Science Foundation of China (71071141), Zhejiang Provincial Natural Science Foundation of China (Z1091224) and the Xinmiao Project of Zhejiang Province.

References

[1] J. Han, M. Kamber. Data Mining: concepts and technique, 2nd edition. High Education Press, Morgan Kaufman Publishers, 2006.

[2] Y.Y.Yao, Granular computing for data mining, Proceedings of SPIE Conference on Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security, Dasarathy, B.V. (2006), p. 1-12.

[3] Y.Y.Yao, On Modeling data mining with granular computing, Proceedings of the 25th Annual International Computer Software and Applications Conference (2001), p. 638-643.

[4] J.Q.Xu, X.Peng, W.Y.Zhao. Program Clustering for Comprehension based on Fuzzy Formal Concept Analysis [J]. Journal of Computer Research and Development, 2009，46 (9): 1556-1566.

[5] Y.Y.Yao, Potential applications of granular computing in knowledge discovery and data mining, in: Proceedings of World Multi Conference on Systemics, Cybernetics and Informations, 1999, 573–580.

[6] R. BelohlaVek. Similarity relations in concept lattices. J. Logic Comput, 2000, 10(6):823–845.

[7] G.Y.Wang, Q.H.Zhang and J.Hu. An overview of granular computing [J] . CAAI Transactions on Intelligent Systems, 2007, 2(6) ： 6-26.

[8] Z.J.Wang, D.C.Zhan. A component retrieval method based on feature matching. International Journal of Information Technology, 2006, 2 (8): 60-72.

[9] G. Widmer, M. Kubat. Learning in the Presence of Concept Drift and Hidden Contexts. Machine Learning, 1996, 23(1): 69-101.

[10] Xiao-Dong Zhu, Zhi-Qiu Huang. Conceptual modeling rules extracting for data streams. Knowledge-Based Systems, 2008, 21:934:940.

©2014 Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.939

A Survey on Nurses' Job Burnout and Factors Analysis in Tangshan Private Hospitals

Yu Su[1], Zhuo Wang[1], Yadong Bian[2], Yanbo Peng[1]

1 Hebei United University, Tangshan of Hebei Province

2 Tangshan Mental&Psychological Hospital, Tangshan of Hebei Province

Email: imsuyu1979@163.com

Keywords: Private Hospitals; Nurse; Job Burnout; Emotional Exhaustion; Depersonalization; Personal Accomplishment

Abstract With transformation of medical model and development of medical and health service, nurses often appear the phenomenon of job burnout. Sampling questionnaire of nurses' formal backgrounds and MBI-SS, we conduct a survey on 280 re-post professional nurses who have been working for one or more years to know the present situation of job burnout and to analyze relative factors, thus providing a theoretical basis for solving the problem of nurses' job burnout in Tangshan private hospitals. The result shows that the job burnout phenomenon is widespread; the job burnout level of male nurses is lower; PD' nurses are easier to feel burnout and ED' nurses have low personal accomplishment; the more satisfaction to higher salary, the lower of their job burnout. All these indicate that, it does help to relieve nurses' sense of burnout, if we arrange their job volume appropriately, improve their treatment.

1 Introduction

Job Burnout refers to the undivided shows a body and mind excessively consumption and exhaustion syndrome under the work pressure for a long time.

With the changes of medical mode and development of medical health service, nursing faces more opportunity and challenges, nurses also bear an increasing pressure from society and work. Under the pressure of society, medical and their career, the majority of nurses feel fatigue in body and mind which severely influence their families and life, which attribute expand of mobility of nursing team and decline of quality of nursing.

In 1970s, American clinical psychologist *Fredenberger* and social psychologist *Maslach* have a profound study and raise the item "Job Burnout", which, as proper noun is listed into the science reference[1]. They designed Maslach Job Burnout questionnaire (MBI) which is the classic test instrument for job burnout field.

In 2001, a questionnaire came from the hospital of America, Canada, German, England and Scotland five countries. Its result shows that hospital's job burnout phenomenon is very serious[2]. Domestic scholar Lixiaomei found nurses' highly emotional exhaustion account for 59.1%. highly depersonalization 34.5%, highly personal accomplishment 53.2% when he study the nursing job burnout[3]. Zhumadian Center Hospital in Henan province has a questionnaire survey to 325 nurses. The result shows that nursing job burnout account for 48.7%[4]. Lixiaowen had a survey to 330 clinical nurses in general hospital and found the nursing job burnout with a high incidence[5].

This article has a survey to nurses in Tangshan private hospitals and analyze their job burnout

condition and relative factors, so as to find the effective strategies and methods to decline nurses' job burnout.

2 Subjects and Methods

2.1 Subjects

The subjects of this research are 280 or the job professional nurses with one or more than one year working experience among 6 private hospitals in Tangshan.

2.2 Methods

Use the multistage sampling method: firstly, select 6 private hospitals by means of sampling method at random; secondly, based on the classification of Internal Department(ID), Surgical Department(SD), Gynecology and Obstetrical Department(GOD), Paediatrics Department(PD), Emergency Department(ED)，ICU and other medical-assisted departments, a certain amount of nurses are elected at random from the above departments as the subjects.

The nurses are surveyed in form of voluntary questionnaires. The questionnaire -consist of 2 parts: questionnaire of nurses' formal backgrounds and MBI-Human Service survey(MBI-SS). MSI-SS was inverted by American psychologists *Maslach* and *Jachson* in 1986[6,7]. It consists of 22 clues and 3 dimensions:①Emotional Exhaustion(EE) ②Depersonalization ③Personal Accomplishment(PA). Every clause has 7 subclauses. The subclauses are scored from 0 to 6. The sum of the clauses in each dimension is the score of each dimension. Each dimension of professional laziness is divided into 3 stages for assessing[2].If the score of EE is below 19, then the degree is light; 19-26, modest; over 26, high. If the score of Depersonalization is below 6, the degree is light; 6-9, modest; over 7, high. If the score of PA is over 39, light; 34-39, modest, below 34, high. Before carrying out formal questionnaire, so subject are elected for initial research, with no limit of credits and the Cronbach's α: 0.738.

2.3 Statistical Analysis

All the materials are analyzed by establishing database in Excel and applying to SPSS13.0 software. Apply *t*-test and variance analysis to compare the state of nurses' job burnout; and apply multiple-liner regression analysis to analyze the factors that affect that state. $P<0.05$ explain that the disparity has statistic significance.

3 Results

This research has distributed 280 questionnaires, retrieving 280 in the rate of 100%. Except for 13 ineffective questionnaires, there are still 267 effective. The effective rate can reach 95.36%.

3.1 The general condition of researching object

Among the 267 effective questionnaires, the youngest researching object is 18 years old and the oldest is 54. The female is in the percentage of 87.6%.The number of nurses in ID is the largest, the number of nurses in ICU is the smallest.

3.2 Present situation of nurses' job burnout

To analyze nurse career, we found that: the nurses have moderate than 83.5% more EE; 83.5% have moderate above to Depersonalization. 95.5% have moderate above to PA declined. Result hint, nurses' job burnout in the private hospitals is generally exist, and the burnout degree presenting in the three dimensions is not optimistic. (chart 1, sheet 1)

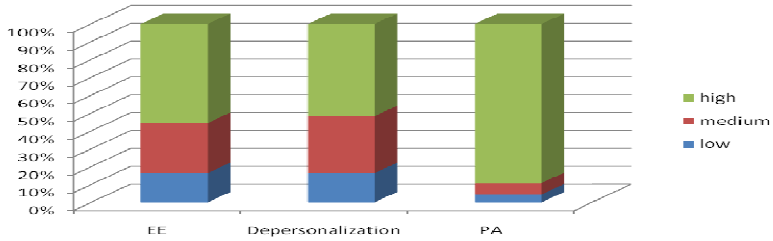


Chart 1 The condition of the burnout degree in the three dimensions

Sheet1 The condition of the job burnout's total score and dimensions

items	*n*	$\bar{x}$±s
EE	267	28.57±10.60
Depersonalization	267	11.24±7.32
PA	267	21.51±9.50
The total score	267	61.32±22.78

3.3 Analysis on influential factors of job burnout

3.3.1 Single factor Analysis

3.3.1.1 The relations between demographic characteristics of subjects and job burnout

The results show that its influence factor is gender in terms of the total score of burnout and EE. The level of fatigue of male nurse is lower than that of female nurses. As for PA, gender is its influence factor. PA of male nurse is lower than that of female nurse. However, age has no effect on both the total score of burnout and every aspect of dimension. The result is showing in the sheet 2.

Sheet 2 Relationship between gender and job burnout（ $\bar{x}$±s）

gender	numbers	The total score	EE	Depersonalization	PA
male	33	49.91±20.35	23.70±9.02	10.00±5.88	16.21±9.25
female	234	62.93±22.68	29.25±10.64	11.42±7.49	22.26±9.32
t		-3.12	2.86	1.04	3.49
P		0.00	0.01	0.30	0.00

3.3.1.2 The Relation between different department and job burnout

Through the analysis, we find in different department the disparity between the total scores of nurses' burnout and the scores of every other dimension has statistic significance. In the aspect of the total scores of burnout, EE and Depersonalization, PD' nurses get the highest scores, while in the aspect of PA, ED' nurses get the lowest scores. See results in sheet 3.

Sheet 3 Relationship between department and job burnout（ $\bar{x}$±s）

department	numbers	The total score	EE	Depersonalization	PA
ID	72	69.92±22.20	34.07± 8.47	11.35±7.12	24.50±10.25
SD	48	59.27±24.87	25.29±12.90	12.52±10.84	21.46±7.40
GOD	20	60.95±22.32	25.85±10.50	11.45±6.92	23.65±7.31
PD	19	80.21±21.50	34.79±9.34	16.00±6.59	19.42±7.68
ED	24	50.96±9.54	25.92±6.84	9.08±2.93	15.96±5.11
ICU	17	69.65±7.79	34.59±7.40	10.00±4.65	25.06±8.18
others	67	49.90±20.70	23.46±8.90	9.90±5.65	16.54±9.29
F		9.50	11.14	2.48	10.09
P		0.00	0.00	0.02	0.00

3.3.1.3 The Relation between the patients' numbers and job burnout

From the results in sheet 4, there is the statistic significance between patients' numbers and their total scores of burnout. More patients, the scores in PA will be the lower, while burnout is the higher; on the contrary, fewer patients, burnout is the lower.

Sheet 4 Relationship between the patients' numbers and job burnout（ $\bar{x}$±s）

patients	numbers	The total score	EE	Depersonalization	PA
＜20	81	64.73±19.87	28.23±9.73	12.32±9.20	24.17±7.78
20～	94	63.09±24.94	29.88±11.06	11.44±7.04	21.77±9.73
40～	67	58.00±22.27	28.32±10.80	10.48±5.48	19.20±10.44
60～	25	53.17±23.13	26.17±10.83	9.33±5.30	17.67±9.27
F		2.31	0.93	1.39	4.93
P		0.08	0.43	0.25	0.00

3.3.1.4 Relationship between satisfaction degree of income and job burnout

We can get something from sheet 5: firstly, as for the total score, EE and Depersonalization, the score about very dissatisfaction to income is obviously higher than other groups , and degree of languor ranks first. Secondly, as for PA, the score about very satisfaction to income is the highest, while degree of languor is the lowest.

Sheet 5 Relationship between satisfaction degree of income and job burnout（ $\bar{x}$±s）

degree	numbers	The total score	EE	Depersonalization	PA
Very satisfied	16	45.31±25.17*	18.92±10.66*	8.00±5.69*	28.38±12.12*
Relative satisfied	96	55.17±20.35	25.88±9.75	9.57±5.54	19.72±9.54
dissatisfied	97	64.01±20.99	29.18±10.06	12.27±6.49	22.57±8.28
Very dissatisfied	58	70.69±23.96	34.29±9.58	13.00±10.21	23.40±10.27
F		9.07	13.14	4.47	2.83
P		0.00	0.00	0.00	0.04

Annotations：*indicates that compared to other three groups, the difference is of statistics importance.

3.3.2 Multivariate analysis

Sheet 6 The multiple linear regression analysis of nurses' job burnout

dimensions	variables	$\hat{\beta}$	SE	$\hat{\beta}_{标化}$	t	*P*
The total score	gender	11.87	5.05	0.18	2.35	0.02
	patients	-0.19	0.08	-0.18	2.53	0.01
	satisfaction degree of income	10.96	3.64	0.24	3.01	0.00
EE	gender	5.69	2.29	0.19	2.49	0.01
	satisfaction degree of income	4.79	1.66	0.24	2.88	0.00
Depersonalization	satisfaction degree of income	1.22	0.58	0.17	2.10	0.04
PA	departments	-0.63	0.22	-0.19	2.83	0.01
	patients	-0.09	0.03	-0.20	3.14	0.00

According to the total score of burnout of MBI and three scores as the dependent variable, and using the single-factor which in the analysis of statistically significant as independent variables, we can do the multiple linear regression analysis.Sheet 6 shows, gender, the number of patients and the satisfaction about income which play a major role on burnout. Factors affecting the level of EE followed by gender and income satisfaction. Depersonalization and income satisfaction level have close relationship. PA is mainly caused by departments and the number of patients.

4 Conclusion

Nursing is one of professions of the highest pressure in the health care industry. The clinical psychological pressure the nurses bear has become an occupational risk, which is prone to extremely physical and mental fatigue and affects the efficiency and job satisfaction.

This study found that job burnout of nurses in private hospitals in Tangshan is prevalent. The total score of job burnout of nurses in those hospitals is significantly higher than the findings of Lichaoping's[8]. The result also showed that pressure nurses are facing in the hospitals are quite high and also confirm that Nursing are a High- pressure job[9].

This research by the single factor analysis and multivariate analysis, studies respectively the impact of the age, gender, the number of patients and income satisfaction factors on job burnout . The results shows that the degree of job burnout of men's is lower than women's; PD' nurses are easier to be tired; ED' nurses have low PA; the more patients, more significant burnout; the higher satisfaction for nurses with income, the lower degree of job burnout .

Therefore, in terms of individuals, nurses should know some concern knowledge of job burnout, study the skills of interpersonal communication. In terms of administrators in the hospitals , the attention should be paid to care, understanding care and support them with human, material and financial resources and appropriate method should be taken such as the establishment of psychological counseling institutions for nurses to provide psychological counseling and advisory services, communicative skills training[10], optimization and adjustment of personnel, improvement of the working environment, development of a rational income distribution system, etc. These all help to reduce the degree of job burnout of nurses.

References

[1] Pines AM.Bumout.In: Goldberger L,Breznitz S.Handbook of Stress The oretical and Clinical spects.2nd ed.New York: The Free Press, 386-402 (1993).

[2] Wang Xiaochun, Gan Yiqun. Review of Foreign Studies on What Causes Job Burnout. Advances in Psychological Science 11(5),567-572 (2003).

[3] Li Xia mei，Liu Yan jun，Job Stressors and Burnout among Staff Nurses，Chinese Journal of Nursing 11(11),646-648 (2000).

[4] Kong Li. the Causes and Countermeasure of Nurses' Job Burnout. Chinese Journal of Rural Medicine and Pharmacy 14(10),65-66 (2007).

[5] Li Xiaowen,Gu Jinjun,Xu Guangyi,etc.Investigation on the Relationship Between Nurse's Job Burnout and Job Stress, Self-esteem, Locus of Control in Polyclinics. Journal of Nursing Science 22(14),1-3 (2007).

[6] Lves Erickson,J．Keeping in touch with staffperceptions of the professional practice environment．Caring Headlines,2000,20：2.

[7] Aiken,L，Havens,D.，Sloane,D．The Magnet Nursing Services recognition Program：A comparison of two groups of magnet hospitals．American Journal of Nursing 100(3),26-36 (2000).

[8] Hua Xiaofen.The Study on the professional pressure and the shortage crisis of nurses in America [J]. Nursing and Rehabiltation Journal 3(2),133-135(2003).

[9] Xi Bo,Li Jun,Wang Cuili. Correlation Between Nurses' Job Burnout and Personality, Self-esteem and Social Support[J]. Journal of Nursing Science 2(10),3-5 (2006).

[10] Yue Shujin，Cheng Jijuan. Influence the factor and the counterplan that the nurse applies the treatment communication technique. Journal of Nurses Training 18(10),891-892 (2003).

doi:10.4028/www.scientific.net/AMM.50-51.944

Modeling Research on Hospitalization Cost in Patients with Cerebral Infarction Based on BP Neural Network.

Lei Zhou, Jianhui Wu, Guoli Wang, Yu Su, Guobin Zhang

United University of Hebei Province,Tangshan, China

Email：ncmcyjsb@163.com

Key Words: BP neural network, Hospitalization cost, Training set, Test set, Neurons

Abstract. To establish the model of hospitalization cost in patients with cerebral infarction based on neural network using artificial network mathematical model, and setting the appropriate parameters. LM, BR, OSS, SCG four algorithms were used respectively to model 3, 5,10,15, 20 neuron number in different hidden layers and comparisons between model fitting ability and its generalization ability were made. The model which involves a hidden layer, 8 input neurons, 15 neurons in output layer network is the optimal model by using OSS system.

Introduction

With the gradual deepening and continuous improvement of socialist market economic system, health system reform has made some progress. With it, various contradictions and problems emerged inevitably, of which the rapid growth of medical costs has aroused extensive attention. In recent years, the growth of medical costs is far higher than the GDP growth rate, so controlling medical costs has concerned people[1].Identification and its function description to influencing factors to medical costs in inpatients play great role in medical security system' operation.

In the previous research to the cost of hospitalization, multiple linear regression as a statistical method was used mostly[2、3]. However, the influencing factor model in hospitalization cost is a typical multi-factor model with complicated action between the factors. In addition, the output factors themselves may also interfere with the efficiency of the existing statistical model fitting, such as confounding factors, multicollinearity and so on, while the existing linear mode couldn't make perfect correction measures to such inference, and couldn't handle the information with interference[4]. Application of neural networks to deal with the problem of nonlinear regression model could overcome the weaknesses of regression model. The property of model includes the distribution of information storage and parallel co-processing, and highly fault-tolerance and learning ability, and fitness to realistic models.

BP neural network is a neural network learning algorithm, and the whole back propagation based artificial neural network, is one of the most widely used neural network models[5]. with minimum mean square error criterion , It carries out weight training using the nonlinear differentiable function. BP neural network learning process consists of two parts: the information being transmitted and the error back propagation. In the process of forward transmission, the input information goes from the input layer to output layer through hidden layer's treatment, and the state of neurons in each layer only influences that of neurons of next layer. If you couldn't get the desired output in the output layer, then transferred to the error back propagation. The error signals are connected back to the original pathway of neurons and neuronal connection weights of each layer

are modified one by one in their return process. This iterative process continues until the error is within the scope permitted.

BP neural network does not need accurate mathematical model and doesn't have any assumption requirements on the data , and its ability in handling nonlinear questions is higher than that of traditional statistics[6], BP neural network has been applied in extensive fields. In the medical field, it plays an important role in disease diagnosis, pathophysiology and disease research, medical imaging research, drug research, and risk factors' analysis of diseases and the prognosis[7、8].

The author hopes to construct BP neural network model using MATLAB software, studies optimization parameters of model in the process of building BP neural network model in hospitalization cost , sets appropriate parameters to construct the fitting model in hospitalization cost based on BP neural network, compiles sensitivity analysis program based on MATLAB, and measures various factors' impact to hospitalization cost by using sensitivity analysis program on the basis of built BP neutral network model.

Constructing process of BP model

Data normalization

Input and output data sets were normalized using the standardized method so that all normalized variables reached to maximum value of 1 and minimum value of 0.

$$s_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

Data partitioning

The improvement of strategies to network generalization ability required different data sets. For example, early stop strategy (LM algorithm) required the data sets be divided into training set, validation set and test set; Bayesian rules (BR algorithm), variable gradient algorithm (SCG algorithm) and Newton algorithm (OSS algorithm) required the data sets be divided simply into training set, and test set.

Owing to continuous debugging process of constructing BP neutral network model, there is no clear theoretical study showing what kind of parameters of methods should be used to the actual data.

Related parameters initializing

(1) Selection of transfer function:

As the sample data were normalized to [0，1], logarithmic S-shaped functions, namely *log sig* functions both in the hidden layer and output layer were selected.

(2) General function approximation problem could be treated with the network in a hidden layer. so this study set a hidden layer

(3)Initial weights and thresholds:

Newff function was used to construct and initialize BP neural network. If the value which MATLAS gave the network default initial weights and thresholds may not be in[-1，1],.each initial weights and thresholds were multiplied by 0.01 artificially so that the weights and thresholds could be set to [-1,1] between initial weights and thresholds

(4)Initializing the learning speed: setting the learning rate of 0.01

(5)Setting the network convergence of error performance for the SSE, and setting its target value of 0.05.

(6)Maximum training step: Setting the maximum training step as 1000. When the network training steps got to the size of 1000, the training of the training set stopped.

Modeling Results

2459 cases of patients with cerebral infarction (mainly the hospitalization cost and related information) were collected in a third-class A hospital in Tangshan City in 2007 and 2008. All information was re-encoded by computer (See Table 1) and treated with Excel. Cases with missing value and illogical cases were removed and effective samples were 2218, 90.2% of total cases.

Table 1 Influencing factors and quantification methods of hospitalization costs in patients with cerebral infarction

Factors	Code	quantification methods
sex	x1	man=1，woman =2
Marital status	x2	married=1，others=2
age	x3	years old
The first hospitalization	x4	Yes=1，No =2
Rescue or not	x5	No=0，Yes=1
Payment method	x6	Health insurance=1，no health insurance=2
length of stay	x7	days
treatment outcome	x8	cured=1，improved=2，unhealed=3，death=4

In this study, LM, BR, OSS, SCG four algorithms were used respectively to test 3,5,10,15,20 neuron number in the hidden layer with the method of Network random testing for 100 times. The network with better generalization ability and fitting ability were selected according to test results.

Comparison of results of different training methods

(1) Network with 3 neurons in the hidden layer

Table 2 Comparison of results of four algorithms training (3 neurons in the hidden layer)

	Test set R^2				Training set R^2			
	LM	BR	OSS	SCG	LM	BR	OSS	SCG
$\bar{x}$	0.687	0.609	0.699	0.644	0.718	0.752	0.717	0.670
S	0.014	0.002	0.005	0.073	0.020	0.001	0.003	0.089
F	121.69				54.81			
P	0.0000				0.0000			

(2) Network with 5 neurons in the hidden layer

Table3 Comparison of results of four algorithms training (5 neurons in the hidden layer)

	Test set R^2				Training set R^2			
	LM	BR	OSS	SCG	LM	BR	OSS	SCG
$\bar{x}$	0.685	0.590	0.699	0.656	0.722	0.761	0.717	0.689
S	0.009	0.005	0.004	0.038	0.009	0.004	0.003	0.039
F	612.18				217.63			
P	0.0000				0.0000			

(3) Network with 10 neurons in the hidden layer

Table 4 Comparison of results of four algorithms training (10 neurons in the hidden layer)

	Test set R^2				Training set R^2			
	LM	BR	OSS	SCG	LM	BR	OSS	SCG
$\bar{x}$	0.681	0.606	0.699	0.637	0.721	0.778	0.717	0.674
S	0.120	0.018	0.004	0.033	0.010	0.002	0.002	0.028
F	459.90				804.61			
P	0.0000				0.0000			

(4) Network with 15 neurons in the hidden layer

Table 5 Comparison of results of four algorithms training (15 neurons in the hidden layer)

	Test set R^2				Training set R^2			
	LM	BR	OSS	SCG	LM	BR	OSS	SCG
$\bar{x}$	0.673	0.613	0.700	0.627	0.713	0.784	0.718	0.670
S	0.059	0.016	0.003	0.029	0.067	0.002	0.002	0.026
F	141.86				167.61			
P	0.0000				0.0000			

(5)Network with 20 neurons in the hidden layer

Table 6 Comparison of results of four algorithms training (20 neurons in the hidden layer)

	Test set R^2				Training set R^2			
	LM	BR	OSS	SCG	LM	BR	OSS	SCG
$\bar{x}$	0.670	0.616	0.699	0.636	0.710	0.783	0.718	0.681
S	0.055	0.013	0.005	0.035	0.052	0.002	0.003	0.027
F	123.31				216.81			
P	0.0000				0.0000			

These results showed that after training networks with 3,5,10,15,20 neurons in the hidden layer with LM, BR, OSS, SCG four algorithms, the differences between test set $R^2$and training set R^2 were statistically significant. Depending on test set $R^2$and training set R^2sorting of different neurons in the hidden layer, the study drew a conclusion that OSS algorithm was better than other algorithm in network generalization ability and fitting ability.

Selection of neuron number in the hidden layer

Table 7 Comparison of results of 5 different network training aiming at neurons in the hidden layer

	Test set R^2					Training set R^2				
	3	5	10	15	20	3	5	10	15	20
$\bar{x}$	0.699	0.699	0.699	0.700	0.699	0.717	0.717	0.717	0.718	0.718
S	0.005	0.004	0.004	0.003	0.005	0.003	0.003	0.002	0.002	0.003
F	0.87					2.44				
P	0.4833					0.0457				

When hidden layer neurons were 3,5,10,15 and 20, the differences between network test set R^2were not statistically significant, while there was significant difference in the network training set R^2, which showed that the number of these five groups were not identical(the hidden layer with 3, 5, 10 neurons no difference; the hidden layer with 15,20 neurons no difference; but the others had significant difference) In model evaluation, generalization ability of the model should be considered. From the above table, we could see that when the hidden layer neurons were 15,20, the network training set R^2increased, and network fitting ability improved. When the hidden layer neurons were 3,5,10,15 and 20, there was no difference in network generalization ability . In addition, considering network fitting ability, the study selected 15 neurons in the hidden layer

BP neural network modeling results

A network with better generalization ability was chosen after tried different initial weights and threshold values and different hidden layer neurons. BP neural network model in hospitalization cost of patients in cerebral infarction and its influencing factors was proposed. The parameters of model indicators were as follows:

Table 8 BP neural network model parameter table about hospitalization cost in patients with cerebral infarction

Network structure parameter	Network training parameter	Test set simulation results	Training set fitting results
Hidden layer：1 Hidden layer neurons：15 Input layer neurons：8	Training algorithm：*OSS* Total number of iterations to stop training：28 Learning speed：0.01	R=0.86504 R^2=0.74829 R^2_{adj} = 0.51195	R=0.84747 R^2=0.7182 R^2_{adj} = 0.5387
Output layer neurons：1	Performance function：*SSE* Stop training: *SSE** =1.2857	*SSE* =4.9944e+009 *MSE*=1.7046e+007 *RMSE*=4128.6	*SSE* =2.0119e+010 *MSE* =1.7036e+007 *RMSE* =4127.5

* for normalized data [0,1]

Conclusion

The study fitted hospitalization cost in patients with cerebral infarction and its influencing factors. And in this study, only 8 factors were analyzed, of which 6 were qualitative variables, 2 continuous variables. Only using limited clinical characteristics of patients and their social characteristics to predict changes in the cost of hospitalization will not guarantee the accuracy of prediction. In this predicting process, the model could provide a theoretical reference, but the cost couldn't be accurately predicted.

References:

[1]. Cai Le, Wang Chonghua and Wang Jun. Hospitalization Costs' Structure and Its Influencing Factors [J]. Journal of Hospital Statistics 6 (4),208 (1999)

[2] Zhao Xin-song, Wang Shu-xia, Liu Qian, et al. Analysis of the factors affecting hospitalization cost of patient w ith appendiciti. China Hospital Statistics 15(4),322-324 (2008)

[3] Cun Xiao-le, Femg Xiang-xian, Li Shao-xia. Analysis of hospitalization expenditures among 3287 gastric carcinoma inpatients in general hospital. Chinese Journal of Public Health 2010(6).

[4]. Gao Wei, Nie Shaofa, Shi Lvyuan. Application Study on Survival Analysis with Neural Network l [J]. China Health Statistics 23 (4),76-78 (2006)

[5] Wang Junjie, Chen Jingwu, BP neural network theory and its applications in medical statistics in design skills. China Health Statistics 25(5),547-549 (2008)

[6]. Jin Pihuan, Cao Suhua, Deng Wei. Study on BP Neural Network Construction and Optimization and its Application in Medical Statistics. Fudan University Ph.D. Thesis, 2002

[7] Xu Jian-wei. Application of Bayesian-regularization BP neural network in the prediction of hospital bed. J Shanxi Med Univ 39(9),833-835 (2008)

[8] Guo Li-na, Xiang Bing-ren. A Diagnosis Model of Poisoning Based on BP Neural Network. Progress in Pharmaceutical Sciences 32(2),82-85 (2008)

doi:10.4028/www.scientific.net/AMM.50-51.949

Investigation and Analysis about the Scientific Research Ability of Clinical Teachers Guiding Nursing Undergraduates

Shenglian Dong, Jia Zhao, Qingwen Wu

Nursing Department of Hebei United University, Tangshan,China

Email: dshl5335@163.com

Key Words: nursing research; teaching hospitals; clinical teaching

Abstract. In order to understand the scientific research ability of clinical teachers guiding nursing undergraduates in teaching hospitals, 221 clinical nursing teachers in 3 grade Ⅲ hospitals were investigated. The questionnaires included general information, research participation, paper writing and nursing research knowledge. The *results show that* the scientific research awareness of teachers in these hospitals is deficient. 51.1% of them participated in scientific research, 6.8% took charge of projects, and 51.6% published papers. Teachers with better educational background took part in more research projects and published more papers than the others, and the difference is significant ($P<0.01$). Those with higher professional titles or from affiliated hospitals published more papers than those with lower titles or from non-affiliated hospitals, and the difference is significant ($P<0.01$); however, there is no significant difference in research participation between them.

Introduction

Nursing research is not only the needs to promote the development of nursing, but also the value of it, and it plays a vital role in the whole development of nursing [1]. For our nation's present situation, the nursing undergraduates have gradually become the backbone of nursing profession and their research capacity and level influence the development of nursing largely [2]. At present, scientific research education is being widely carried out among nursing undergraduate students. Generally, the scientific research training is arranged in a particular hospital department in the later period of the undergraduate internship, according to their chosen subject and interest. What's more, it calls for direct guidance from the clinical nurses in the teaching hospital. The scientific research ability of these teachers exerts immediately influences on the effects of scientific research training [1]. In order to understand the research ability of clinical teachers in teaching hospitals, we conduct the following investigation, so as to promote practice teaching and clinical teaching ability.

Object and Method

Object

Clinical teaching nurses in 3 third-grade class-A hospitals were chosen. One hospital is affiliated to our college and the other two are the non-affiliated. 240 questionnaires were handed out and 221 of them can be counted as valid. The recovery rate is 92.1%.

Method

The questionnaires were designed by us and the content includes general information, cognitive situation of scientific research, and participation in nursing research, paper writing and publishing. These questionnaires are completed with factual information.

Statistical Methods

Data were analyzed with SPSS13.0 software. Descriptive analysis method was used to analyze these data and X^2 method was used to verify these data.

Result

General Information

The interviewees' age ranged from 22 to 50, and the average age was 35.4±6.9. Educational background: technical secondary school :16 （11.8%）junior college 97（43.9%）; undergraduate 107（48.4%）; master 1（0.5%）.Professional titles: Nurse 33(14.9%）,Nurse Practitioner 85(38.5%) ,Supervisor nurse 91(41.2%) ,Associate chief nurse 11 (5.0%) ,Professor of Nursing 1 (0.5%) . Departments: medical department 72, surgical department 61, obstetrics and gynecology department 32, pediatrics department 25, emergency room 14, ICU 9, others 8.

Research Participation and Paper Writing of Clinical Teaching Nurses

There are 113 people (51.1%) who took part in scientific research, including 15 project leaders (6.8%) and others are participants. 55 of them participated in research for the purpose of promotion, while the 57 left do it for improving their scientific research ability. 114 people (51.6%) published papers in professional periodicals and there are 327 articles published in all, which on average are 1.5±2.2 papers per person. However, 45 people (20.4%) written some papers but were not published. What is more, there are 72(32.6%) people did not write any paper. Among the 114 people, 32(28.1%) published one paper, and 72 (63.2%) have 2 to 5 publications, the left 5 people (8.8%) published more than 5 papers. The situation of research participation and paper publishing were compared and analyzed in terms of educational background, professional titles and different hospitals. And the result was verified with X2 method. In terms of educational background, we divided the interviewees into 2 groups, junior college or below and or above, and found out that people with better educational background participated more scientific researches and got more papers published. The figures are statistically significant (P<0.01). As for professional titles, we divided interviewees into groups of Nurse Practitioner or below and Nurse-in-charge or above. And we found out that people with higher professional titles got more papers published. The figures are statistically significant. However, there is no significant difference in the percentage of scientific research participation.(P>0.05). Within the 3 hospitals we surveyed, 2 non-affiliated teaching hospitals can be counted as one group, while the affiliated teaching hospital can be seen as one group. We compared clinical teaching personnel from the two groups and found that teaching personnel from affiliated teaching hospital got more paper published than non-affiliated teaching hospital. The result was statistically significant; However, there is no significant difference in the percentage of scientific research participation, P>0.05 (see chart one)

Table 1 Comparison of Research Participation and Paper Publishing Rate between Clinical Teachers of Different Attributes

project		Number of People	Research Participation				Paper writing				
			Participants	Non-participants	X2	P	Published, not written	not published but written ,		X2	P
Educational Background .Professional title	Specialist or below	113	46	67	10.054	<0.01	53	13	47	9.490	<0.01
	Undergraduate or above	108	67	41			61	22	25		
	Nurse Practitioner or lowers	118	54	64	2.920	>0.05	44	23	51	20.965	<0.01
	Professor of Nursing or above	103	59	44			70	12	21		
Hospital	Directly-affiliated Hospitals	107	50	57	1.609	>0.05	76	8	23	32.180	<0.01
	Non-directly-affiliated Hospitals	114	63	51			38	27	49		

The Clinical Teaching Teachers' Cognize Perceive and Need for the Scientific Research of Nursing

The survey result shows that 205(92.8%) of the clinical teaching teachers hold the opinion that there is difficulty in scientific research of nursing. Table 2 shows the result of the grasp of the process and the need for the knowledge of the scientific research.

Table 2 the Clinical Teaching Teachers' Cognize Perceive and Need for the Scientific Research of Nursing（n=221）

Project		Number of people	Percentage（%）
the Scientific Research Procedure	grasp	52	23.5
	know	116	52.5
	no idea	53	24.0
needs to the knowledge of research	selected subjects	68	30.8
	Document Indexing	35	15.8
	paper's writing	11	5.0
	Statistical analysis	9	4.1
help to the scientific research	the support of the leaders	160	72.4
	fund subsidization	183	82.8
	cooperation with colleagues	152	68.8
	expert guidance	184	83.3

Discussion

The Clinical Teaching Teachers' Research Awareness Needs to Be Improved.

In the survey, 51.1% of nurses involved the research; only 6.8% are the host of project host, and the rest are participants. Li Shanshan and etc who are nurse of grade Ⅲ hospitals have involved in nursing research activities [4], which shows that the awareness of nursing research is still indifferent, and they do not realize that nursing research pay great role of improving their quality and social status, and the important role of promoting the development the social benefits [5]. Therefore, the training of nurses' research awareness, especially clinical teaching teachers, should be further strengthened in order to enable them to declared subjects and participate in scientific research actively.

The Clinical Teaching Teachers' Research Capability Needs to Be Enhanced Contrarily.

For the survey, 205 nurses (92.8%) considered that it was difficult to engage in nursing research, and only 23.5% of the nurses have mastered the research process. The demand of scientific knowledge is mainly on the knowledge of choosing research topics and searching document. Doing scientific research need the help of expert guidance, financial support, leadership support, colleagues' corporations, which consistent with other documents' report. [10] The promotion of nurses' research quality is the foundation of advancement of their research ability [9]. Nursing directors should, in light of different levels of nursing personnel and various needs of scientific research knowledge and implement layered training, hold discussions on the general problems of nursing personnel and then carry out targeted tutorship concerning specified difficulties. It's all necessary to hold lectures about choosing research subjects, document retrieval, statistics knowledge and paper writing. What's more, they should hire proficient teachers or relevant personnel to provide overall process and comprehensive help. [10]

Pay attention to the Management of Clinical Teaching Teachers and Enhance the Overall Competence of Teaching Nurses.

Because of our nation's low level of nursing education, clinical nursing research is still lagging behind. [12-15]Survey results show that highly educated and high-grade nurses carry out scientific research obviously better than the low educated and low-grade nurses, which consistent with other documents' report. [12-15]For the survey, within our 221 interviewees, only 48.9% of all nurses possess bachelor degree or above and more than 103 own intermediate professional titles, accounting for 46.6%. The majority only own tertiary education. People with bachelor degrees or above participated more actively in scientific research and published more papers than those with lower education degrees. People with higher professional titles or from affiliated hospitals published more papers than those with lower professional titles or from non-affiliated hospitals but there are no significant differences with regard to scientific research participation. At present the educational background of clinical teaching nurses from teaching hospitals and the overall research ability are relatively poor, especially the research guiding ability of teaching nurses can not adapt to the needs of undergraduates, and the situation may do harm to the cultivation of undergraduates' research ability. Hence, teaching hospitals should step up efforts to cultivate the overall competence of teaching personnel, select the nurses with better educational background and rich clinical experiences to assume the work of teaching and guiding undergraduate students. What's more, teaching hospitals should also strengthen the further education of nursing personnel, actively absorb more nurses with higher academic levels and promote the education levels of nursing team. With the continuous advancement of undergraduate nurses, more nurses with master degree joining the teaching team, frequent training programs targeting teaching nurses and the accumulation of teaching nurses' clinical experience, teaching nurses will finally make significant progress in research guiding ability.

References

[1]Huang Xianwei,Analysis and Thinking about the Nursing Undergraduate' Status of Nursing Research in Our Hospital .Medical Information No2,352-355 (2010)

[2] LI Xinhui,YAN Li,CHEN Lili. Improving Baccalaureate Nursing Students'Research Ability by Using Student Research Program and Undergraduate-master Collaboration.Journal of Nursing Science Aug. 24(16)(Surgery Edition),87-88 (2009)

[3] Chen Changxiang, Wu Qingwen, Li Jianmin. Investigation and analysis on the status quo of nursing scientific research of clinical nursing personnel in teaching hospitals. Journal of Nursing Training 20(3),253-255 (2005)

[4]LI Shanshan,WANG Zhiwen,LI Li,et .A survey on the present nursing research abilities of senior nurses and above in Class-I A hospitals . Chinese Nursing Management 7(2), 32-34 (2007)

[5] Liang Shuang. Zhou Ying, perimeter United States, et al. Analysis Status of Clinical Nursing Research and Management Ideas . Qi Lu Nursing Medicine 15(22),33-34 (2009)

[6] Chai Yongping,Cheng Jinlian. A survey and analysis of cognition of nursing staff toward nursing scientific research. CHINESE NURSING RESEARCH 23(5) ,1239-1240 (2009)

[7]Dong Shenglian, Xing Fengmei, Chen Changxiang。Investigation on nursing research in teaching hospitals. Journal of Nursing Science 25(11),45-47 (2010)

[8] Qiu Meining, Chen Fang. A survey and analysis 0 cognition of nurses in primary hospitals toward nursing scientific research.FAMII, Y NURSE 6(9),2270-2272. (2008)

[9] Ding Cuixiang, Liang Yunxia,Yang Zhaoxia etc. Investigation and analysis on the effective shortcut of promoting nurses' scientific research ability. China Practical Medical 2(33) ,33:31-32 (2007)

[10] Zhu Liying. Investigation on the need of nursing personnel's learning needs about scientific research. Chinese Journal of Modern Nursing 13(8),2301-2302 (2007)

[11] Lu Yuzhen, Zhen Yan. Investigation on nurses' conceptions about nursing research projects. Journal of Nursing Science 23(12) ,62-63 (2008)

[12]Yang Mingying, Zhang Yue, Yuan Wei., States quo and constraining factors of nurses'scientific researches in one hospital of Yunnan province

[13] Cui Xiang, Liang Yunxia, Cai Suling, et al. Nurses' Cognitive Attitude and Level of Scientific Research and Related Investigation and Research EJ3. Chinese Practical Nursing Magazine 22(10),11-14 (2006)

[14] Wang Yi, such as, Huanghua Lan. Investigation of the Situation and Needs of Scientific Nursing Research of 16 municipal hospitals . Nursing Research 19(4) ,586-588 (2005)

[15] Lian Dongmei. Status and Ideas of Chinese Nursing Research. Journal of Nursing Chinese Practical Nursing Magazine 22(1),38-40 (2006)

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.954

Analysis to Postgraduate Employment Plight Under Dual-economy

Huilan Li ; Lei Zhou; Chunling Sun

United University of Hebei Province

Tangshan 063000, China

Key words: dual-economy; postgraduate; employment; plight; statistics

Abstract: Owing to the income gap between urban and rural, east and west, within and outside the system caused by the dual economy of our society, and the asymmetry between job opportunities and postgraduate enrollment scale, quality and requirements and so on , the structural plight in postgraduate employment appeared. In order to grasp the extent of postgraduate employment difficulties, trends and influencing factors, postgraduate employment rate, job satisfaction and the fluctuation of postgraduate enrollment in author's university were analyzed and surveyed with samples, and the questionnaire results were analyzed through multivariate statistical analysis and social science statistics software in order to deeply analyze the causes of the plight of postgraduate employment and current employment problems and provide the basis for policy-setting and theoretical guidance.

1.Introduction

The huge income gap between urban and rural, east and west, within and outside the system caused by the dual economy of our society made orientation of postgraduate employment imbalanced, and the asymmetry between job opportunities and postgraduate enrollment scale, quality and requirements and so on made postgraduate employment into the dilemma gradually. Employment of university students had been the focus of academic field, but more focuses were put on college students' employment and less on postgraduate employment. Current postgraduate employment research were involved in the general introduction to employment status , analysis to existing problems and some suggestions. In a word, these results provided an important reference for further research from different perspectives, and the downside was that the countermeasures recommended were most relatively broad and general. In addition, comprehensive analysis of multiple factors especially that of the socio-economic structural factors was scarce. What's more, statistical and quantitative analysis were not sufficient and some theoretical point of view lacked data supporting.

2.Postgraduate employment status

With the large enrollment of postgraduate for many years, the number of postgraduates increased dramatically, while the addendum of employment agencies were limited, especially the number of jobs that postgraduates were optimistic were getting down instead of adding up.

In the term of employment rate, till June 2006, the national 116 college postgraduate employment rate over the years was over 95.6 percent,showing the plight of postgraduate jobs didn't exist in general. Analysis to the employment rate of university graduates from Guangdong Province in December 2007 indicated that the employment rate of medicine students was in the lowest ranking, 5 percentage points lower than the average[1], which showed there existed significant differences between postgraduates from different majors; In the term of the

employment field, postgraduates tended to go to work in universities , research institutes and other institutions, while these fields preferred PHD, which made postgraduates in an embarrassing situation; in the term of employment flowing, the big difference between the urban and rural environmental resources caused by the dual economy made the flow of postgraduate employment more imbalanced. And almost no student was willing to work in rural enterprises.

3.Representation of postgraduate employment plight

3.1Lower employment rate.

In recent years, the employment rate of postgraduates in Jiangsu Province has been decreasing. Compared with the previous year 2004 to 2006, the employment rate decreased by 0.47, 0.66 and 1.29 percent respectively[2]. The author's university postgraduate employment statistics showed that since 2004. postgraduates' employment rate in all majors decreased 0.69 percent annually. Less employment rate and more enrollment rate indicated that the plight of postgraduate employment was getting more serious.

3.2 Decreased job satisfaction

The "Minnesota Satisfaction Questionnaire (MSQ in short)" was used to conduct the random survey aiming at the first year job satisfaction to 2009 and 2004 postgraduates after their graduation from 16 universities in Hebei province (see Table 1). The results showed that there existed significant difference between job development, pay, benefits, environment existing in the above two samples. Most of the 2009 session of postgraduates didn't think they found the ideal jobs, while the 2004 postgraduates were satisfied with the jobs, which showed that in recent years, postgraduates under the growing pressure had to lower their expectations to employment.

3.3Increased cost

In recent years, postgraduates paid more in job searching, particularly female postgraduates.. Compared with three years ago, the time for first signing job contacts for postgraduates from Hebei Province delayed 2 to 3 months averagely. In order to find jobs, some postgraduates utilized all networks of resources and paid additional social cost. The sex discrimination in the job market made female postgraduates in employment difficulty. Thus, in order to increase their chances of employment, some postgraduates had to continue to pursue a Ph.D degree, in particular female postgraduates. .

4. Multivariate analysis to employment plight

4.1 Market factors

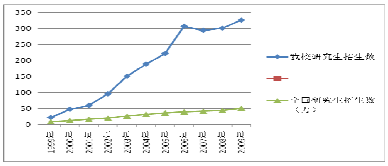
Table 1 Comparison of job satisfaction between 2009 session and 2004 session postgraduates from 16 universities in Hebei Province (P <0.05)

Variables		2009 session group（n1=200）	2004 session group（n2=100）	Z value	P value
Busy working	Very satisfied	32	61	-11.684	0.000
	Not sure	81	17		
	Not very satisfied	87	22		
Opportunities to work independently	Very satisfied	41	79	-4.123	0.000
	Not sure	50	10		
	Not very satisfied	109	11		
Opportunities to do different tasks	Very satisfied	69	72	-3.340	0.001
	Not sure	38	19		
	Not very satisfied	93	9		
Opportunities to become a member of a group	Very satisfied	87	62	-10.713	0.000
	Not sure	42	13		
	Not very satisfied	71	25		

Superiors' way to treat staff	Very satisfied	41	66	-10.578	0.000
	Not sure	84	17		
	Not very satisfied	75	17		
Decision-making competence of managers	Very satisfied	68	53	-11.414	0.000
	Not sure	59	15		
	Not very satisfied	73	32		
Not doing things against conscience	Very satisfied	81	78	-1.042	0.298
	Not sure	60	10		
	Not very satisfied	59	12		
Stable employment provided by the job	Very satisfied	53	81	-4.715	0.000
	Not sure	31	9		
	Not very satisfied	116	10		
Opportunities to do things for others	Very satisfied	64	48	-14.412	0.000
	Not sure	79	31		
	Not very satisfied	57	21		
Opportunities to order others to do something	Very satisfied	33	52	-6.915	0.000
	Not sure	46	12		
	Not very satisfied	121	36		
Opportunities to develop abilities	Very satisfied	41	74	-5.678	0.000
	Not sure	42	11		
	Not very satisfied	117	15		
Ways that company policies are put into practice	Very satisfied	62	61	-9.387	0.000
	Not sure	57	14		
	Not very satisfied	80	25		
pay and workload	Very satisfied	46	72	-7.725	0.000
	Not sure	21	11		
	Not very satisfied	133	17		
Opportunities for promotion	Very satisfied	49	64	-7.890	0.000
	Not sure	24	13		
	Not very satisfied	127	23		
Opportunities to use one's own judgments	Very satisfied	51	69	-3.495	0.000
	Not sure	59	12		
	Not very satisfied	90	19		
Opportunities to do things in one's own way	Very satisfied	44	57	-7.019	0.000
	Not sure	47	14		
	Not very satisfied	109	29		
Working condition	Very satisfied	40	71	-8.395	0.000
	Not sure	22	13		
	Not very satisfied	138	16		
Colleague relationship	Very satisfied	64	51	-14.411	0.000
	Not sure	59	23		
	Not very satisfied	77	26		
Praise from good work	Very satisfied	69	57	-14.421	0.000
	Not sure	61	10		
	Not very satisfied	70	33		
Sense of achievement derived from the work	Very satisfied	36	73	-8.692	0.000
	Not sure	22	9		
	Not very satisfied	142	18		

4.1.1 Imbalance between supply and demand.

First, the supply increased sharply. With the growing enrollment of students over years, the average scale of graduate enrollment increased over 26.9 percent annually. Take the author's university for example, in 1999, the number of enrollment was 22 , and in 2009, the number was 340, which showed the growing rate of enrollment increased 15.5 times in this period. (See Figure 1)



Graph 1: Comparison between the number of national graduate enrollment and that of graduate enrollment in North China Coal Medical University from 1999 to 2009

According to the statistics from the Ministry of Personnel in 2002, the supply and demand ratios of

professional positions were 8.25:1，3.17:1，2.83:1，2.54:1 in the fields of arts, sports, medicine, secretarial and management and so on , which showed clearly that supply exceeded demand in employment market[3]. Second, the demand decreased. Postgraduates were in an awkward position in the job market. On the one hand, universities, research institutes were their relatively favorite working place. On the other hand, the buyer's market made the employers prefer doctoral students. Additionally, some employers in marketing preferred college students instead of postgraduates owing to lower costs and stability. Third, the diploma screening signal distorted. Screening Hypothesis indicated that the level of education was the expression of personal ability, and education provided a market signal of classification based on individual ability. On the postgraduate employment market, information of trading parties was not the same. Under such circumstances, education as a filter signal on the labor market conveyed information of the potential employees' labor productivity to employers. The enrollment expansion led to talents' mixed quality, even unqualified talents existed. Diploma distortion as a screening signal made wise employers determine the demand price according to average pricing principles so as to reduce the errors caused by the loss of screening, which resulted in the declined needs to postgraduates in the same price level[4].

4.1.2 Market barriers.

According to a 2006 survey in job markets of Wuhan university students, postgraduates were victims of discrimination in height or gender as high as 64.68 percent , in household register as high as 22.94 percent in their job searching.

Post advertisements generally had clear requirements of age, but for most postgraduates with jobs , they can't meet the requirement. Female graduates in school accounted for 32.4 percent of the total number in 1999. and 39.91percent. in 2003. Current difficult employment for female postgraduates had formed a strange cycle: difficult employment for college graduates forced them to be a postgraduate , and difficult employment for postgraduates forced them to be a graduate with doctor degree, which showed the higher the education level, the greater proportion of girls, the more difficult employment; the biggest impediment to employment in today's labor market was the household register restrictions. Many jobs made clear requirement for the household register. For example, Chaoyang District of Beijing needed 123 civil servants in 2006, in which 57 positions required the candidates with permanent household register in Beijing .

4.1.3 Inadequate information

At present, job site market of postgraduates was not perfect, and labor supply and demand information exchange in the period of transition was still in the non-institutionalized situation and lack of openness and easy access. Some popular and tight demand-side information as a scarce resource was not always fully available to the market, resulting in the chance of blindness in both supply and demand options and reducing the effectiveness of market allocation of human resources.

4.2 Individual factors

Professional competence of some postgraduates were in decline. Large-scaled enrollment of graduates resulted in the poor training quality of postgraduates significantly. In addition, postgraduates' limited experience also led them have no chance with their favorite posts. The detachment between the quality structure of postgraduates and demand of the market would inevitably lead to their decreased core competitiveness; On the other hand, career psychology of postgraduates also made the employment in plight. Demands in professional counterparts, job stability in the part of postgraduates were still generally high; Some students took graduate entrance examination with utilitarian purposes, like easy job searching , employment, promotion and so on. Impure motives led them not to study hard, which not only reduced the value of diploma, but also

caused them break contracts after signing them frequently which induced employers to weaken the function of postgraduate education.

4.3 Factors in training units.

4.3.1 Detachment between talents training and job market.

Professional setting, and training styles in universities were in a stereotyped way. And large enrollment within the framework of the existing disciplines led to disciplinary structural imbalance. The "products" with homogenization and uneven academic structure deviated from the individual demand of market.. Therefore, postgraduate social adaptability was poor, resulting in their difficulties in finding jobs.

4.3.2Lagging employment promotion mechanism

The current postgraduate career guidance had many deficiencies: career guidance was only confined to graduating class with lower coverage , often limited to low-level basic services instead of high-level, individual-oriented instruction. Although most colleges and universities offered employment guidance courses, teachers read the book directly, and the teaching materials were boring. In addition, unskilled staffs with poor professional quality in career guidance programs also made the guidance difficult to play a role in employment promotion.

4.3.3Lack of training in ability.

Postgraduate education put more emphasis on theory teaching instead of ability training. Large enrollment made one tutor guide a number of graduate students, who had fewer opportunities of practice. However, the U.S traditional university education focused on thinking and ability training to students, which were our shortcomings in China's higher education.

In fact, in spite of the growing pressure of employment today, there were many postgraduates who were willing to become the city "ants". Analysis of postgraduate employment couldn't be divorced from the dual characteristics of our social economy. Here we could use Todaro model of rural residents flowing under the background of dual economy to analyze the postgraduate employment. In the Todaro model, rural –urban population emigration in the short time indicated increasing function between urban and rural net income. Namely, the greater the net difference between urban and rural income, the more the number of population in the rural-urban influx and vice versa. Taking into account the reality of the dual economy, postgraduates tended to follow the principle of Todaro model in their job searching because middle and large cities could provide higher employment household register income, more social security benefits and greater wage income within the system than the small towns. Therefore, the root of difficult employment situation for postgraduates was the great income gap in the dual economy ,which led to a large number of students prefer employment within the city system.

References:

[1] Announced Employment Rate of University Graduates in Guangdong Province in 2007 [EB / OL]. [2008-06-03]. Http://fjgb.com.

[2] Xie Jingxian. Jiangsu Graduate Students Loving Home in Difficulty in Finding Job [N]. Modern Express ,2007-03-30.

[3] Qian Jie, etc.. Postgraduate Employment: Crossroad Between Ideal and Reality [J]. Education and Occupation, 2007 (7), p71.

[4] Sun Baicai etc. Quality of Postgraduate Education and Postgraduate Employment: Diploma Screening Perspective [J]. Higher Education Exploration, 2007 (3), p23.

doi:10.4028/www.scientific.net/AMM.50-51.959

BP neural network and multiple linear regression in acute hospitalization costs in the comparative study

Jianhui Wu,Guoli Wang,Jing Wang and Yu Su

Hebei Province Key Laboratory of Occupational Health and safety for Coal Industry,

Division of Epidemiology and Health Statistics, North China Coal Medical College, Tang Shan 063000, China

Email:wujianhui555@163.com

Key word: BP neural network multiple linear regression LM Algorithm Prediction

Abstract. The BP neural network is the important component of artificial neural networks, and gradually becomes a branch of the computation statistics. With its many characteristics such as large-scale parallel information processing, excellent self-adaptation and self-learning, the BP neural network has been used in solving the complex nonlinear dynamic system prediction. The BP neural network does not need the precise mathematical model, does not have any supposition request to the material itself. Its processing non-linear problem's ability is stronger than traditional statistical methods. By means of contrasting the BP neural network and the multi-dimensional linear regression ,this article discoveries that the BP neural network fitting ability is more stronger, the prediction performance is more stable, may be further applied and promoted in analysis and forecast of the continual material factor .

Introduction

Because the BP neural network its may to the linearity or the misalignment multivariable, in does not suppose the prerequisite in the situation to carry on the statistical analysis, the same tradition's statistical method needs the variable which analyzes to tally certain condition to compare has its own merit[1-2]. The BP neural network has materially realized the mapping function from the input to the output, but the mathematical theory had proven that it has function that realizes any complex misalignment mapping. So it especially qualifies to solve the complex question of internal mechanism. We do not need to build the model and understand its internal procedure. We only need the input and obtain the output. So long as the structure of the BP neural network is outstanding, the questions less than 20 input functions can restrain to the lowest error or so within 50000 times learning. Moreover theoretically, a three hierarchical neural network can approach the given function by the random precision. This is the very attractive expectation[3-4].

Principle of the BP neural network

The BP neural network is composed of the forward-propagating and the ant propagation. In the process of forward-propagating, the infed information treated with in implicit hierarchical unit passes from input level to the output level. Each neuron's condition only influences next neuron's condition. If it cannot obtain the expected output in the output level, it will transfers to the ant

propagation--- makes the error signal return along the connection way, and makes the error signal be the smallest by modifying connection weights between each neuron. The BP algorithm is a learning algorithm proposed for multi-layered perceptions (MLP). The BP neural network's transmission function among the concealed levels uses continuous and differentiable nonlinear function, and usually is the sigmoid function. The transmission function among the output levels may uses the linear function or the sigmoid function, which is determined by the distribution of the output layer's vector. Basic steps: ①Initialization weight and threshold value, $w_{ji}(0), \theta_j(0)$is is small stochastic value. ②Importing the training samples: Input vector x_k ， k=1， 2， …,P; Expected output d_k, k=1， 2， …,P； Iterating ③ to ⑤for each input sample. ③The condition of computer network's actual output and implicit strata unit: $o_{kj} = f_j\left(\sum_i w_{ji} o_{ki} + \theta_j\right)$。 ④Computation training error: Output level $\delta_{kj} = o_{kj}(1-o_{kj})(t_{kj}-o_{kj})$， Concealed level $\delta_{kj} = (1-o_{kj})\sum_m \delta_{km} w_{mj}$。 ⑤Revision weight and threshold value: $w_{ji}(t+1) = w_{ji}(t) + \eta\delta_j o_{ki} + \alpha\left[w_{ji}(t) - w_{ji}(t-1)\right]$ $\theta_j(t+1) = \theta_j(t) + \eta\delta_j + \alpha\left[\theta_j(t) - \theta_j(t-1)\right]$ ， η for the length of stride， t for the times of iteration. ⑥Every time that k experiences from 1 to P, judging whether the target satisfies the requirement of accuracy or not. May see, the BP neural network model makes the function question of input and output sample's function transfer to a nonlinear optimized question, and has used the ordinaries gradient descent in the optimized techniques. If we regard the neural network as the mapping between the inputs and the output, this mapping is an altitude nonlinear mapping.

BP neural network and multiple linear regression modeling

Example synopsis

The suitable controllable policy of medical expense plays a key role in effectively operating the medical security system. Therefore seeking the suitable methods of anglicizing the factors which influence the hospital expense has become the question which urgently waits to be solved. [5] Formerly, massive researches for the hospital expense use the method, such as the multiple linear regressions, the variance analysis and so on, which request to satisfy the independent hypothesis. [6] However, the same expense category's case often has the similarity in hospital expense, and its observating value is not fully independent. Because has not considered the data hierarchy structure's existence in this kind of material, the multiple linear regressions and the variance analysis cause to mistaken analysis for influential factors, and will draw the completely opposite conclusions [7-8]. This article takes the hospitalized expense as a dependent variable, the influencing factor as the independent variable, and separately establishes the BP neural network model and the multiple linear regression models.

The result of the BP neural network model

Using the LM algorithm to establish the BP neural network model, model various parameters is as follows.

Table 1 in hospital expense BP neural network model parameter list

Network architecture parameter	Network training parameters	Simulation results of test set	Training set fitting results
Hidden layers: a layer	Training Algorithm: LM algorithm	R =0.813	R =0.84747
Hidden layer neurons： 35	A total cessation of training iterations: 18	R^2 =0.72	R^2 =0.793
Input layer neurons: 10	Learning speed： 0.01	R^2_{adj} = 0.704	R^2_{adj} = 0.782
Output layer neuron： 1	Performance function： SSE	SSE =1.8133e+009	SSE =4.6429e+009
	Stop training SSE^*=1.21781	MSE =1.646e+007 $RMSE$ =5736.7	MSE =37036e+007 $RMSE$ =6337.9

* The SSE for the normalized [0, 1] in terms of data.

The Results of Multiple Linear Regression Model

The multiple linear regression result showed that the hospital expense's influence is the days in hospital, the payment way, the secondary diagnosis, the surgery situation, the condition in hospital, the occupation, from large to small in turn, and the days in hospital plays the most important role in hospital expense.

Table 2 analysis result for hospital expense influencing factor multivariate linear regression

Variable	B	Beta	t	Sig
Constant term	3783.923		8.854	0.000
Payment way	-589.798	-0.343	-14.329	0.000
Secondary diagnosis	2991.692	0.327	13.477	0.000
Surgery situation	2519.017	0.238	10.334	0.000
Hospitalization days	91.080	0.560	21.377	0.000
In hospital condition	-294.522	-0.064	-2.830	0.005
Occupation	-151.387	-0.064	-2.822	0.005

Model comparison

Model fitting ability comparison

When establishing the BP neural network model, we can use the LM algorithm, divide the data set into the test collection and the training regulations. The multiple linear regression model uses the same data to be divided into the test collection and the training regulations. Two models' fitting targets are as follows:

Table 3 model fitting targets

Indicators	Multiple linear regression model		BP neural network model	
	Test set	Training set	Test set	Training set
R^2	0.151	0.485	0.72	0.79
R^2_{adj}	0.123	0.48	0.704	0.782

Compared with two kinds of model's fitting targets, regardless of being tests the collection or the training regulations, the BP neural network model's fitting ability is better than the multiple linear regression models. Some scholars' research conclusion is the same [9].

Model forecast performance comparison

According to the random principle, we draw 500 samples randomized, separately use two kinds of models to forecast, extract the relative error compared with the real value, carry on the pair the T-test, and thus compare the forecasting qualities of the two groups of model. The result is as follows:

Table 4 multiple linear regressions and the BP neural network forecast that the performance compares

Methods	case	mean（$\bar{x} \pm s$）	difference（$\bar{x} \pm s$）	t	P
Multiple linear regression	500	0.572 ± 0.007	0.148 ± 0.113	1.31	0.000
BP neural network	500	0.32 ± 0.16			

by Table 4 knows: the relative error average value of the multiple linear regression model is 0.572, and the BP neural network model's is 0.32，t=1.31，$P < 0.000$。The above result indicated that two methods have the difference, and the BP neural network has better predictive ability through two groups of relative error average value's size.

Conclusion

BP neural network can be due to the linear or nonlinear multi-variable in the absence of the prerequisite conditions for statistical analysis, compared with traditional statistical methods has its advantages. A good BP neural network can even incomplete or biased information is not the case can be made correct predictions. Using the data of hospital expense and the influencing factor, the article has separately established the BP neural network model and multiple linear regression mold, and obtains two kinds of models' the goodness of fit targets R^2、R^2_{adj}. Comparing the model forecasting performance, I find that the stability and the predictive ability of the BP neural network is better than the multiple linear regression model. Therefore I think that the BP neural network can grasp the intrinsic regularity of data, especially complex misalignment relations more easily than multiple linear regressions. And because the BP neural network model does not need to consider

that whether covariate is independent, whether the strain capacity satisfies normal distribution and so on, in hospital expense and continual material, I suggested that we use the BP neural network model. Although the BP neural network has many advantages, but it is currently not as the traditional statistical methods are well known and applied, in the BP neural network before being widely accepted, there is much work to do[10]. Such as BP neural network is still in a "black box" stage: With the number of input variables have to come out after the corresponding results and this during the process and its clinical significance still needs to be improved and clarified. BP neural network at the same time as a new statistical method of testing its weight coefficient and confidence interval calculation, network training with the best principles, the choice of the number of hidden layers of meaning and interpretation of both their specific needs further study.

References

[1] Ma Yinqi, Zhou Lifeng, Jin Peihuan. Artificial neural network application in statistics,*Medical Information* 6(11),21-24 (1998)

[2] Wang Jing, Li Man, Multivariate Statistical Methods in the Study of hospital fees progress. *China Health Statistics* 26(1),91-95 (2009)

[3] Wang Junjie, Chen Jingwu, BP neural network theory and its applications in medical statistics in design skills. *China Health Statistics* 25(5),547-549 (2008)

[4] Su Gaoli, Deng Fangping, Based on the MATLAB language improved algorithm of BP neural network. *Bulletin of science and technology* 19(2),130-135 (2003)

[5] Zhang Juying, Wei Jian, Yang Shuqin, Neural network model factors in the hospitalization cost analysis. *Chinese Journal of Hospital Administration* 18(3),143-145 (2002)

[6] Yuan Qing, Yuan Kuichang, Diabetic patient's hospital fees multiple regression analysis. *China Health Statistics* 25(4),360-362 (2008)

[7] Fang Jiqian, Health Statistics, Fifth Edition, *People's Medical Publishing House* (2008)

[8] Liu Biyao, Shen Yi, Based on BP Neural Network Modeling of hospital costs. *Zhejiang University, a master's degree thesis* in 2006.

[9] Zhou Lifeng, Gao Ersheng, Jin Pihuan. Comparison Between Neural Networks And Multivariate Linear Regression Method. *Modern Preventive Medicine* 25(3),272-274 (1998)

[10] Bartfay E, Mackillop WJ, Peter JL. Comparing the predictive value of neural network models to Logistic regression models on the risk of death for small-cell lung cancer patients. *Eur J Cancer Care (Engl)* 15(2),115-124 (2006)

doi:10.4028/www.scientific.net/AMM.50-51.964

Comparison of BP Neural Network Model and logistic regression in the analysis of influencing factors of violence in hospitals

Jianhui Wu, Guoli Wang, Xiaoming Li and Sufeng Yin

Hebei Province Key Laboratory of Occupational Health and safety for Coal Industry, Division of Epidemiology and Health Statistics North China Coal Medical College, Tang Shan 063000, China.
Email：wujianhui555@163.com

Key Words: BP neural network; logistics regression; hospital violent; ROC curve

Abstract: Collecting violence cases for medical personnel from different levels of the hospital of Tangshan, we create a model for influential factors of hospital violence, and respectively with BP Nerve Network and logistic regression, by sensitivity, specificity and ROC curve, it is compared with two methods,in order to discovering effective analytical method . The training set and testing set sensitivity of BP Neural Network Model are 0.916 and 0.935,and the specificity is 0.447 and 0.526,the area of ROC curve is 0.769 and 0.785； for logistic regression Model ,for its the training set and testing set, sensitivity is 0.907 and 0.925, the specificity is 0.377 and 0.404, the area of ROC curve is 0.663and0.666. In hospital violence influencing factors, the forecast capability of BP Neural Network Model is better than logistic regression Model and it has farther extend value.

Introduction

The violent incident disturbed the natural order and the operators'life,it has brought some negative influences, so it has became an important issue[1-2].For control violence-occupational hazard in hospitals,it is important to seeking hazard factors. Logistic regression[3-5] has always been used for analysis influential factors of workplace violence, but it asked absolute information and S curve relation between independent variable and dependent variable , according with multiplicative model for various independent variables, and so on, the influencing factors don't meet these conditions, so the result is dissatisfy.

The characteristics of BP network include distributed information stored, parallel information disposal manner, tolerance of fault, the ability of self-organization and self-learning, and adaptive and nonlinear processing capability. Compared with traditional medicine statistical method, the artificial neural network needn't exact math model, it hasn't request for variable, so it can make up the deficiencies of traditional statistical method, especially, when the intent not achieve or the effect not good by statistical method in existence, the artificial neural network model has beautiful effect. In the field of statistics, it has been applied to resolve the problem about forecast or distinguish.

The research sets up a model for influencing factors of hospital violence by BP Neural Network Model and logistic regression we compare the two models, the intention is discovering effective analytical method.

The principle of BP neural network

BP neural network is a nonlinear dynamic system [10,11], by a number of functional distribution of single neurons in the parallel composition, BP algorithm from the data stream prior to the calculation (forward propagation) and the error signal back-propagation 2constitutes a process .Positive communication, the propagation direction of input layer →hidden layer→ output

layer, each layer neurons under the influence of the state of only a layer of neurons. If not desired in the output layer output error signal is turned back propagation process. By alternating the two processes, the implementation of the right vector space error function gradient descent strategy, search for a group of dynamic iteration the weight vector, the network error function reaches a minimum, thus completing the process of information extraction and memory. The neural network is widely applied in nonlinear system modeling, control and identification etc for it has better capability[12-14].

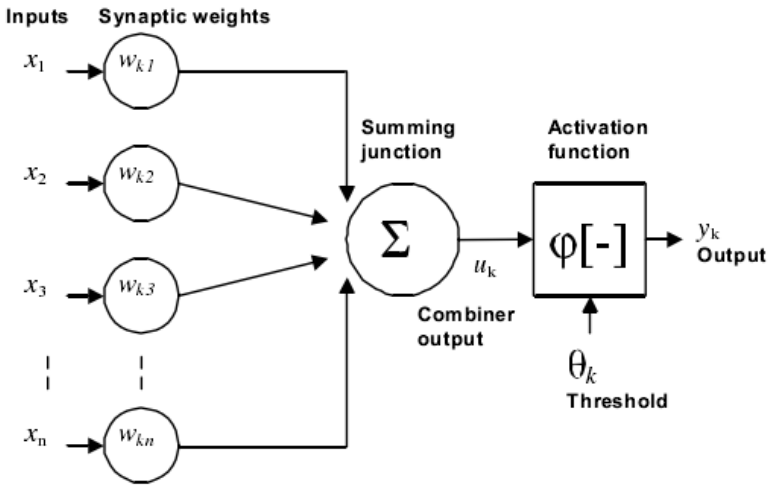


Figure 1.Model of the neuron in an artificial neural network

The model of BP Neural Network Model and logistic regression

Problem description

By the sampling methods of whole group, we conducted an investigation into the medical staff for 2 grade A hospital and grade 2 a hospital in tangshan city,it includes: ①the feature of social demographic, Such as gender, age and marital status, education and title and so on; ②job context, include level of hospital,work department and so on. 1331 people were investigated, the effective questionnaire was withdrawn 1,235,and the response rate is 92.79%.

Table1.The quantitative method of hospital violence and their influence factors

number	variable	valuation
X1	hospital rank	0= secondary hospital;1= tertiary hospital
X2	age	1=~30;2=~40; 3=~50 4=51~
X3	sex	0=female;1=male
X4	diplomas	1= secondary school and below;2= undergraduate;3=postgraduate and over
X5	title	1=no title;2=primary;3= intermediate level;4= associate senior;5= senior
X6	marriage	1= unmarried;2= married;3= widowed or divorced
X7	emergency	0=no； 1=yes
X8	peditatrics	0= no ;1= yes
X9	stomatology	0= no ;1= yes
X11	gynecology	0= no ;1= yes
X12	surgery	0= no;1= yes
X13	assistant department	0= no;1= yes
X14	other department	0= no;1=yes

The result of logistic regression model

These factors as independent variable in Table1. Whether violence take place,it as dependent variable,we progress Logistic regression analysis,the result is:

Table2. the result of Logistic regression analysis

variable	B	*P*	Exp(B)	95% C.I.for EXP(B)	
				Lower	Upper
hospital level	-0.92	0.00	0.40	0.31	0.51
other department	-0.40	0.01	0.67	0.51	0.89
emergency	1.53	0.00	4.62	2.80	7.64
marital status	0.49	0.00	1.64	1.28	2.09
constant term	0.021	0.93	1.02		

From the Table2,we can clearly see,after implementing Logistic regression analysis, factors of hospital violence included hospital level, other department, emergency, marital status.

Modelling BP Neural Network

These factors as independent variable which are hospital level, other department, emergency, marital status. Whether violence take place, it as dependent variable, a method of systematic sampling has been adopted ,using the method, data is divided into the training set and testing set, the BP Neural Network model is established by Levenberg-Marquardt[15]（LM）optimization algorithm. As follows,Table3 is main parameter.

Table3 the BP Neural Network model parameter of influential factors of violence

Network Structural parameters	Network Training Parameter	Simulation result	Fitting- result
number of hidden layer：1	Training Algorithm：LM	SSE=43.03	SSE=129.35
neuron number of hidden layer：7	iteration times： 17	MSE=0.17	MSE=0.18
neuron number of input layer：4	Learning Speed： 0.01	RMSE=0.42	RMSE=0.42
neuron number of output layer：1	performance function：SSE		

The model compare

The data of the training set and testing set is fitted by two models, as follows, the Table4 regard sensitivity, specificity and area of ROC curve.

Table 4.the comparative result of the training set for two models

model	total percentage	sensitivity	specificity	area of ROC curve
Logistic regression	741	0.907	0.377	0.663
BP Neural Network	741	0.916	0.447	0.769

Table5. the comparative result of the testing set for two models

model	total percentage	sensitivity	specificity	area of ROC curve
Logistic regression	247	0.925	0.404	0.666
BP Neural Network	247	0.935	0.526	0.785

From Table 4 and Table 5,we can see whether it is the training set or testing set, BP Neural Network is superior to Logistic regression.

Conclusion and Outlook

The paper establish BP Neural Network model and logistic regression model by the data of influential factors of workplace violence, it is compared with two methods by sensitivity, specificity and ROC curve, we discover BP Neural Network is superior to Logistic regression. And Bp the nerve network model does not need to consider the conditions which is independent variable and so on, so Bp the nerve network should be widely used for the judge classifications.

References

[1] CDC.Violence-Occupational hazard in hospitals[R]. DHHS (NIOSH)-Publication, No 2002-101.2002,1-10.

[2] Sofield L, Salmond SW. Workplace violence. A focus on verbal abuse and intent to leave the organization. Orthop Nurs.2003,22(4): 274-283.

[3]Susan Megison, Lynn Westbrook. Hospital Dissemination of Information Resources on Intimate Partner Violence: Statewide Analysis of Texas Emergency Room Staff. Journal of Hospital Librarianship. 2009,9(3):231-248.

[4]Yang Xiaoduo, Jiang Yafang, Li Chun, et al. The prevalence and influential factors of workplace violence against nurses in general hospitals. Chinese Journal of Nursing,2009,44(10):877-881.

[5]Liu Yao, Deng Ling, Yu Min, et al. Investigation of psychological violence at hospital workplace in Guangzhou. Chinese Journal of Public Health,2009,25(9):1050-1051.

[6]Zhu Taiqi，Shi Hui . The artificial neural network theory and application..Science press，2006：21.

[7]Li Lixia, Wang Tong ,Fan Fengxi. Strategies about how to Select the Appropriate Parameters for BP Neural Network. Modern Preventive Medicine，2005, 32(2):128-130.

[8]Wang Junjie, Chen Jingwu.BP Neural Network Apply in Disease Prediction. Journal of Mathematical Medicine，2008,21(3):258-259.

[9]Vincenzo Savarino ,Pietro Dulbecco . Can artificial neural networks be beneficial in diagnosing gastro-oesophageal reflux disease. European Journal of Gastroenterology & Hepatology ,2005, 17:599–601.

[10] Hornik K, Stinchcome M, White H.Multilayer feedford networks are universal approximaters. neural networks 1989,2:359~379.

[11]Therneau TM, Grambsch PM, Fleming TR. Martingale-based residuals for survival models. Biometrika 1990,77:147~153.

[12]Wang Xiaohua, He Yigang. Optimal design of frequency-response-masking filters using neural networks. Acta Electronica Sinica.2008,36(3):486 一 489.

[13]S S Ge, C C Hang, T H Lee, et al. Stable Adaptive Neural Network Control. Boston, MA: Kluwer, 2001.

[14]Tian Guang, Qi Feihu. Feature transformation and SVM based hierarchical pedestrian detection with a monocular moving camera. Acta Electronica Sinica 2008 , 36(5):1024 一 1028.

[15]Robert Hecht-Nielsen .Theory of the Backpropagation Neural Network. Department of Computer Science and Engineering, University of California at San Diego, Preliminary Manuscript, January 2006.

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.968

Application of the BP nerve network model in the analysis of factors affecting hospitalization expenses

Guoli Wang,Jianhui Wu ,Ting Li and Xiaoming Li

Hebei Province Key Laboratory of Occupational Health and safety for Coal Industry,

Division of Epidemiology and Health Statistics, North China Coal Medical College, Tang Shan 063000, China

Email: wujianhui555@163.com

Key Words: the BP nerve network model; Hospitalization expenses; Influential factor

Abstract. Collected 2218 medical records of cerebral infarction patients from 2007 to 2008 in a third-class hospital in Tang Shan as sample. A statistical analysis was made of the hospitalization expenses and its influencing factors by BP nerve network model to identify the factors affecting hospitalization expense by using BP nerve network model. The results showed that the major factors impacting the hospitalization expense included age, hospitalization day, therapeutic result and emergency medical treatment. The sensitivity were: age is 0.91426, hospitalization day is 0.32131, therapeutic result is 0.19722, emergency medical treatment is 0.27564. BP nerve network model is a powerful analytical tool for hospitalization expenses and its influencing factors analysis.

Introduction

In recent years, the medical costs have risen sharply. The Ministry of Health promulgated a report in 2009 showed that the average hospital expenses of city is 8958 yuan, and the average hospital expenses of village is 3685 yuan[1]. High medical costs, especially the hospitalization expense give people brought serious economic burden. Control of the high medical costs has become a hot social concern[2].

In some studies of medical expenses, the analysis methods were multiple linear regression and analysis of variance[3-5], however, the medical expenses don't obey the normal distribution[6], and affected by various factors. In this case, the traditional linear structure model is difficult to accurate fitting. It is necessary to consider the suitable statistical methods for cost data characterri- stics.

Back Propagation neural network, refers to based on the BP algorithm multi-layered perception, the BP algorithm is one kind based on the erroneous anti propagation principle training algorithm, it take the least error mean square as the criterion, and uses the misalignment differentiable function to carry on the weight training. Namely the erroneous counter-biography error anti propagation algorithm's learning process is composed of the information forward-propagating and error's anti propagation two processes[7]. Compares with the traditional statistical method, the BP neural network does not need the precise mathematical model, does not have any supposition request to the material itself, its processing non-linear problem's ability is higher than traditional generally the statistical method.

Cerebral infarction is one of three big cause of death diseases, cerebral infarction patient's in hospital expense accounts for various countries medical expense 2%~4%, along with the world population aging, this proportion will further grow[8]. In our country, cerebral infarction the total

disease incidence rate not yet to see the report, according to the regional investigation, every year approximately is 166.07～199.96/100,000[9], if each brain stem dead patient is hospitalized the expense is 8400 Yuan, every year the country, the society and the family altogether spend 195.2～23,520,000,000 Yuan, takes in the development the large population, this is a huge medical expenditure. In view of this, this article take cerebral infarction patient's in hospital expense as an example, carries on the discussion on the BP neural network model in this aspect's application, provides in the methodology for the similar research the reference.

Model construction

The BP neural network is one kind of multi-layered forward feed neural network, the weight adjustment uses the anti propagation the learning algorithm, his/her neuron's transition function is the S form function, the output is 0~1 between continual quantities, may realize from the input to the output random misalignment mapping[10]. Through to network parameter's selection, after having determined the network layer, the concealment level nerve number, the initial weight, the study speed, the expected error and the maximum length of stride, construction neural network. After definite network structure, carries on the training using the input output sample, is also makes the adjustment to the network, the multiple relapse, until sample restraining, causes the input output mapping relations which the network realizes assigns, thus discovers the most important influencing factor.

We fitted 1 concealed level list to output the BP neural network, the maximum training length of stride was 1 000, study speed by 0.01, expected error basis actual situation in 0.00~0.02 scope selection, after newff function construction and initialization BP neural network, caused the weight and the threshold value supposes as far as possible[- l,1]Between. The algorithm selection uses stops the strategy early step tangent BP training function Trainoss, take the sex, the age, the marital status, in hospital number of times, the rescue, the payment way, in hospital number of days and the treatment result and so on as the input variable, in hospital expense is the output variable, is carrying on MATLAB7.0 and SAS10.0the analysis.

(1)BP net training regulations, the verification collection and test collection's determination divides into the data set the training regulations (60%), confirmation collection (20%), test collection (20%), prevents the model fitting. Carries on the first sample data set according to the input variable sorting, selects in turn according to the equal-space, record number 2, 7, 12, 17 .....For test collection, 4, 9, 14, 19 .....For confirmation collection, other for training regulations

(2)input output variable pretreatment in this research, because the input variable including the classified variable (sex, marital situation, extension does turn over to situation, whether there is to rescue, payment category) and the continuous variable (in hospital number of times, age, in hospital number of days), the output variable (in hospital expense) is the continuous variable, to cause the input output variable to fall, in the transfer function changes on the quickest gradient, uses the standard reduction to input the output variable to carry on the normalization, after causing various variables normalization, the maximum value is 1, the minimum value is 0.

(3)BP neural network model establishment each nerve number determination: Regarding the input level, its nerve number basis actual problem decides, is equal the independent variable which said in the statistical method; Regarding the output level, its nerve number also acts according to the actual problem to decide, is equal the dependent variable which said in the statistical method. But concealed the level nerve number goal hypothesis still not to have the unification formula at present, generally according to the experience to establish like may use the trial and error method, represented the input level nerve number by m, n represents the concealment level nerve number,

then supposed n=log2m first, if the network restraining effect were not good, then increased the nerve number again, each time increased the nerve number, until increased the network error to reduce not the obvious convergence rate again to be slow, then might consider that the stop increased; Or according to the Komlogorov theorem, to can simulate this continuous function precisely theoretically, should suppose n=2m+1.

Example analysis

The material from the Tangshan some Third-level First-class Hospital 2007~2008 year brain stem dead patient's medical record information, the rejection flaw case, does not conform to the logical case, effective case total 2218 examples. Withdraws from his system in primary data after reorganization, carries on the analysis with the Excel2003 database to compile. Patient medical record material including information and so on all-in cost, sex, age, marital status, in hospital number of times, rescue, payment way, in hospital number of days and treatment result. And, masculine 1254 people, feminine 964 people, average age (65.65±11.29) year old, average in hospital date (17.05±9.22) day; Is marrying 2199 people, unmarried, divorces and loses one's spouse equals 19 people; First is hospitalized 1568 examples, must first is hospitalized 650 examples; In hospital period carries on rescues the patient 199 examples, has not carried on rescues the patient 2019 examples; Health insurance payment patient 1995 examples, non-health insurance payment patient 223 examples; Cures 1628 examples, changes for the better 561 examples, 11 examples, have not died 18 examples. Patient's average in hospital expense is 8371.22 Yuan, lowest all-in cost 1035.71 Yuan, highest all-in cost 41791.22 Yuan.

Because in hospital expense and in hospital number of days is the non-normal distribution material, therefore after its normalization, carries on the analysis. In this research uses the variable name and the quantification method see Table 1

Table 1 selection medical record home page project and quantification method

Factor	Code	quantification method
Sex	x1	Male =1，Female =2
Marital status	x2	marriage =1，Others =2
Age	x3	Age
Thetimes of In hospital	x4	One times =1，≥two times =2
Rescue	x5	No =1，yes =2
Payment way	x6	Health insurance =1，Non-health insurance =2
In hospital days	x7	day
The result of Treatment	x8	Cure =1，better =2，not recovered =3，Death =4

In hospital expense BP model building After the attempt different initial weight and the threshold value as well as the different concealed level neuron, has chosen the promoted performance and the fitting performance good network finally, established the brain stem dead patient to be hospitalized the expense and the influencing factor BP neural network model, the model each parameter target see Table 2.

Table 2 in hospital expense BP neural network model parameter

Network architecture parameter	Network training parameter	Test collection simulation result	Training regulations fitting result
Concealed layer：1	Training algorithm：*OSS*	R=0.8621	R=0.84331
numbers of concealed	The number of altogether	R^2=0.74321	R^2=0.71118
level nerve：15	iterative when stop train：20		
Numbers of Input	Study speed：0.01	$R^2_{adj} = 0.41534$	$R^2_{adj} = 0.52435$
level nerve ：8			
Numbers of Output	Performance function：*SSE*	SSE=5.2389e+009	SSE=2.081e+010
level nerve：1			
	when stop train SSE=1.22463	MSE=1.6738e+007	MSE=1.7327e+007
		RMSE=4091.2	RMSE=4162.6

In hospital expense influencing factor sensitivity analysis result uses the sensitivity analysis to carry on the analysis to in hospital expense each influence factor sensitivity, reflects each influencing factor by this to in hospital expense influence size. The result see Table 3. x3 (age), x7 (in hospital number of days), x8 (treatment result) and x5 (rescue) occupies all factor influence to in hospital expense influence the major part, prompts them is in hospital expense important influencing factor.

Table 3 in hospital expense influencing factor influence sorting

Influence sorting	Influencing factor	sensitivity
1	age	0.94897
2	The times of In hospital	0.16101
3	The result of Treatment	0.15227
4	Rescue	0.14537
5	number of In hospital	0.09421
6	Marital status	0.08733
7	Payment way	0.06751
8	Sex	0.01391

Conclusion

In hospital expense's influencing factor pattern is one kind of typical multi-factor joint influences pattern, between the factor function way is complex, outputs between the factor itself also possibly to have certain relational disturbance existing statistical model the fitting efficiency, like promiscuous factor, multiple collinearity and so on, the existing linear model limits to this kind of disturbance's revision measure, also does not have processes this kind to have the disturbance information method well, the application neural network processes the non-linear problem to be able to overcome the regression model weakness.

The BP neural network through to has the representative example study and the training, can grasp the thing substantive characteristics, trains good BP the network, theoretically can approach any input (independent variable) and the output (strain capacity) between random misalignment mapping, has strongly from the organization, the auto-adapted ability, has the high fault tolerance. This model in the application the independent variable may be continual, may also be separate, does

not need to consider whether the independent variable does satisfy the normality and the variable conditions and so on independence, may distinguish the variable the complex misalignment relations, Especially uses the existing statistical method to be unable when serves the purpose or the effect is not good, uses this model often to receive the very good effect. But the traditional t analysis often has each kind of supposition condition request to the data distribution, when does not satisfy the request, possibly (how can make the variable transformation to the primary data including to cause misalignment relations transformation for linear relationship), but which one kind of functional generation in fact chooses is very difficult, possibly must use the complex statistical method. But the BP net takes a misalignment mathematical model, has the superiority in these aspects, it is good at dealing with the complex fuzzy mapping relations especially, does not need to know the data the distributed form. The BP neural network takes a standard misalignment mathematical model, has the broad application prospect in statistics.

Reference

[1] Ministry of health of the people's republic of china. http://www.moh.gov.cn/publicfiles/business/htmlfiles/mohbgt/s3582/200902/39201.htm

[2]Xing Hai-yan, Ren Guang-yuan. Analyzing of influential factors of hospital expense in patients with acute appendicitis. Modern Preventive Medicine 34（20）,3983-3985,(2007).

[3] Wei Feng-jiang, Cui Zhuang, Li Chang-ping, et al. Analysis of Hospitalization Cost and Relative Factors on Acute Appendicitis. Modern Preventive Medicine 37(5),847-849 (2010).

[4] Yang Li, Xiao Yong-hong, Nie Yao, et al. Impact of misuse of antimicrobial therapies on inpatient costs. , Journal of Peking University (Health Sciences) 2010(3).

[5] Cun Xiao-le, Femg Xiang-xian, Li Shao-xia. Analysis of hospitalization expenditures among 3287 gastric carcinoma inpatients in general hospital. Chinese Journal of Public Health 2010(6).

[6] Zhang Juying, Wei Jian, Yang Shuqin. Application of the nerve network model in the analysis of factors affecting hospitalization expenses. Chinese Journal of Hospital Administration 18(3),143-145 (2002).

[7] Jiangang Yang. Artificial neural networks practical course.zhejiang: Zhejiang University publishing house, 2001:45.

[8] Wentworth DA. Atkinson RP. Implement of an acute stroke program decreases hospitalization costs and length of stay[J]. Stroke 27(6),1040-1043 (1996).

[9] Wenzhi Wu, Kaiping Wu, Zhen Hong, Our country three city development society crowd intervenes nine year apoplexy disease incidence rate change [J]. Chinese Senior citizen cardiovascular magazine 4(1),30-33 (2002).

[10] Dong Yi. artificial neural networks algorithm model study of Regression analysis. China health statistics 15,44-46 (1998).

doi:10.4028/www.scientific.net/AMM.50-51.973

Application of factor analysis in the research of hospital medical expenses

Sufeng Yin[1], Jianhui Wu[1], Xiaojing Wang[2], Sha Li[3]

[1].Hebei Province Key Laboratory of Occupational Health and safety for Coal Industry, Division of Epidemiology and Health Statistics North China Coal Medical College, Tang Shan 063000, China.

[2]. Kailuan Hospital Affiliated North China Coal Medical University, Tang Shan

[3]. TangShan Gongren Hospital,Tang Shan 063000, China.

Email：ysfdzy@yahoo.com.cn

Keywords: medical expense of inpatients, factor analysis, influencing factor

Abstract. Case records of inpatients in a hospital in Tangshan in 2007 and 2008 are chosen for factor analysis. This paper aims at finding the latent factors influencing inpatients' medical expense. By analysing the factors which controlling and influencing inpatients' medical expense, the practical value of factor analysis in the study of medical expense is evaluated, and an analysis method in line with the structure of hospital medical expenses and its characteristics is tentatively explored. The paper comes to the conclusion that the factor of basic expense, the operation factor and the examination factor are the common factors which control and influence inpantients' medical expense. It is implied that factor analysis is a proper method used in research of hospital medical expenses.

Introduction

The increasing health expense, especially the medical expense, as a global issue, has become a huge burden on the national financial expense.[1] Therefore, it is of great significance to identify, control and adjust factors of the high medical expense through research on factors influencing the medical expense.

As an important method of multivariate statistical analysis, in recent years factor analysis has been widely employed by medical researchers and satisfactory progress and results have been achieved. But the application of factor analysis in medical expenses is still in its initial stage.

This paper aims at identifying the common factors through factor analysis of medical expenses of inpatients, evaluating the application of factor analysis in medical expenses of inpatients. At the same time, this paper tentatively explores an analysis method according to the hospital medical cost structure and characters.

The basic thought of factor analysis and the analysis process[2] [3]

The basic thought of factor analysis

Factor analysis, a multivariate statistical analysis for dimensionality reduction, is used here, by studying the inter-dependency in the correlation matrix of various indexes, to identify the common factors which control all variables. Every index is expressed as the linear combination of the common factors, in order to show the correlativity between original variables and factors.

The process of factor analysis

Suppose there are N samples and P indexes. $x(x_1, x_2 \cdots x_p)$ is the observable random variable,

and the common factor we are looking for is(F_1，F_2，… F_m). Then the model：

$$X_1=a_{11}F_1+a_{12}F2+...+a_{1m}F_m+\varepsilon_1$$

$$X_2=a_{21}F_1+a_{22}F2+...+a_{2m}F_m+\varepsilon_2$$

$$...$$

$$X_p=a_{p1}F_1+a_{p2}F2+...+a_{pm}F_m+\varepsilon_p$$

is called factor model which can also be expressed with a matrix：$X=AF+\varepsilon$

After the correlation coefficient matrix R of variables is calculated, the latent root of correlation coefficient matrix R and its corresponding unit proper vectors can be acquired. And based on that, the factor loading matrix can be calculated.

Usually the number of factors is determined according to the latent roots and the accumulated contribution rate of factors. The common practice is to adopt the ones with an accumulated contribution rate of over 80%. In actual practice, factor rotation is needed.There are two types of factor rotation: orthogonal rotation and oblique rotation. So usually the orthogonal rotation is adopted.

Orthogonal rotation usually includes quartimax, varimax, and equanmax.

In the process of factor analysis, each common factor can also be expressed as the lineal combination of variables, so that the value (factor score) of each common factor can be evaluated by the observed value of variables. Its mathematic model is $F_1=b_{i1}X_1+ b_{i2}X_2 +\cdots + b_{im}X_m(i=1, 2, \cdots m)$, in which F_1 is the i_{th} factor score. Factor scores are the final result of factor analysis.

Case analysis

Case introduction

3934 discharged cases in 2007 and 2008 are taken from the database of abstracts of medical records of a hospital of Grade 3, Class A in Tangshan city.. Valid records are taken for further study, and all the fees are renamed: bed fee, drug fee, transfusion fee, operation fee, testing fee, examination and treatment fee, radiation fee. All of them are numerical variables..

Factor analysis of fees of inpatients with the top 10 diseases

First, get the correlation coefficient matrix of all fees

The correlation coefficient examination shows that correlation exists among all variables but between blood transfusion fee and examination & treatment fee.

Principal component analysis and quartimax orthogonal rotation

Principal component analysis is made to all data and therefrom find the common factors, as is shown in Table 1.

Table 1 Result of principal component analysis

No	Latent root	Contribution rate(%)	Accumulated contribution rate(%)
1	2.6168	0.3738	0.3738
2	1.0388	0.1484	0.5222
3	0.9363	0.1338	0.6560
4	0.8287	0.1184	0.7744
5	0.7966	0.1138	0.8882
6	0.4399	0.0620	0.9501
7	0.3496	0.0499	1.0000

Under the principle of the maximum accumulated contribution rate, the minimum common factors and the latent root bigger than 1, three common factors are taken for analysis. The accumulated contribution rate of these three common factors is 65.60%.While, with principal

component analysis, the initial model of factors is built, as is shown in Table2：

Table 2 the factor model（initial model）

Variable	Factor 1	Factor 2	Factor 3
Bed	0.69231	-0.06541	0.24051
drug	0.82051	-0.15872	-0.08 67
transfusion	0.25746	0.80800	-0.12743
operation	0.41878	0.53899	0.21035
testing	0.83275	-0.10780	-0.21163
examination	0.40963	-0.15245	0.74125
radiation	0.60108	-0.17636	-0.46536

As is shown in Table 2, some original indexes are heavy loaded on both factors. For example, the operation fee is heavy loaded on both Factor 1 and 2, which is hard to explain. In order to further study the practical significance of common factors, the quartimax orthogonal rotation is undertaken, as is shown in Table 3.

Table3 factor model with quartimax orthogonal roatation

Variable	Factor 1	Factor 2	Factor 3
Bed	0.558 60	0.144 35	0.456 67
drug	0.810 76	0.072 49	0.207 41
transfusion	0.080 69	0.839 27	-0.156 53
operation	0.174 38	0.645 50	0.251 10
testing	0.853 72	0.117 59	0.085 00
examination	0.149 31	0.010 03	0.847 40
radiation	0.749 24	-0.027 15	-0.216 48

As is shown in Table 3, the common factor 1 controls mainly the bed fee, the drug fee, the testing fee, and the radiation fee, so it is called the basic expense factor; the common factor 2 controls mainly the operation fee and the blood transfusion fee, so it is called the operation factor; the common factor 3 controls mainly the examination and treatment fee, so it is called examination and treatment factor. [4] It is obvious that the basic expense factor is the most influential in the total medical expense, among which the drug fee accounts for the most part.

The factor score equation is built according to the standard factor score coefficient as follows:

$F_1=0.1652X_1+0.3525X_2-0.0529X_3-0.0598X_4+0.3921X_5-0.0987X_6+0.4226X_7$

$F_2=0.0280X_1-0.0638X_2+0.7653X_3+0.5546X_4-0.0232X_5-0.0524X_6-0.1264X_7$

$F_3=0.3344X_1+0.0386X_2-0.2111X_3+0.1845X_4-0.0940X_5+0.8125X_6-0.3661X_7$

The factor score is calculated according to the equation above.Therefore, diseases should be taken into consideration while working on controlling the medical expenses because different diseases are treated with different methods which cause different expenses. The differences are reflected in the factor scores of different diseases in this study.[5]

Conclusion

Correlation, which usually exists among factors of various medical expenses. Therefore, it is unreasonable to focus on one single fee and ignore its correlation with other fees while studying the hospital medical expenses. Neither is it reasonable to analyse the total expense only since a lot of

information may be ignored. Factor analysis is proved to be a proper method because the correlation among variables are taken into consideration and the common factors which can control and influence all factors can be identified, without loosing the original information of variables. The result of this study shows that there are three main factors influencing the medical expenses. After knowing about these three factors, we can know better about the latent factors controlling the hospital medical expenses, so that the expenses can be reasonably employed and controlled.

References

[1]Zhao Cun Xian, Chen Jun Guo. Study and Strategy of Policy on Increasing Medical Expense. *ChongQing Medicine* 38(1),10-11 (2009)

[2]Zhang Xiao Dong, Han Fu Rong.Application of Factor Analysis in Assessment of Performance Excellence. *Journal of BeiJing University of Technology* 33(10),1117-1120 (2007)

[3]Zhang Lin Quan.Research and Application Based on Factor Analysis.Natural Sciences *Journal of Harbin Normal University* 25(5),60-63 (2009)

[4]Li Guang Rong,Wu Lu Nan. *Combined Analysis and Forecast of Implications on Hospital Cost. Value Engineering* 29(2),63-64 (2010)

[5] Sun Jia Nan,Hou Lei. Factor Analysis of Medical Expenses of Respiratory Disease. *Chinese Journal of Hospital Statistics* 16(2),100-102 (2009)

doi:10.4028/www.scientific.net/AMM.50-51.977

Based on BP neural network discrete data forecast

Wang Jing Wang Guoli Wu Jianhui, and Su Yu

Hebei Province Key Laboratory of Occupational Health and safety for Coal Industry,

Division of Epidemiology and Health Statistics, North China Coal Medical College, Tang Shan 063000, China

Wujianhui555@163.com

Keywords: BP neural network Discrete data Residual Prediction

Abstract. Artificial neural network is based on human brain structure and operational mechanism based on knowledge and understanding of its structure and behavior of simulated an engineering system. BP artificial neural network is an important component of neural networks, as it can on the linear or nonlinear multivariable without preconditions in the case of statistical analysis, with the traditional statistical methods, analysis of the variables need to be consistent with certain conditions compared to its own advantage. The BP neural network does not need the precise mathematical model, does not have any supposition request to the material itself. Its processing non-linear problem's ability is stronger than traditional statistical methods. This article uses two groups of data to establish the BP neural network model separately, and carries on the comparison to the model fitting ability and the forecast performance, discovered BP neural network when data distribution relative centralism fits ability, forecasts the stable property. But the predictive ability is unable in the discrete data application to achieve anticipated ideally.

Introduction

The artificial neural network is a mathematical model that simulates the biology neural network. It is the network which constitutes by many simple neurons. Between various neurons it is connected through the weight. Each neuron contains a part of network information. The neural network has parallelism, the misalignment, the fault tolerance as well as auto-adapted learning capability. It is widely used in the processing pattern recognition, approximation of function, the classification, the forecast and so on[1]. The BP neural network is one of the most widely used neural network models. It through the multi-layer forward neural network's erroneous ant propagation, promotes the widrow-Hoff study rule the neural network which produces to the multi-layered network and the misalignment differentiable excitation function, can form any continual misalignment mapping willfully approximate[2]. Compared with the traditional continuous material multi-factor analysis method, the BP neural network does not need the precise mathematical model, and it does not have any supposition request to the material itself. Its processing non-linear problem's ability is stronger than traditionally statistical method. The former massive research which are based on the BP neural network method. They had not considered that the data distributed tendency influence to its predictive ability's. They only carried on the discussion to the BP neural network application [3-4].

BP neural network principle

The BP neural network is a kind of artificial neural network and a kind of multi-layered perception. It is composed of an input level, or several concealed levels and an output level. The learning process includes signal forward-propagating and error's ant propagation. When forward-propagating, the input sample spreads from the input level, after concealment level cascade processing rear-drive to output level. If output level actual output and expectation output not symbol, then carries on error's ant propagation. Error's ant propagation is the outlet error by some form through the concealment strike input level cascade counter-biography, and shares the error for each level all units, obtains each level the error signal, this error signal is the achievement revises various units weight the basis. This kind of signal forward-propagating and the erroneous ant propagation's each weight alignment procedure carries on again and again, achieves the acceptable degree until the outlet error, or carries on the study number of times which until establishes in advance.

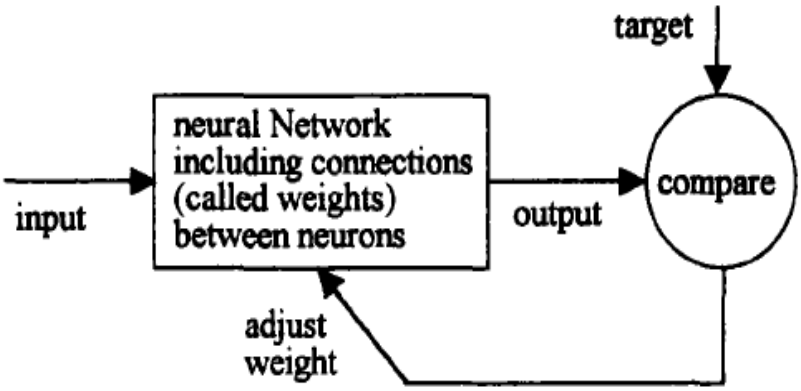


Figure 1 neural network work flow

BP neural network modeling

Example synopsis

The former literature reports that the days in hospital is the major effect factor in hospital expense. Chronic illnesses separate more discrete along with the increase of hospital days. Therefore I choose two single diseases with acute appendicitis, and diabetes was modeled [5-6]. This research material is from 2006~2009 year A group of acute appendicitis which are provided in the Hebei Province Tangshan 2 General hospital (the ICD10 code is K35) and B group of diabetes (the ICD code is E11) patients' home page material of medical record. The total cases is 2317, and the whole materials is coded by the computers, and then eliminate the cases which has the flaw value and does not conform the logic. The effective samples are 2170 (A group of 973, B group of 1197), which occupy 93.65% of the whole cases.

Table 1

Target	A（ Acute appendicitis）	B（ Diabetes）
Example number	973	1197
Mean value	6166	8179
Standard deviation	4420	11278
Coefficient of variation	71.68	137.89
Skewness	4.00	14.91
Kurtosis	28.37	294.29

Analysis method

First divides the data set, divides into the training sample collection ((to use to network training obtains network weighting factor), the confirmation sample collection (to use in preventing network to plan Taiwan excessively) and the test sample collection (uses in examining neural network forecast effect). Uses the BP neural network to carry on screening separately to two group of data variables, after screening the variable constructs BP the neural network the weighting factor expression. Inspects two models R^2, R^2_{adj} and the relative error appraises the fitting and the predictive ability and in the different sample stability.

Model comparison

Table 2 BP neural network model result various parameters are as follows.

Parameter	A group		B group	
Concealed layer	1		1	
Concealed level nerve number	35		35	
Training algorithm	LM		LM	
Iterative number of times	18		17	
Study speed	0.01		0.01	
Stop training SSE*	1.21781		0.07068	
Finally	Training set	Test set	Training set	Test set
R^2	0.79	0.72	0.87	0.52
R^2_{adj}	0.785	0.704	0.869	0.499
SSE	4.6429e+009	1.8133e+009	5.7652e+009	1.9734e+010

* The SSE for the normalized [0,1] in terms of data

Model fitting ability comparison

When establishing the BP neural network model, we can use the LM algorithm, divide the data set into the test collection and the training regulations. Two models' fitting targets are as follows:

Table 3 model fitting targets

Target	BP neural network model			
	A group		B group	
	Test set	Training set	Test set	Training set
R^2	0.72	0.79	0.52	0.87
R^2_{adj}	0.704	0.785	0.499	0.869

Compared with two kinds of model's fitting targets, regardless of group A or group B, the BP neural network model's fitting ability achieves ideally anticipated. Some scholars' research conclusion is the same [7].

Model forecast performance comparison

According to the random principle, we draw 500 samples randomized, carry on the forecast separately using the BP neural network, extract the relative error with the real value's comparison, and carry on the pair the T-test, Then we compare the forecasting qualities of the two groups of model. The result is as follows:

Table 4 homogeneous processing method group comparison

Methods		mean（$\bar{x} \pm s$）	difference	t	P
BP neural network	A group	0.52 ± 0.11	0.156	-1.34	0.181
	B group	0.364 ± 0.036			

Carries on the independent sample the T-test, may know by Table 4: The BP neural network does not have the difference during two groups, may think BP neural network when processing different dispersion pattern material the predictive ability has the good stability. But a data distribution more centralized A group finds by inspection the influential person from the fitting target to be higher than B group of predictive ability.

Conclusion

BP neural network learning process is an iterative process, the forward propagation and back propagation component. The basic idea of the algorithm is based on the output layer error, starting from the output layer, in turn, adjust the network weights and thresholds, and finally make the minimum mean square error of the output. Samples on some kind of "learning", the network is the network held the weight of adaptive learning obtained through the correct internal representation. This is the characteristics of the unknown sample data entry trained network, the network properties of the samples can be automated reasoning and recognition. Using the data of hospital expense and the influencing factor, the article has separately established the BP neural network model and multiple linear regression mold, and obtains two kinds of models' the goodness of fit targets R^2、R^2_{adj}. Comparing the model forecasting performance, we find that the stability and the predictive ability of the BP neural network performed well in the relative centralism's data set. Regarding identical sample centralized non-training sample, network, no matter what however can give the correct input - output relations ability to be called the network pan-ability. Network pan-ability relies on the network structure and the training sample characteristic. The network structure with the hidden nodal point number, the implicit strata number and the hidden node's function characteristic concerns. In general, the network hidden nodal point number and the implicit strata number are smaller; network pan-ability is stronger. But training sample's characteristic is decided by the data distributed tendency, in data set material its network pan-ability is stronger [8-9]. And because the BP neural network model does not need to consider that whether covariate is independent, whether the strain capacity satisfies normal distribution and so on, in hospital expense and continual material, I suggested that we use the BP neural network model.

References

[1] Ma Yinqi, Zhou Lifeng, Jin Peihuan. Artificial neural network application in statistics.*Medical Information* 6(11),21-24 (1998)

[2] Wang Junjie, Chen Jingwu, BP neural network theory and its applications in medical statistics in design skills. *China Health Statistics* 25(5),547-549 (2008)

[3] Gao Wei, Nie Shaofa, Shi Lvyuan. Neural Networks in Application of Survival Analysis. *Chinese Journal of Health Statistic* 23(4),76-78 (2006)

[4] Su Gaoli, Deng Fangping, Based on the MATLAB language improved algorithm of BP neural network. *Bulletin of science and technology* 19(2),130-135 (2003)

[5] Zhang Juying, Wei Jian, Yang Shuqin, Neural network model factors in the hospitalization cost analysis. *Chinese Journal of Hospital Administration* 18(3),143-145 (2002)

[6] Liu Biyao, Shen Yi, Based on BP Neural Network Modeling of hospital costs. *Zhejiang University, a master's degree thesis* in 2006.

[7] Zhou Lifeng, Gao Ersheng, Jin Pihuan. Comparison Between Neural Networks And Multivariate Linear Regression Method. *Modern Preventive Medicine* 25(3),272-274 (1998)

[8] Bartfay E, Mackillop WJ, Peter JL. Comparing the predictive value of neural network models to Logistic regression models on the risk of death for small-cell lung cancer patients. *Eur J Cancer Care (Engl)* 15(2),115-24 (2006)

[9] Fatih S. Erol, Hadi Uysal, Ucman Ergiun, et al. Prediction of Minor Head Injured Patients Using Logistic Regression and MLP Neural Network. *Journal of Medical Systems* 29(3),205-215 (2005)

doi:10.4028/www.scientific.net/AMM.50-51.982

The comparative study of workplace violence in both the state-owned hospitals and private hospitals

Liqun Yu [1,a], Xiaoxia Tang [2,b], Jianhui Wu [1,c], Shi Chen [1,d]

[1]Department of Preventive Medicine, North China Coal Medical University, Tangshan, China

[2]Department of Respiratory, Kailuan Hospital Affiliated to North China Coal Medical University, Tangshan, China

[a]liqunyu@yahoo.com.cn , [b]tangxiaoxia@yahoo.cn , [c]wujianhui555@163.com , [d]chenshonest@163.com

Keywords: state-owned hospitals; private hospitals; violence; workplace; Epidemiology

Abstract To understand the prevalence of workplace violence of the state-owned hospitals and private hospitals and to compare the characteristics of them in Tangshan city. Cluster sampling method were used to investigate the three state-owned hospitals and five private hospitals, medical staff who suffer violence in the workplace from April 2008 to April 2009 in Tangshan city. The incidence rate of workplace violence in state-owned hospitals(62.98%) was higher than private hospitals（13.55%）. The type of the violence is mainly psychological violence both in state-owned hospitals and private hospitals. And state-owned hospitals' incidence of psychological violence (46.48%) was higher than private hospitals（13.55%）. Both state-owned hospitals and private hospitals, the violence mainly occurred during daytime hours. Violence of state-owned hospital most occurred in wards（47.92%）, while private hospitals most occurred in medical rooms（32.35%）. The most important reasons for workplace violence in State-owned hospitals and private hospitals were the perpetrators of alcohol (42.81%) and patients without improvement or self-opinion that there is no improvement（57.35%）, respectively. Private Hospitals' incidence of workplace violence is lower than the state-owned hospitals' in Tangshan city, According to the characteristics of various hospitals, we should implement interventions to reduce the incidence of violence in the hospital.

Introduction

Workplace violence means employees suffered threats and harassment that are harmful to their safe and health from external or internal colleagues in their workplace[1]. Medical personnel are high-risk groups of non-lethal workplace violence[2,3]. Hospital workplace violence that seriously hampers the hospitals' normal working order and brings health services in many negative effects, has become an important public health problem[4,5,6]. This study had investigated three state-owned hospitals' and five private hospitals' medical staff from April to May 2009 in Tangshan City to understand the workplace violence' epidemiological characteristics and reasons that medical staff suffered in private hospitals, and compared with state-owned hospitals, so as to provide the basis for prevention and control of the hospital workplace violence.

Subjects and methods

Subjects

Cluster sampling method was used to investigate the three state-owned hospitals and five private hospitals' medical staff from April to May 2009 in Tangshan city. This study distributes 897 questionnaires, and returned 844 valid questionnaires, and response rate was 94.09%

Methods

Both conform to the following two kinds of cases are included in this hospital workplace violence statistics. (1)Violent incidents in line with WHO on the definition of workplace violence, which

divided into psychological violence and physical violence(psychological violence is through verbal or non-verbal infliction of pain on the medical staff; Physical violence including beating, kicking, shooting, bar, pushing, biting and other acts of violence)[7]. (2) Violent incidents happened in April 2008 ~ April 2009. According to the relevant documents of WHO and the definition of workplace violence in survey design, Included demographic characteristics, (Sex, age, education, job title); Occupation; the violence that suffered by study subjects (the violence's manner, time, place and reason). Under the guidance of the same trained investigators, study subjects self-administered the questionnaire.

Statistical Analysis

The effectiveness of the forms will be checked unified by professional, after excluding invalid forms, all data entered using the Excel2003 software and build database. Data will use SAS8.0 software to apply statistical analysis, calculating the incidence rate of workplace violence in hospital and examining by $\chi 2$ test among all groups.

Result

Basic information

This study survey 497 medical staff in state-owned hospitals, including 108 males, 389 females, sex ratio 1:3.60; and survey 347 medical staff in private hospitals, including 31 males, 316 females, sex ratio 1: 10.19, the basic information are in Table 1.

Table 1. Basic information

groups		Private hospitals		Stated-owned hospitals	
		Surveyed number	Constituent rate (%)	Surveyed number	Constituent rate (%)
Age	<30	233	67.15	235	47.28
	30～	51	14.70	153	30.78
	40～	15	4.32	73	14.69
	50～60	48	13.83	36	7.24
Education	technical secondary school and below that education	113	32.56	174	35.01
	universities and higher education	234	67.44	323	64.99
Title	free professional title	101	29.11	56	11.27
	the junior professional titles	170	48.99	226	45.47
	intermediate grade professional titles	73	21.04	167	33.60
	Deputy High professional titles and above	3	0.86	48	9.66
Sex	Male	31	8.93	108	21.73
	Female	316	91.07	389	78.27
Marital status	unmarried	183	52.74	159	31.99
	married	164	47.26	329	66.20
	Divorced	0	0	9	1.81
Type of occupation	Doctor	96	27.66	212	42.66
	Nurse	189	54.47	234	47.08
	Medical technicians stuff	62	17.87	51	10.26

Information of workplace violence in hospital

The incidence rate of workplace violence in state-owned hospitals(62.98%) was higher than private hospitals（19.60%）, showed remarkable statistical significance. The incidence rate of psychological violence and physical violence of state-owned hospitals was higher than private hospitals, both showed remarkable statistical significance, the results are in Table 2.

Table 2. Information of workplace violence in hospital

Types of violence	Private hospitals		State-owned hospitals		χ^2	P
	n	%	n	%		
psychological violence	47	13.55	231	46.48	98.4617	<0.0001
physical violence	21	6.05	82	16.50	20.3110	<0.0001
Total	68	19.60	313	62.98	155.2845	<0.0001

Characteristics of hospital workplace violence

Males (51.61%) were more vulnerable than women (16.46%) to hospital violence in private hospitals, showed remarkable statistical significance. Women that suffered the workplace violence in state-owned hospitals (62.72%) were higher than women in private hospitals (16.46%), showed remarkable statistical significance. The unmarried and married workplace violence incidence in private hospitals were 9.84% and 30.49%, both lower than that in state-owned hospitals, both showed remarkable statistical significance. The married workplace violence incidence rate(30.49％)was higher than the unmarried (9.84%) in private hospitals, showed remarkable statistical significance. In private hospitals, the highest incidence rate of violence was"40 years old" group (40.00%), followed by "50 to 60 years old" group (37.50%) group," the lowest was"<30 years old" (15.02%), showed remarkable statistical significance. In state-owned hospitals, the highest incidence rate of violence was"30 years old" group (74.51％), followed by"40 years old" group (64.38％), showed remarkable statistical significance. In each age group of State-owned hospitals' violence incidence rate were higher than private hospitals, showed remarkable statistical significance. In private hospitals the violence incidence rate of free professional title was 13.86% and 18.82% for the junior professional titles, intermediate grade and above is 28.95%, all lower than the corresponding titles of state-owned hospitals(that were 50.00%，59.73%，69.77%), showed remarkable statistical significance.

In private hospitals the violence incidence rate of doctors, nurses, medical technicians were 37.50%, 9.52% and 22.58%, all lower than the corresponding occupational categories in state-owned hospitals(that were 60.38%, 65.38%, 62.75%)，showed remarkable statistical significance.

In private hospitals ，technical secondary school and below that education the violence incidence rate were 20.35%, universities and higher education were 19.23% ，all lower than the corresponding degree that in state-owned hospitals（62.64%, 63.16%），all showed remarkable statistical significance.

Time and place that violence occurred

In private hospitals, the main sites of violence were the Medical Office Room (32.35%), outpatient (30.88%), and ward (20.06%), while in state-owned hospital the main place were wards (47.92%), nursing Office Room (27.16%), and emergency room (21.41%), the results are in Table 3.

Table 3. Time that violence occurred

Time	Private hospitals		State-owned hospital	
	Number of suffer violence	Constituent rate (%)	Number of suffer violence	Constituent rate (%)
Day shift	51	75.00	168	53.67
Night work shift	14	20.58	134	42.81
Off work	3	4.42	13	3.52

χ^2= 11.5764 P=0.0031

The reasons of violence

In private hospitals, the main reasons that victims suffered violence were " no better or patients themselves feel the disease that does not get better", "perpetrators of alcohol" and "did not meet the requirements of the perpetrators", while in state-owned hospital main reasons of violence were "the perpetrators of alcohol abuse", " no better or patients themselves feel the disease that does not get better", and "did not meet the requirements of the perpetrators", the results are in Table 4.

Table 4. The reasons of violence

The reasons of violence	Private hospitals		State-owned hospitals	
	n	%	n	%
perpetrators of alcohol	33	48.53	134	42.81
no better or patients themselves feel the disease that does not get better	39	57.35	114	36.42
did not meet the requirements of the perpetrators	30	44.12	108	34.50
High medical expenses	20	29.41	65	20.77
Attempts to restrict the behavior of the perpetrators （for example：smoking, drinking）	15	22.06	50	15.97
Perpetrators of mental disorders	11	16.18	47	15.02
Patients died	12	17.65	39	12.46
Dissatisfied with the services	16	23.53	38	12.14
Waiting too long	6	8.82	37	11.82

Conclusion

Violence is to cause injure that purpose to self-inflicted harm or does harm to others[8]. violence has become the first cause of death between 15 to 44 years old, more people brought about various kinds of non-fatal injuries by violence[9]. Many reports indicate that workplace violence of the health system is on the rise in recent years[10,11]. The state-owned hospitals has a high incidence rate of workplace violence in Tangshan, and psychological violence and physical violence were both higher than private hospitals. That female medical staff of state-owned hospitals had a high violence incidence rates to women in the private hospital medical staff, and nurses had a higher incidence rate. The most important reasons for workplace violence in private hospitals was patients without improvement or self-opinion that there is no improvement, and in State-owned hospitals the most important reasons for workplace violence was perpetrators of alcohol.

References

[1] Mayhew C., Chappell D.: Violence in the workplace. *Med J Aust* 183(7), 346-347 (2005)
[2] Ng K., Yeung J., Cheung I., et al.: Workplace violence-a survey of diagnostic radiographers working in public hospitals in Hong Kong. *J Occup Health* 51(4), 355-363 (2009)
[3] Khalil D.: Levels of violence among nurses in Cape Town public hospitals. *Nurs Forum* 44(3), 207-217 (2009)
[4] Findorff MJ., McGovern PM., Wall M., et al.: Risk factors for work related violence in a health care organization. *In Prev* 10, 296-302 (2004)
[5] Kam C., Chong V., Onch S., et al.: Workplace violence directed at nursing staff at a general hospital in southern Thailand. *J Occup Health* 50(2),201-207 (2008)
[6] Barling J., Dupré KE., Kelloway EK.: Predicting workplace aggression and violence. *Annu Rev Psychol* 60, 671-692 (2009)
[7] World Health Organization: New research shows workplace violence threatens health services. *Press Release WHO/37* (2002)

[8] Forster JA., Petty MT., Schleiger C., et al.: Know workplace violence: developing programs for managing the risk of aggression in the health care setting. *Med J Aust* 183(7), 357-361 (2005)
[9] Sevgi C., Cihad D., Fiennur D., et al. : Violence towards workers in hospital emergency services and in emergency medical care units in Samsun: an epidemiological study. *Ulus Travma Acil Cerrahi Derg* 14(3), 239-244 (2008)
[10] Child RJ., Mentes JC.: Violence against women: the phenomenon of workplace violence against nurses. *Issues Ment Health Nurs* 31(2), 89-95 (2010)
[11] Gallant-Roman MA. : Strategies and tools to reduce workplace violence. *A AOHN J* 56(11), 449-454 (2008)

doi:10.4028/www.scientific.net/AMM.50-51.987

Oxidative damage of cardiovascular system in rats following lead exposure

Licheng Yan, Yanshu Zhang, Lin Yao, and Houjun Xu

Hebei Province Key Laboratory of Occupational Health and safety for Coal Industry,

Division of Sanitary Toxicology and Sanitary Chemistry, North China Coal Medical College,

Tang Shan 063000, China

yanlicheng@126.com

Keywords: Lead; Cardiovascular toxicity; Oxidative damage

Abstract. The aim of the current study is to figure out the relationship between the oxidative damage in cardiovascular system and lead exposure. Oxidative damage was investigated in the heart and aorta of rats exposed to lead acetate for 6 weeks. Male wistar rats were randomly divided into three groups: low-dose (10 mg/kg), high-dose (40 mg/kg) lead acetate and control groups. The rats in low-, or high-dose groups were received 10 mg/kg and 40 mg/kg lead acetate by garages, and the rats in control group were administrated same amount distilled water only. We compared the three groups in terms of the level of T-SOD, NOS, and MDA in serum and heart, aorta and histological examination. We found the activity of T-SOD, NOS in rats treated with low and high doses of lead acetate were significantly lower than in the control group. The content of MDA in the low and high doses of lead acetate groups was significantly higher than in the control group. In addition, the microscopic morphology of heart and aorta in the low- and high-dose groups showed differences from the control. Oxidative damage may be involved in the process of heart and aorta dysfunction induced by lead exposure.

1 Introduction

Although blood lead levels of individual in industrialized countries have declined over the past decades, pockets of high lead exposure and widespread low-level lead exposures still persist [1]. Lead is potentially preventable exposures that may explain population variation in cardiovascular disease rates [2,3]. However, after more than 100 years since initial reports suggested a link between lead exposure and cardiovascular outcomes, the mechanism of lead to cardiovascular disease is still incompletely understood. The currently available evidence indicated that oxidative stress played a major role in atherogenesis and evolution of acute ischemia, and resultant cardiac and vascular remodeling. A line of studies showed that lead exposure has been associated with an increased incidence of clinical cardiovascular end points such as coronary heart disease and arterial disease [4,5,6]. Researches have been carried out to explore the relationship between role of ROS toxicity and heart and vascellum dysfunction [7,8]. However the role of ROS in cardiovascular damage induced by lead exposure has not been clear yet.

In present study, rats are treated with various concentrations of lead acetate for 6 weeks, and the relationship between the oxidative damage in cardiovascular system and lead exposure will be investigated.

2 Materials and Methods

Animals. Male wistar rats (n=24) purchased from Vital River laboratory, Beijing. Upon arrival, the rats were housed in a temperature-controlled, 12/12 light/dark room, and acclimated for one week prior to experimentation. They were allowed to have free access to distill water and pelleted rat chow.

Lead administration. The rats, weighting 180–200g, were randomly divided into three groups: low-dose (10 mg/kg), high-dose (40 mg/kg) lead acetate and control groups. The rats in low-, or high-dose groups were received 10 mg/kg and 40 mg/kg lead acetate by garages once a day respectively, 5 time per week, and the rats in control group was administrated same amount distilled water only. All the rats freely received the same standard laboratory feed. All animal experiments were conducted in compliance with Guide of the Care and Use of Laboratory Animals of the Research Council.

Blood collection and organ preparations. At the end of the 6-week exposure period, the rats were sacrificed to attain blood and heart, aorta sample. Several indices including the level of T-SOD, NOS, and MDA in serum and heart, aorta were measured.

Histological examination. The hearts and aorta were fixed in Bouin's fixative solution and washed with saturated lithium carbonate in 70% ethyl alcohol for 24 h to remove the excess fixatives. The tissues were dehydrated through increasing concentrations of ethyl alcohol (70%, 80%, 90%, and 100%), cleared in xylene, infiltrated in paraffin and Paraplast by an automatic tissue processor, and then embedded in the paraffin wax with an embedding machine. The tissue blocks were cut at 1μm in thickness on a rotary microtome, stained with hematoxylin and eosin (H&E), and observed under a light microscope.

Statistical analysis. Statistical analyses were carried out with SPSS software (version 13.0). We used the analysis of variance (ANOVA) and Duncan's multiple comparison to test for significant differences between groups.

3 Results

Effects of lead on body weights in rats. Rats were examined after 6 weeks of lead exposure. Body weights in the high-dose group were significant lower than those of the other groups. No statistically significant differences were found between the body weights of the low-dose and control groups.

Changes of T-SOD, NOS, and MDA in rats. Several reports have suggested that oxidative stress may be involved in lead toxicity [9,10].Table 1,2,3 showed the changes of T-SOD, NOS, and MDA in rats treated with different concentrations of lead for 6 weeks. In serum the activity of T-SOD, NOS in rats treated with low and high dose of lead acetate were significantly lower than that in the control group. Rats treated with a high dose of lead acetate showed a statistically significant decrease in the activity of T-SOD, NOS compared with that in the low dose of lead acetate. In serum the content of MDA in the low and high doses of lead acetate groups were significantly higher than in the control group. Rats treated with a high dose of lead acetate showed a statistically significant increase in the content of MDA versus the low dose of lead acetate. The same trend has been showed in heart and aorta samples (Table 1,2,3).

Table 1 Effect of lead on T-SOD, NOS, and MDA in serum($\bar{x}$ ±s, n=8)

Groups	T-SOD（U/ml）	NOS（U/ml）	MDA（nmol/ml）
Control	126.26±3.38	8.45±0.97	3.09±0.14
Low dose	96.84±10.39*	5.60±0.90*	4.07±0.43*
High dose	54.17±7.85*#	3.46±0.85*#	5.52±0.53*#

* Compared with the control group, $p<0.05$

Compared with the low-dose group, $p<0.05$

Table 2 Effect of lead on T-SOD, NOS, and MDA in heart homogenate ($\bar{x}$ ±s, n=8)

Groups	T-SOD(U/mgprot)	NOS(U/mgprot)	MDA(nmol/mgprot)
Control	63.56±7.83	6.58±1.88	5.02±0.51
Low dose	54.77±2.83*	3.76±.038*	6.93±0.34*
High dose	25.10±2.91*#	3.15±0.56*#	7.97±0.41*#

* Compared with the control group, $p<0.05$

Compared with the low-dose group, $p<0.05$

Table 3 Effect of lead on T-SOD, NOS, and MDA in aorta homogenate ($\bar{x}$ ±s, n=8)

Groups	T-SOD(U/mgprot)	NOS(U/mgprot)	MDA(nmol/mgprot)
Control	44.46±3.43	6.77±1.20	4.79±0.35
Low dose	26.71±6.51*	4.65±0.46*	5.37±0.51*
High dose	12.75±3.30*#	4.09±1.09*#	7.10±0.23*#

* Compared with the control group, $p<0.05$

Compared with the low-dose group, $p<0.05$

Histopathological results in the heart and aorta.

Microscopically, the heart and aorta of the control group showed no definite histopathological abnormalities. However, the heart of animals in the low- and high-dose group showed the pathological myocardial fiber turbid and inflammatory cell infiltration (Figs. 1). No definite histopathological abnormalities were observed in the aorta of the low- and high-dose groups (Figs. 2). In experimental animal models of atherosclerosis and hypertension, associations between oxidative stress and impaired endothelial function have been demonstrated [11].In present study, the same results didn’t be showed, and further experiment should be performed..

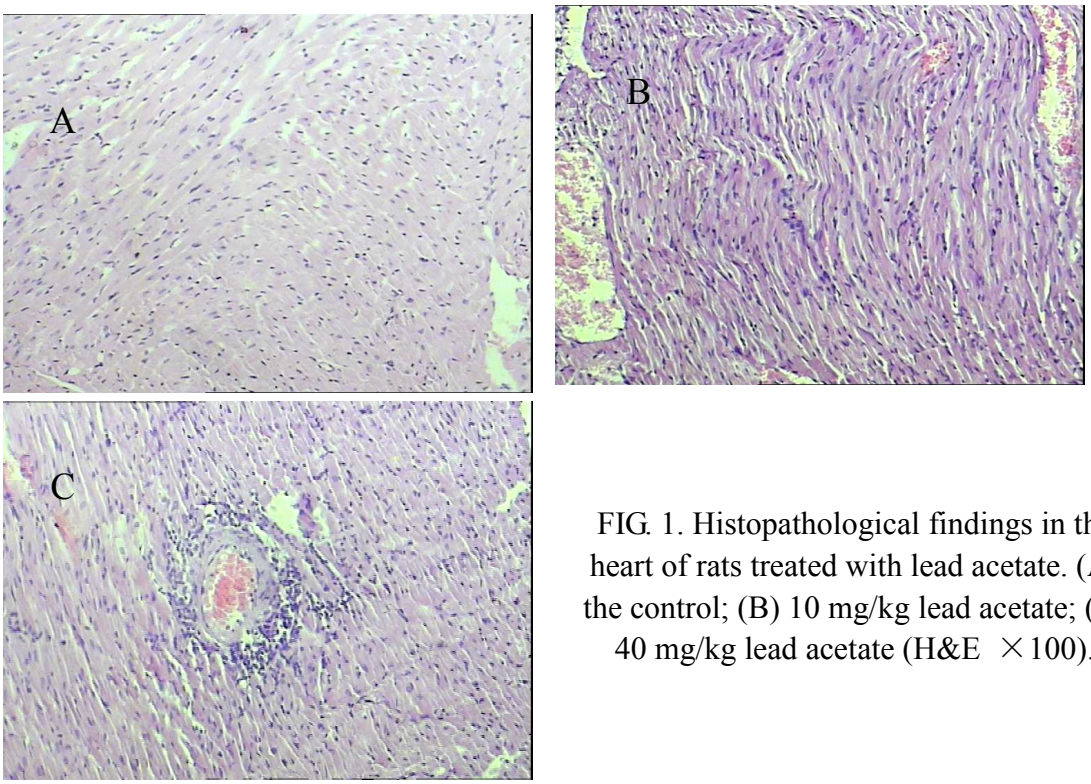


FIG. 1. Histopathological findings in the heart of rats treated with lead acetate. (A) the control; (B) 10 mg/kg lead acetate; (C) 40 mg/kg lead acetate (H&E ×100).

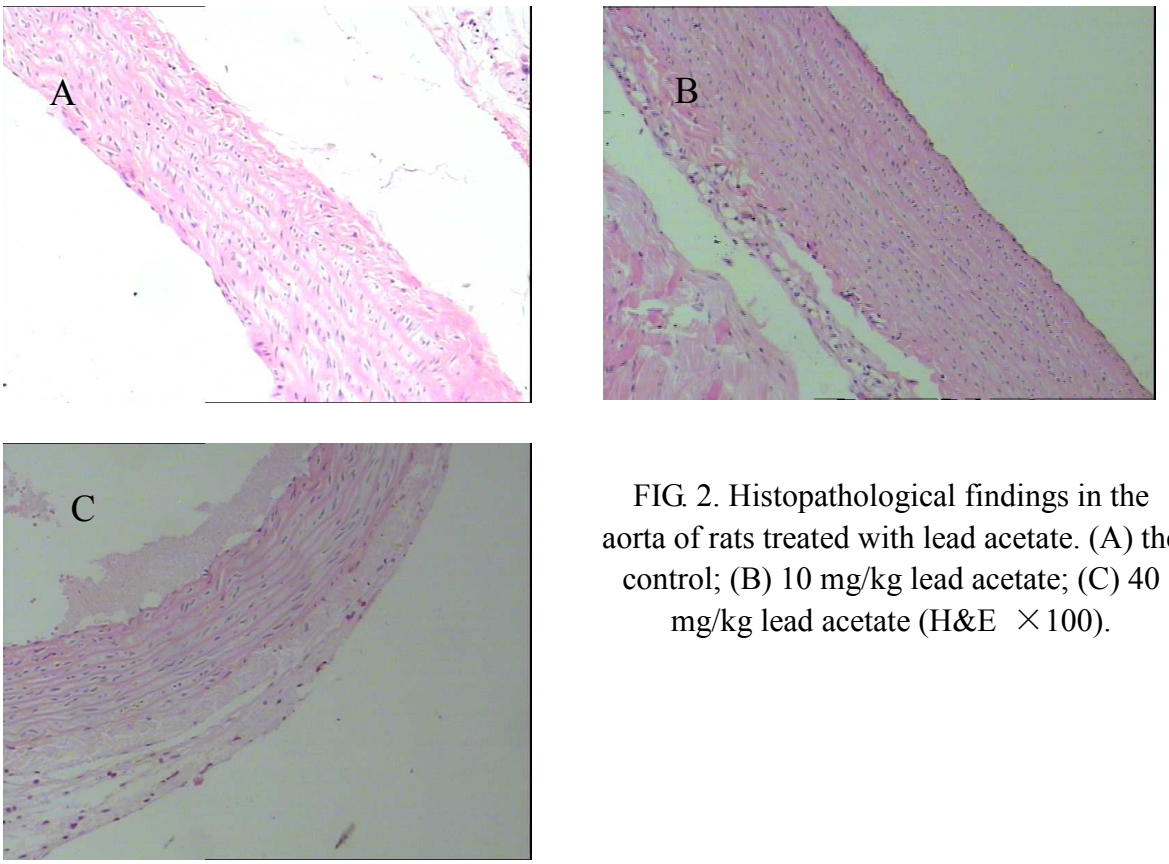


FIG. 2. Histopathological findings in the aorta of rats treated with lead acetate. (A) the control; (B) 10 mg/kg lead acetate; (C) 40 mg/kg lead acetate (H&E ×100).

4 Conclusion

In the present study, we found the activity of T-SOD, NOS in rats treated with low and high doses of lead acetate were significantly lower than the activity in the control group. The content of MDA in the low and high doses of lead acetate groups was significantly higher than in the control group. In addition, the microscopic morphology of heart and aorta in the low- and high-dose groups showed differences from the control. Oxidative damage associated with lead may be proposed as a possible mechanism of lead cardiovascular toxicity.

References

[1]. Pirkle, J.L., Brody, D.J., Gunter, E.W., et al.: The decline in blood lead levels in the United States. The National Health and Nutrition Examination Surveys (NHANES). JAMA 272(4),284–291(1994)

[2]. Bhatnagar, A.: Environmental cardiology: studying mechanistic links between pollution and heart disease. Circ Res 99(7),692–705(2006)

[3]. Weinhold B.: Environmental cardiology: getting to the heart of the matter. Environ Health Perspect 112(15),A880–887(2004)

[4]. Menke, A., Muntner, P., Batuman, V., et al.: Blood lead below 0.48 micromol/L (10 microg/dL) and mortality among US adults. Circulation. 114(13),1388-1394(2006)

[5]. Navas-Acien, A., Selvin, E., Sharrett, A.R., et al.: Lead, cadmium, smoking, and increased risk of peripheral arterial disease. Circulation 109(25),3196–3201(2004)

[6]. Schober, S.E., Mirel, L.B., Graubard, B.I., et al. Blood lead levels and death from all causes, cardiovascular disease, and cancer: results from the NHANES III Mortality Study. Environ Health Perspect 114(10),1538–1541(2006)

[7]. Sennoun, N., Meziani, F., Dessebe, O., et al.: Activated protein C improves lipopolysaccharide-induced cardiovascular dysfunction by decreasing tissular inflammation and oxidative stress. Crit Care Med 37(1),246-255(2009)

[8]. Yu, M.A., Sánchez-Lozada, L.G., Johnson, R.J., et al.: Oxidative stress with an activation of the renin-angiotensin system in human vascular endothelial cells as a novel mechanism of uric acid-induced endothelial dysfunction. J Hypertens 28(6),1234-1242(2010)

[9]. Gautam, P., Flora, S.J.: Oral supplementation of gossypin during lead exposure protects alteration in heme synthesis pathway and brain oxidative stress in rats. Nutrition 26(5),563-570(2010)

[10]. Bhatti, P., Stewart, P.A., Hutchinson, A.: Lead exposure, polymorphisms in genes related to oxidative stress, and risk of adult brain tumors. Cancer Epidemiol Biomarkers Prev 18(6),1841-1848(2009)

[11]. Cai, H., Harrison, D.G.: Endothelial dysfunction in cardiovascular diseases: the role of oxidant stress. Circ Res 87(10),840–844(2000)

doi:10.4028/www.scientific.net/AMM.50-51.992

Multi-sensor Network Design Based on Intelligent Node

Z.X. Tian[a], Z. Dong[b], G.Y. Yang[c], C.J. Chen[d]

Dalian Scientific Test and Control Technology Institute，P.O. Box 67 Dalian, 116013, PRC

[a]tianzuoxi@sina.com.cn, [b]Sonar_D@163.com.cn, [c]yanggyok@163.com.cn,

[d]ccjiu2000@yahoo.com.cn

Keywords: computer network; network design; intelligent node; DSP microprocessor; CAN-bus

Abstract. To resolve the problem of data transmitting in multi-sensor network possessing of a lot of sensors and processing sensors signals on the same time base and transmitting the results remotely, a three classes' bus topology structure based on intelligent node was designed. The system comprised scene devices, transmission cable, and console. The scene devices included several gateways, more nodes and sensor groups. The console communicated remotely with gateways via Ethernet, and gateway communicated with nodes via CAN bus, each node managed a group of sensors via RS-485 bus. The designs of intelligent node were presented in detail including hardware design, transport protocols, as well as program flow. Above conceptual design was applied successfully in a certain multi-sensor network system principle prototype consisting of a console, a gateway and three nodes possessing of ten sensors. The results prove the conceptual design feasible.

1 Introduction

As modern industry technology progresses, more sensors are employed for information collection and system control in scene. The structure and transport protocols of sensors network are differ in thousands ways up to the use demands. Traditional RS-485 field-bus can only construct master-slave system, which needs encode station address and restricts communication distance and node amount. Light network based on EPON or GPON is also used in distributed system, but high cost of light network facility restricts its application especially for the system employing a mass of nodes. CAN-bus (Controller Area Network) abolishes traditional code of station address and can work in multi-master mode, with advantages of high communication speed, easy implementation and high performance-cost ratio[1,2]. CAN-bus is a technology of serial communication network effectively supporting distributed or real time control system. CAN-bus has been applied in multitudinous fields, rapidly extending from primal motorcar industry to aircraft industry, railway transport, industry control, medical equipment and security defense, etc[3,4]. In some occasions, there are a lot of sensors, and signals are to be processed on the same time base between sensors, and the results are to be transmitted remotely. In this case, simplex structure and transport protocols are not competent. A three classes' bus topology structure based on intelligent node is designed to meet these demands in this paper. The conceptual design is presented and prove feasible through a principle prototype comprises three nodes.

2 Design of Network Topology Structure

A schematic diagram showing the network topology structure is shown in Figure 1. The system comprises scene devices, transmission cable, and console. The scene devices include several gateways, more nodes and sensor groups. A three classes' bus topology structure is designed

employs Ethernet as trunk and CAN-bus as first class branch and RS-485 as second class branch. The console communicates remotely with gateways via Ethernet, and gateway communicates with nodes via CAN bus, each node manages a group of sensors via RS-485 bus. The quantities of gateway and node, as well as sensor in group are configured according to practical instance.

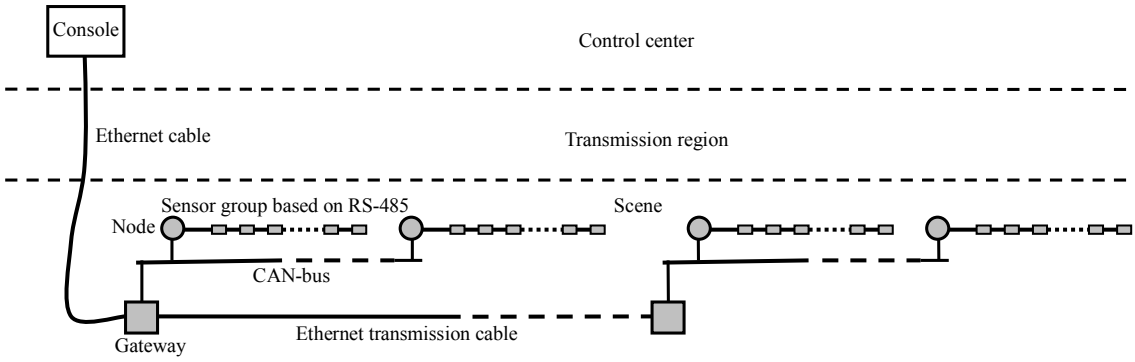


Fig. 1 Schematic diagram of the network topology structure

3 Design of Intelligent Node

3.1 Hardware Design

DSP microprocessor is kernel of the node, which samples and process signals of sensor group on the same time base via RS-485 bus, and sends results to the gateway via CAN bus when need. A schematic diagram of hardware composition and interfaces is shown in Figure 2.

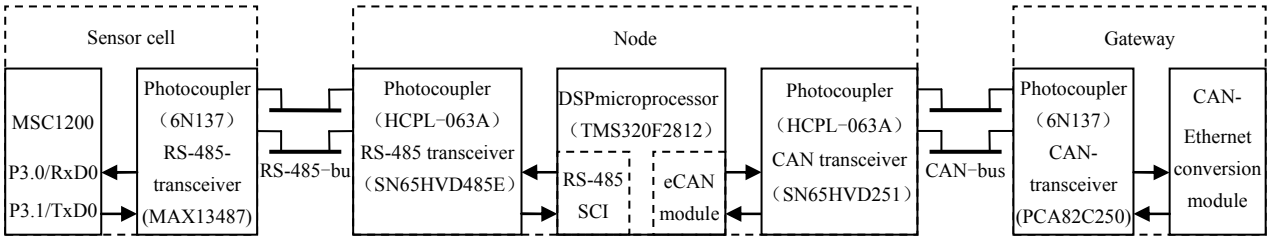


Fig. 2 Schematic diagram of hardware composition and interface of intelligent node

Microprocessor in node adopts high-performance 32-bit fixed-point DSP chip of TMS320-F2812 with 150MHz master clock frequency, 150MIPS processing capability, 128k×16bit flash devices, 18k×16bit SRAM, two serial communications interfaces (SCIs, standard UART), enhanced controller area network (eCAN), serial peripheral interface (SPI), and up to 56 general purpose I/O (GPIO) pins. MSC1200 single chip processor is kernel of the sensor cell, which integrates precision analog-to-digital converter (ADC), 8051 microcontroller and flash memory on chip. MSC1200 samples and stores sensor signals, DSP reads data of every sensor cells in turn at required rate via RS-485 bus. It is to be attended that sampling frequency of sensor cell must be at least m times DSP reading rate, as DSP must complete fetching data of m sensor cell in turn during one sampling period. Actually fetching data operation of DSP corresponds to sampling again. It is impossible to start ADC only at the moment of each fetching data operation, because ADC works unsteadily and has high noise at start moment. MSC1200 stores the data from A/D into one of two registers alternately, which ensures DSP could get the newest data at each fetching data operation. All sensor cells data in DSP can be considered synchronous, and most delay between them is less than one sampling period. DSP processes data of sensor group, and transmits the results to gateway via CAN bus. Gateway transforms CAN formant data of nodes to Ethernet formant data, and transmits them

to the console.

3.2 Transport Protocols

There are three protocols to be designed, RS-485 bus protocol between node and sensor cell, CAN bus protocol between node and gateway, and CAN-Ethernet transform protocol. One of two SCIs in TMS320F2812 can be used for RS-485 communication to external after being driven by RS-485 transceiver. The eCAN module compatible with the CAN2.0B standard in TMS320F2812 can be used for CAN communication to external after being driven by CAN transceiver. All SCI and CAN use GPIO MUX, so a external control chip is needed. A universal CAN-bus device is used to implement protocol transform from CAN to Ethernet.

▪ RS-485 Protocols

RS-485 is half-duplex asynchronous serial communications interfaces, and its transmitting and receiving use the same bus of twisted-pair in common controlled by PD01 of GPIO in TMS320F2812. RS-485 works at transmitting mode when PD01is high level, and at receiving mode when PD01is low level. TMS320F2812 communicates with each sensor cell at master-slave mode in turn according to its address. The basic unit of data is called a character and is one to eight bits in length. Each character of data is formatted with a start bit, one or two stop bits, and optional parity and address bits. A character of data with its formatting information is called a frame and is shown in Figure 3.

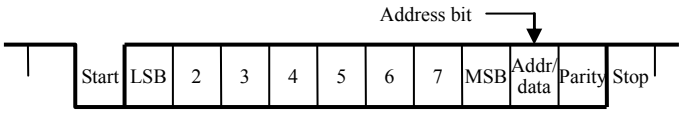


Fig. 3 RS-485 data frame

▪ CAN Protocols

Data link layer is kernel of CAN protocols, thereinto Logical Link control (LLC) accomplishes the functions of filtration, over loading informing, management resuming etc. Medium Access control (MAC) accomplishes the functions of data packing / unpacking, frame coding, medium access managing, error detecting, error message token, response, series-parallel conversion etc. All these functions are implemented via message transmitting.

There are two message frame formats in CAN2.0B protocol, difference between them is the length of identifier field, standard frame has 11 bits, and extended frame has 29 bits. Whatever frame format has four types of frame, Data frame, Remote frame, Error frame, Overload frame.

Data frame is kernel of CAN bus, through which the functions of data sharing, data exchanging and data transmitting are implemented between nodes. Standard / extended Data frame comprises seven different fields and is shown in Figure 4.

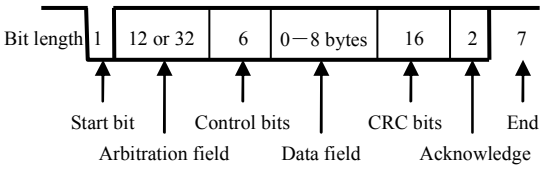


Fig. 4 CAN data frame

The function of data link layer is implemented by eCAN controller in node. The interior structure of eCAN controller integrated in TMS320F2812 is shown in Figure 5[5], what to be needed is only reading or writing the memories when TMS320F2812 communicates with exterior.

▪ CAN-Ethernet Transform Protocols

CAN−Ethernet gateway transforms data from CAN−bus to Ethernet via CAN data package mode. This mode adopts fixed 20 bytes length message format encapsulated with all CAN−bus 2.0 A / B message information to transmit data from frame to frame. The data structure is defined as follows[6]:

[1 byte]：CAN port indexing.

[4 bytes]：CAN package ID，the lower byte is anterior and the higher byte is posterior, standard frame is 11 bits and extended frame is 29 bits.

[1 byte]：data length，one to eight bytes.

[8 bytes]：CAN package data，can fill residuary bytes with arbitrary data if data length is less than eight.

[1 byte]：1 represents extended frame and 0 represents standard frame.

[1 byte]：remote transmission indicating bit, 0 represents not RTR data package and 1 represents RTR data package.

[4 bytes]：reserve.

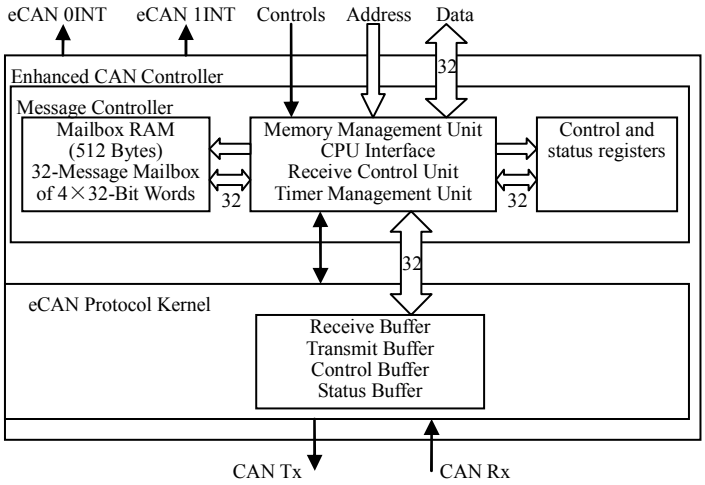


Fig. 5 Interior structure of eCAN controller

3.3 Program Flow

There are two communication modes between node and gateway, receiving mode and transmitting mode, normal mode is transmitting. Receiving mode adopts interruption request (IRQ) way, DSP responds to the IRQ of receiving message from eCAN controller, reads data of mailbox RAM, decodes and executes commands. Transmitting mode adopts state inquiring way, DSP writes information into mailbox RAM of eCAN controller to transmit when DSP finds that the process results of sensor signals need to be sent. The operations of transmitting and receiving are implemented by eCAN controller. The node fetches data of sensor cells via transmitting and receiving IRQ of RS-485 communication interface. The program flow is shown in Figure 6.

4 Application Results

Above conceptual design is applied successfully in a certain multi-sensor network system principle prototype. It consists of a console, a gateway and three nodes possessing of ten sensors. The distance between gateway and console is 2km. The sampling frequency of ADC is 30Hz in sensor cell, and the rate of DSP reading each sensor cell data is 1ps in node. Communication speed of RS-485 bus is 9.6kbps, and that of CAN bus is 50 kbps.

5 Summaries

The practical application results indicate that the multi-sensor network designed based on intelligent node and three classes' bus transport modes can synchronously meet these demands of processing signals between sensors on the same time base, possessing a lot of network nodes, and transmitting data to a long distance.

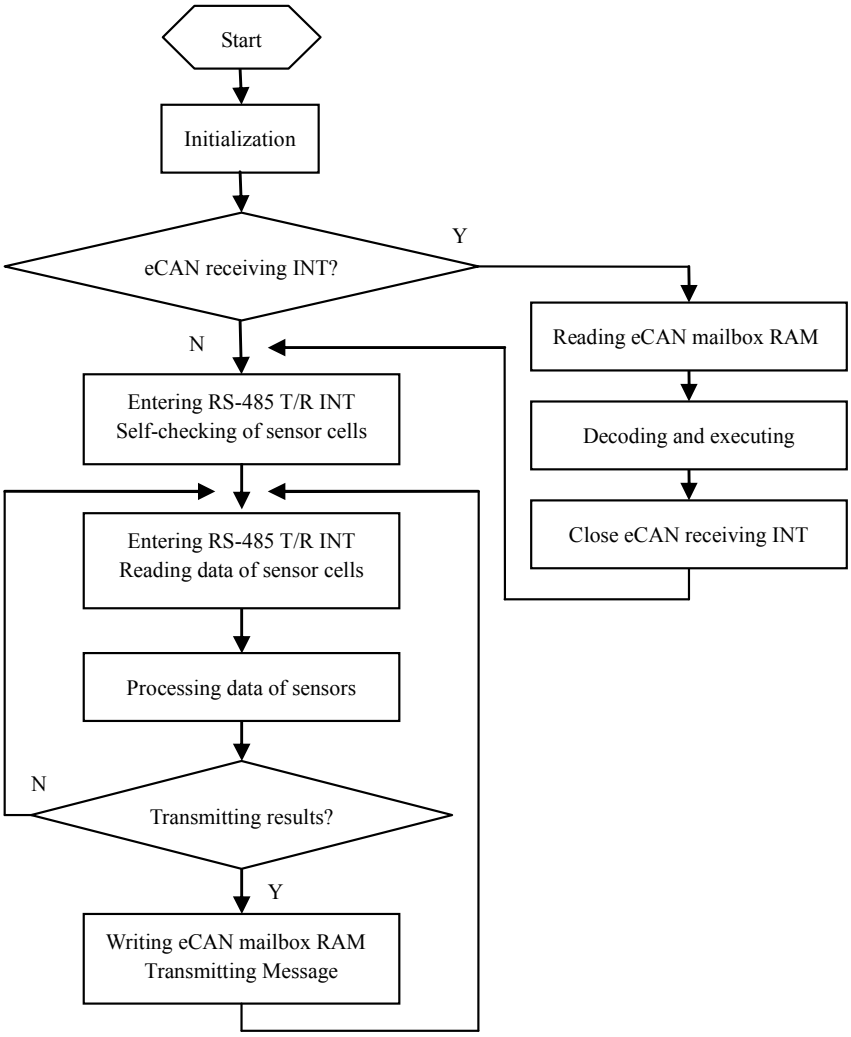


Fig. 6 Program flow of node

References

[1] Y.T. Rao, J.J. Zou and J.H. Wang: *Principle and Application of Field CAN-bus(Ver 2)*. (Beijing Aircraft and Spaceflight University Publications, PRC 2007)

[2] L.G. Zhou: *Principle and Application of Field-bus iCAN*. (Beijing Aircraft and Spaceflight University Publications, PRC 2007)

[3] W.J. Zhang, L.H. Liu: *Design of Multi-sensor Real-time Monitoring System Based on CAN Bus*. Computer Measurement & Control. Vol. 14 (6) (2006), p. 716-718

[4] B.X. Hu, F. Zhang and A.J. Liu: *Design of Gyro Dynamical Test System Based on CAN-bus and DSP*. Computer Measurement & Control. Vol. 14(12)(2006), p. 1611-1613

[5] *TMS320x281x,280x DSP Enhanced Controller Area Network (eCAN) Reference Guide*, Texas Instruments Incorporated Publishers, 2005 revise, p. 1-7

[6] *CAN-bus products technical configuration manual (Ver 2.5)*, Bei jing Kinghwa Technology, Co. LTD Publishers, p. 60-64

doi:10.4028/www.scientific.net/AMM.50-51.998

Damping Parameter Identification of Large Span Cable-stayed Bridges

Zhanyu Bu[1, a], Xu Xie[2,b] and Yong Ding[1,c]

[1]Faculty of Architecture, Civil Engineering and Environment ,Ningbo University, China

[2]College of Civil Engineering and Architecture, Zhejiang University, China

[a]buzhanyu@nbu.edu.cn, [b]xiexu@zju.edu.cn, [c]dingyong@nbu.edu.cn

Keywords: cable-stayed bridge, non-proportional damping, damping parameter identification

Abstract. The damping makeup of large span cable-stayed bridges is complicated. The damping matrices of such structures cannot be easily accurately determined. On assumption of the damping coefficients of substructures in cable-stayed bridge (such as girder, tower, cable and support et al) are constants, a method to estimate the damping coefficients and damping matrices of such structures is presented. The method is based on several known modal damping ratios and the calculation of complex eigenvalue and eigenvector utilizing the state-space methodology. Numerical simulation was carried out by the example of Tatara Bridge. Through comparison between the calculated damping ratios utilizing the approximated damping matrices and the field test data, the method is proved valid.

Introduction

The damping of cable-stayed bridge can be mainly classified into four types by their sources, namely viscous damping, Coulomb damping, structural damping and radiant damping[1,2]. The characteristic of damping of different components in cable-stayed bridges varies very much, which consists the nonlinearity of the damping of whole bridge[3].

Rayleigh had verified that if the damping matrix is the linear combination of mass matrix and stiffness matrix, then system has classical normal modes[4]. T. K. Caughey[5], T. K. Caughey and M. E. J. O'Kelly [6]studied the relation between damping matrix and mass and stiffness matrix. But in reality the damping system has complex mechanics, and generally doesn't have classical normal modes. There are mainly two methods to solve non-classical damped system problems. The first is the state-space method [7] and the second is the perturbation method[8].

In this paper, the damping characteristic of large span cable-stayed bridges is analyzed by complex eigenvalue calculation method. Then a method to get the estimation of the damping matrices according to field test data of damping ratio is proposed[9].

State-space method for non-proportional damping system

The earthquake governing equation of motion for a multi-degree-of-freedom (MDOF) structure system can be written in matrix form as

$$[M]\{\ddot{V}\}+[C]\{\dot{V}\}+[K]\{V\}=-[M]\{r\}\ddot{v}_g \tag{1}$$

where $[M]$, $[C]$ and $[K]$ are symmetric mass, damping and stiffness matrices respectively. $\{V\}$ is a vector of structural displacement responses, $\{r\}$ is modal participation vector and $\ddot{v}_g$ is a vector of the earthquake acceleration excitation.

If damping matrix $[C]$ can't be decoupled by the real normal modes which decouples the mass and stiffness matrix, it is usually difficult to solve equation (1) directly. Thus equation (1) can be deduced into single order state space, then equation (1) becomes

$$\{\dot{Z}\}-[A]\{Z\}=\{E\}\ddot{v}_g \tag{2}$$

where

$$\{E\}=\begin{bmatrix}\{0\}\\ -\{r\}\end{bmatrix} \tag{3}$$

$$\{Z\}=\begin{Bmatrix}V\\ \dot{V}\end{Bmatrix} \tag{4}$$

$$[A]=\begin{bmatrix}[0] & [I]\\ -[M]^{-1}[K] & -[M]^{-1}[C]\end{bmatrix} \tag{5}$$

Solving equation (2) one can get the modal vector $\{\Phi^{(n)}\}$ and eigenvalue $[\lambda^{(n)}]$, which can be written as[10]

$$\{\Phi^{(n)}\}=\begin{bmatrix}\{\phi^{(n)}\} & \{\overline{\phi^{(n)}}\}\\ \lambda_n\{\phi^{(n)}\} & \overline{\lambda_n}\{\overline{\phi^{(n)}}\}\end{bmatrix},\text{n=1,2,…,N} \tag{6}$$

$$[\lambda^{(n)}]=\begin{bmatrix}\lambda_n[I] & [0]\\ [0] & \overline{\lambda_n}[I]\end{bmatrix} \tag{7}$$

In which, the superscript "—"denote the conjugate of the complex number, $\{\phi^{(n)}\}$ is the n-th eigenvector of original structure, eigenvalue λ_n can be written in the form of complex,

$$\lambda_n=\alpha_n+i\beta_n \tag{8}$$

In which,

$$\alpha_n=-\xi_n\omega_n \tag{9}$$

$$\beta_n=\omega_n\sqrt{1-\xi_n^2} \tag{10}$$

Where ω_n is the natural circle frequency of vibration of the n-th mode, ξ_n is the modal damping ratio of the n-th mode, from equation (9) and (10) one can obtain

$$\omega_n=\sqrt{\alpha_n^2+\beta_n^2} \tag{11}$$

$$\xi_n=\frac{-\alpha_n}{\sqrt{\alpha_n^2+\beta_n^2}} \tag{12}$$

Damping coefficients estimation method

For system with viscous damping, the damping ratio of the n-th mode can be written in the form below,

$$\xi_n=\frac{1}{2\omega_n}\frac{\bar{\phi}_n^T C\phi_n}{\bar{\phi}_n^T M\phi_n} \tag{13}$$

In equation (13), ϕ_n is the modal shape vector of the n-th mode, $\bar{\phi}_n$ is the conjugate vector of ϕ_n. Assuming the damping coefficient of material is constant, the structure can be divided into substructures with approximate same damping. The damping matrix can be written in the form below

$$C = c_1\overline{C}_1 + c_2\overline{C}_2 + \cdots + c_m\overline{C}_m \tag{14}$$

In equation (14), $c_1, c_2, \ldots, c_m$ is the damping coefficients of component material in structure, $\overline{C}_1, \overline{C}_2, \ldots, \overline{C}_m$ is the unit damping matrix of substructure with the same material in structure, which is similar with mass matrix but substituting density ρ with damping coefficients c_m, m is the number of material. Thus, equation (13) can be written in the form below,

$$\xi_n = \frac{1}{2\omega_n} \frac{\sum_{i=1}^{m} c_i \overline{\phi}_n^T \overline{C}_i \phi_n}{\overline{\phi}_n^T M \phi_n} \tag{15}$$

The damping coefficients can be derived from the solution of the simultaneous equations below,

$$\begin{cases} \xi_{r1} = \left(\dfrac{1}{2\omega_{r1}} \dfrac{\overline{\phi}_{r1}^T \overline{C}_1 \phi_{r1}}{\overline{\phi}_{r1}^T M \phi_{r1}}\right) c_1 + \left(\dfrac{1}{2\omega_{r1}} \dfrac{\overline{\phi}_{r1}^T \overline{C}_2 \phi_{r1}}{\overline{\phi}_{r1}^T M \phi_{r1}}\right) c_2 + \cdots + \left(\dfrac{1}{2\omega_{r1}} \dfrac{\overline{\phi}_{r1}^T \overline{C}_m \phi_{r1}}{\overline{\phi}_{r1}^T M \phi_{r1}}\right) c_m \\ \xi_{r2} = \left(\dfrac{1}{2\omega_{r2}} \dfrac{\overline{\phi}_{r2}^T \overline{C}_1 \phi_{r2}}{\overline{\phi}_{r2}^T M \phi_{r2}}\right) c_1 + \left(\dfrac{1}{2\omega_{r2}} \dfrac{\overline{\phi}_{r2}^T \overline{C}_2 \phi_{r2}}{\overline{\phi}_{r2}^T M \phi_{r2}}\right) c_2 + \cdots + \left(\dfrac{1}{2\omega_{r2}} \dfrac{\overline{\phi}_{r2}^T \overline{C}_m \phi_{r2}}{\overline{\phi}_{r2}^T M \phi_{r2}}\right) c_m \\ \cdots \\ \xi_{rm} = \left(\dfrac{1}{2\omega_{rm}} \dfrac{\overline{\phi}_{rm}^T \overline{C}_1 \phi_{rm}}{\overline{\phi}_{rm}^T M \phi_{rm}}\right) c_1 + \left(\dfrac{1}{2\omega_{rm}} \dfrac{\overline{\phi}_{rm}^T \overline{C}_2 \phi_{rm}}{\overline{\phi}_{rm}^T M \phi_{rm}}\right) c_2 + \cdots + \left(\dfrac{1}{2\omega_{rm}} \dfrac{\overline{\phi}_{rm}^T \overline{C}_m \phi_{rm}}{\overline{\phi}_{rm}^T M \phi_{rm}}\right) c_m \end{cases} \tag{16}$$

In which $\xi_{r1}, \xi_{r2}, \ldots, \xi_{rm}$ is the modal damping ratio of experiment of the 1,2,…m order mode respectively. With the damping coefficients derived from equation (16) the structural damping matrix can be formed.

Identification of damping parameters of Tatara Bridge

Tatara bridge is a three span continues cable-stayed bridge measuring 1480 m in total length and 890 m in center span, linking Ikuchijima Island in Hiroshima Prefecture and Ohmishima Island in Ehime Prefecture, located almost at the center of the Honshu-Shikoku Bridge Onomichi-Imabari Route[11], structural elevation is shown in Fig. 1.

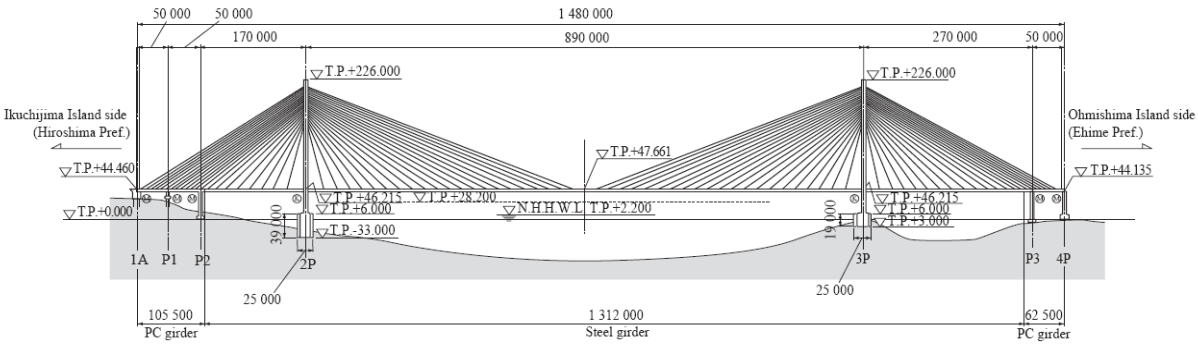


Fig. 1 Elevation of Tatara Bridge(Unit: mm)

The simulation utilizes rod system model. The model has 239 nodes, 410 elements, 2063 freedoms in total. The comparison of field test results [12]and the calculated modal frequencies are listed in Tab. 1.We can conclude from Tab. 1 the calculated results are close to that of test, which validates the correctness of the simulation model.

The theory that estimates damping coefficients and damping matrices by modal damping ratio of experimental data stated above is utilized in the damping parameter identification of Tatara Bridge.

Tab. 1 The comparison of field test results and calculated modal frequencies of Tatara Bridge

Modal description	Modal shape	Field test results（Hz）		Calculated frequencies（Hz）	Logarithm camping δ	Damping ratio ξ
		Vibration test	Micro-tremor test		Test results	Test results
Vertical bending	First symm.	0.226	0.227	0.22599	-	-
	Fist antisymm.	0.263	0.264	0.26744	0.0185	0.002944
	Second symm.	0.348	0.348	0.33626	0.0102	0.001623
Lateral bending	First symm.	0.097	0.102	0.08648	0.1235	0.019652
	First antisymm.	0.248	0.220	0.24024	0.0711	0.011315
	Second symm.	0.470	0.497	0.40484	0.0188	0.002992
Torsion	First symm.	0.497	0.497	0.58743	-	-
	First antisymm.	0.831	0.821	0.87725	-	-

The Tatara Bridge is divided into five substructures according to its composite and material, i.e. concrete girder, steel girder, steel tower, support and cable. Assuming the damping coefficients of every substructure is constant, we can calculate the undamped complex modal shape vectors of the cable-stayed bridge with equation (1) to equation(12).Then based on the given data of modal damping ratios from field test(Tab. 2), for instance the damping ratio of second mode(first symm. lateral bending),the fourth mode(first antisymm. lateral bending),the fifth mode(first antisymm. vertical bending),the sixth mode(second symm. vertical bending) and the eighth mode(second symm. lateral bending) are known, using equation (16) we can obtain the damping coefficients of the five substructures, as shown below

c_{gh} =0.135285(10^4Ns/m^2)

c_{gg} =0.0407357(10^4Ns/m^2)

c_t =0.0165237 (10^4Ns/m^2)

c_p =0.326908 (10^4Ns/m^2)

c_c =0.000430619 (10^4Ns/m^2)

where c_{gh} is damping coefficient of concrete girder, c_{gg} for steel girder, c_t for steel tower, c_p for support, c_c for cable, applying the damping coefficients above to equation (14) forming the damping matrices. Using equation (1) to equation (12), the calculated modal damping ratios are given in Tab. 2.The damping matrices distribution is given in Fig. 3.Inspection of Tab. 2 shows that although the relative error of the results is big compared with that of experiment, the real error is acceptable.

Tab. 2 The modal frequency and modal damping ratio of first iteration

Mode sequence	Modal frequency（Hz）	Modal damping ratio		
		Computed results	Test results	computed/test
1	0.032832	0.048225529	-	-
2	0.087298	0.021389021	0.019652	1.09
3	0.226644	0.008380064	-	-
4	0.240509	0.007367612	0.011315	0.65
5	0.266018	0.006928983	0.002944	2.35
6	0.333892	0.005376892	0.001623	3.31
7	0.394419	0.004521273	-	-
8	0.407168	0.004179792	0.002992	1.40
9	0.451339	0.004084927	-	-
10	0.464612	0.004156080	-	-

Conclusions

In this paper, a method of damping matrices identification of cable-stayed bridge is proposed.

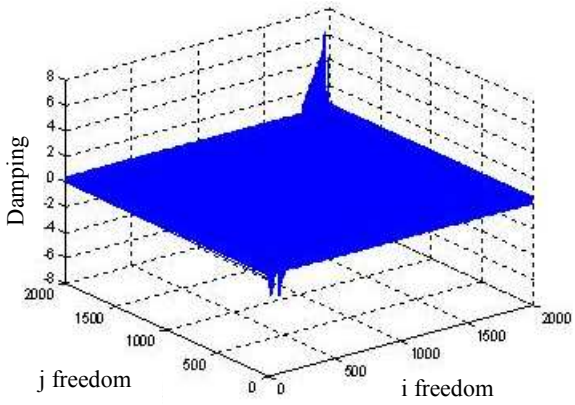


Fig. 3 The damping matrices of first iteraiton(unit:10^4Ns/m)

Cable-stayed bridges can be divided into substructures with approximately same damping property. By assuming the damping coefficients of substructures are constant, the damped frequencies and modal eigenvectors can be got by state-space complex eigenvalue calculation method suggested by Foss[7]. With the theory proposed by the author, one can obtain the approximation of damping matrix of the whole bridge. Through comparison between the calculated damping ratios utilizing the approximated damping matrices and the field test data, the method is proved valid.

This paper is sponsored by the Startup Foundation of Ningbo University (No. 2010148).

References

[1] Xie Xu. Seismic response and earthquake resistant design of bridges. China Communications Press , Beijing(2006).

[2] M. Liu and D.G. Gorman. Formulation of Rayleigh Damping and its extensions. Computer and Structures, Vol. 57(1995),No. 2,p.277-285.

[3] Yoshikazu Yamada and Kenji Kawano. Dynamic response analysis of systems with non-proportional damping. Proceedings of the 4th Japan Earthquake Engineering Symposium. Tokyo(1975), p.823-830.

[4] Lord Rayleigh. Theory of Sound. Dover Publications Inc., New York, N.Y.(1945),Vol. 1.

[5] T.K. Caughey. Classical normal modes in damped linear dynamic systems. Journal of Applied Mechanics,June,(1960),p.269-271.

[6] T.K. Caughey,M. E. J. O'Kelly. Classical normal modes in damped linear dynamic systems. Journal of Applied Mechanics, September, (1965) ,p.583-588.

[7] K.A. Foss. Co-ordinates which uncouple the equations of motion of damped linear dynamic systems. Journal of Applied Mechanics, September,(1958),p.361-364.

[8] D.L. Cronin. Eigenvalue and eigenvector determination for nonclassically damped dynamic systems. Computers & Structure. Vol.36(1990),No. 1,p. 133-138.

[9] BU Zhan-Yu. Cable-deck Coupled Vibration and Damping Characteristic Research in the Seismic Response Analysis of Cable-stayed Bridges. Doctor thesis of Zhejiang University (2005).

[10] R.M. Lin, J.Zhu. On the relationship between viscous and hysteretic damping models and the importance of correct interpretation for system identification. Journal of Sound and Vibration, 325 (2009):14–33.

[11] S. Yoshiyuki, H. Nobuyuki, M. Tomoji, A. Masatake. Erection of "Tatara Bridge".IHI Engineering Review, Vol.36,No.2(2003):p. 65-84.

[12] Y. Manabe, H. Yamaguchi, N. Sasaki. Field vibration test of the Tatara bridge. Bridge and Foundation Engineering , No.5 (1999): p. 27-30.

doi:10.4028/www.scientific.net/AMM.50-51.1003

Research on Seismic Performance of Prestressed Precast Reinforced Concrete Intelligent Structure

Jianqiang Han [1, 2, a], Zhenbao Li [1] and Xiaosheng Song [1, 2]

[1] Beijing Key Lab of Earthquake Engineering and Structural Retrofit,
Beijing University of Technology, Beijing, China 100124

[2] Earthquake Engineering Research Center of Hebei Province,
Hebei Polytechnic University, Tangshang, China 063009

[a]tshjq@139.com

Keywords: intelligent Structure, prestressed, precast, reinforced concrete, structure system

Abstract. This thesis studies deeply the crack development characteristics, failure pattern, hysteresis curve and the displacement ductility of this new prestressed precast reinforced concrete intelligent structure, by analyzing one prestressed precast reinforced concrete frame under low reversed cyclic load test. Prestressed precast reinforced concrete frame is a new assembly architecture intelligent structure. We build a model using finite element analysis software to the test piece model analysis, the analysis result agree well with the experimental results. Experimental studies indicate that this new prestressed precast reinforced concrete intelligent structure has a good seismic performance. This prestressed precast reinforced concrete frame is a new kind of structural system complying with the development of architectural industrialization, which is worthy of popularization and application in the earthquake area.

Introduction

The prestressed precast reinforced concrete frame is a new kind of structural system assembly, complying with the development of architectural industrialization. This structural system is of various advantages---- less steel consuming, lower self-weight, higher construction efficiency, better quality, lower labor consuming, energy saving, more environment friendly and higher social and economic benefits, which are in favor of sustainable social development. At present, this structure has been applied in various building constructions in Japan and the United States[1-2], having shown its unique advantages. The researches in this area in China are mainly conducted in the node specimen test[3-4], however, the overall framework of the prestressing force pilot studies are relatively rare. Our project group fabricated frame of this new node specimens and has systematically studied it[5-7], obtaining the seismic performance of this new structural system. This paper is based on the test of the overall effect of the cyclic loading on a two floors of a two-span prestressed frames fabricated with reinforced concrete frame, researching its overall structure seismic performances---- the prefabricated crack case, failure mode, ductility, energy dissipation, and stiffness degradation of the new system.

Experimental Design

Specimen design and loading device. Test specimen is fabricated with two floors and two spans of prestressed reinforced concrete frame structure. The concrete grade of frame basis is C30. The concrete grade of columns and beams is C50. The channels for Strand were reserved in the framework of prefabricated columns and beams. When the prefabricated components meet the required strength, we fabricate the frame with three steps. The first step is hoisting the prefabricated components. The second step is perfusing the beam end node joints with high strength mortar. And the final step is tensioning strand. To ensure that the concrete beam end assembly-node bound has better strength in fabricated frame beam ends, we paste two layers CFRP over the range of the beam ends. Precast

concrete specimen reinforced concrete frame design dimensions, reinforcement and the loading device are shown in Fig. 1.

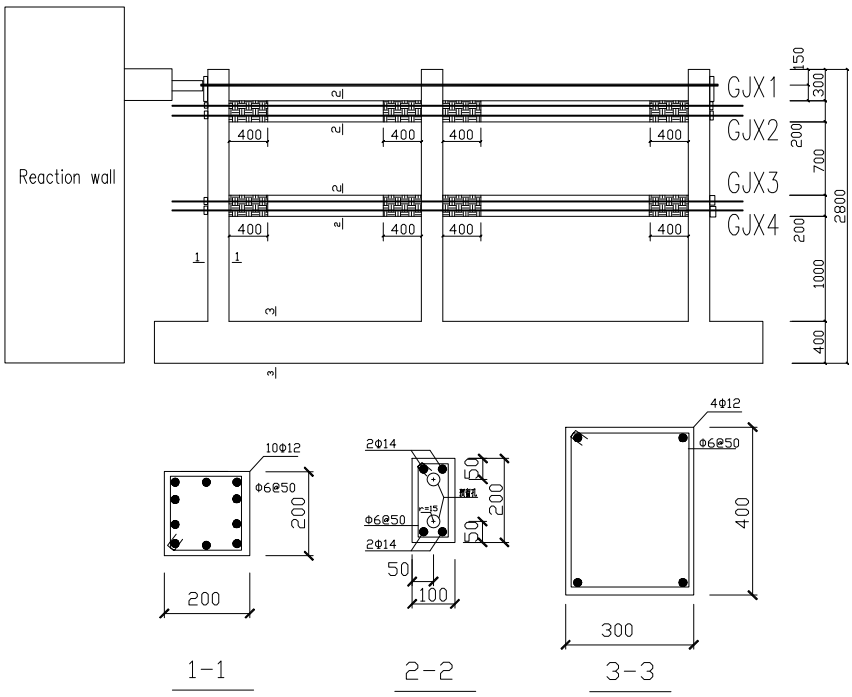


Fig. 1 Specimen dimensions

Loading program. Prestressed precast reinforced concrete frame specimen was loaded by load - displacement hybrid control method[8]. When the final bearing capacity of the specimen decreases to about 85% of the ultimate load , the specimen is regarded as damaged and the teat terminates. Test measurements included the control section main bar strain, crack and corner nodes, and the overall displacement of the framework. Loading program is shown in Fig. 2.

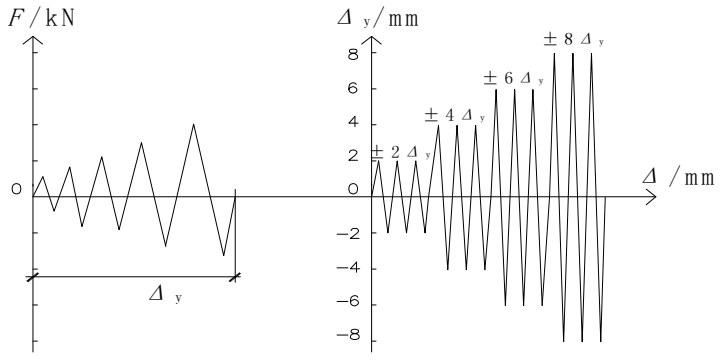


Fig. 2 Loading program

Fig. 3 Node beam damage characteristics

Test results and analysis

Failure mechanism. From the characteristics of the destruction of the prestressed precast reinforced concrete overall framework , at the beginning period of loading, there is no significant changes in the framework of beams; cracks appears later; the structure distorts little; hysteresis curve shows elasticity rule; the residual deformation of the specimen is small. With the load and horizontal displacement cycles increasing, the end of the beams appears cracks, and becomes the plastic hinges. The tangent slope of the hysteresis curve decreases obviously. When the horizontal load adds up to the limit load, the bearing capacity decreased slightly with the increase of the horizontal displacement.

The structure shows good ductility. The destruction of the prestressed precast reinforced concrete of the overall framework is characterized by the appearance of a wide crack at the end of the beam , as shown in Fig. 3. At the same time, a number of micro-column cracks appear at the root. At the later period of loading, the column root bears vertical force, and the bar yields and shows plastic hinges. Failure mechanism of this structure consistents with the principles of "strong column and weak beam".

Hysteresis curves and skeleton curves. Through data processing, the *P-Δ* hysteretic curve of the specimens is obtained, as shown in Fig. 4; the specimen skeleton curve shown in Fig. 5, where *P* is the level force applied on the top of the framework, and *Δ* is the framework top horizontal displacement.

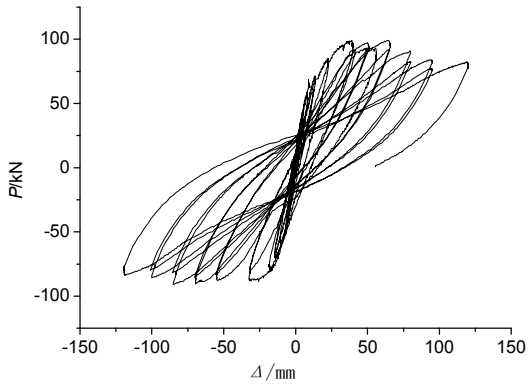


Fig. 4 P-Δ curves of specimen

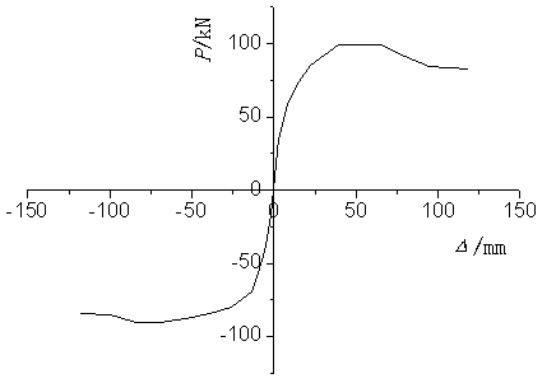


Fig. 5 The skeleton curve of the sample frame

Energy dissipation. Under cyclic loading, each unloading cycle hysteresis loop area enclosed by the structure size is equal to a cycle in which non-elastic deformation energy value. Hysteresis loop area reflects the structure of the cumulative total energy consumption of the size of the plastic. The loading process of each specimen varies during the test. Therefor, the hysteresis curve of this paper takes the envelope curve of the area surrounding the specimen to represent the value of energy dissipation . Prestressed precast reinforced concrete frame specimens of energy consumption data is shown in Tab. 1. From Tab. 1, we can conclude that the prestressed precast frame has a better energy dissipation. As long as the concrete beam end is better constrained, we can ensure that the Prestressed precast frame possesses good ductility and energy dissipation.

Tab. 1 Framework for energy-consuming specimen

Specimen	YZKJ
Load displacement	118.25
(mm)	-118.98
Energy consumption	8897.4
(Nm)	9430.3
Total (Nm)	18327.7

Displacement and ductility characteristics. The overall framework of the specimen cracking, yield, ultimate and damage control point when the characteristics of the corresponding value, shown in Tab. 2.

Tab.2 The sample control points eigenvalues of YZKJ overall framework

Loading direction	Cracking		Yield		ultimate		damage		Ductility coefficient
	Disp. /mm	Load /kN	Disp. /mm	Load /kN	Disp. /mm	Load /kN	Disp. /mm	Load /kN	
Forward	1.96	27.1	9.14	68.2	84.60	90.8	118.30	83.9	12.9
Reverse	-3.10	-25.6	-13.86	-69.8	-64.40	-99.3	-118.90	-82.4	8.5

Finite element analysis. ABAQUS finite element analysis software is used to establish the prestressed precast frame structure solid model, using the contact element and connecting the unit assembly components. Among the components, a method of the "hard contact" with friction is applied between the fabricated frame end of the beam and column nodes. The concrete and the loading constrain steel hoop on the top of the frame use three-dimensional solid element C3D8R; bar and strand tendons use space two node linear truss element T3D2; binding steel hoops and the concrete face are bound together by tie means; steel and steel wire are placed into the concrete by using embedded mode; load point role of the level of binding steel hoop centers a reference point; reference point and the load surface of the steel hoop are bound together by coupling approach. Before loading, we change the component's prestressed concrete temperature field to establish the inside pre-stress; calculation analysis is conducted with displacement control, by using multiple step command to simulate the experimental system under cyclic loading rules. During the analysing process of the model of ABAQUS, we use the actual arc length method to solve if the structure may enter the declining stage of the case. After ABAQUS finite element analysis on the results, the skeleton curve of finite element analysis results and experimental contrast curve can be obtained. as shown in Fig. 6.

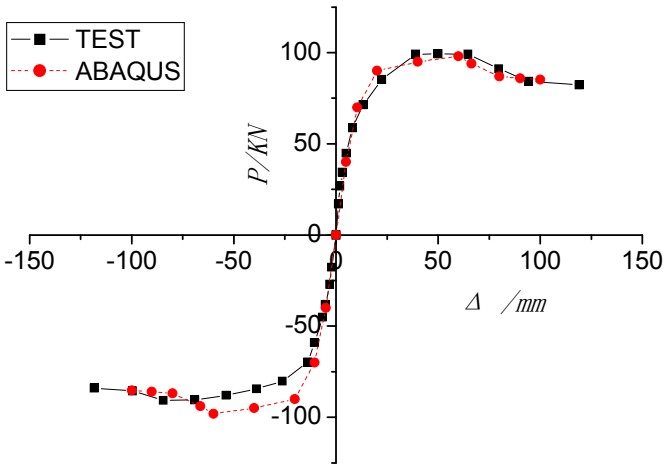


Fig. 6 Base shear force and displacement curve

Conclusions

By this assembling prestressed reinforced concrete intelligent frame structure test and finite element analysis, the following results can be concluded:

(1) This new prestressed precast reinforced concrete intelligent frame structure has a good energy dissipation capacity;

(2) The pre-fabricated steel wire compression enables to effectively improve the overall performance of the framework and significantly improve the cracking load of the structure ;

(3) hysteresis curves showed obvious mold rope phenomenon, and has good resilience and good ductility ;

(4) through the Prestressed precast reinforced concrete frame structure finite element analysis, the experimental results agree well with the theoretical results.

In summary, this new prestressed precast reinforced concrete intelligent frame structure has better seismic performance for the earthquake zone in the popularization and application of a consistent development of new industrial building structure system.

References

[1] M. J. Nigel Priestley: PCI Joumal. Vol. 47 (2002), p. 66

[2] Blandon, John J. and Mario E: PCI Joumal. Vol. 50 (2005), p. 56

[3] B. Liu, J. Tian and Y. Zhang: Journal of Building Structures (In Chinese). Vol. 28 (2007), p. 74

[4] B. Liu, Y. Zhang and Z. Jin: Journal of Building Structures(In Chinese). Vol. 6 (2005), p. 60

[5] J. Han and Z. Li: Journal of Industrial Construction(In Chinese). Vol. 39 (2009), p. 100

[6] T. Dong: Experimental and Theoretical Research on Seismic Performance of Precast Prestressed Concrete Frame Structure with Unbonded Post-tensioned Tendons[D]. Beijing University of Technology. 2006, 10.

[7] D. Wang: Research of Seismic Performance and Reinforcement Performance of Precast Prestressed Concrete Connections with Unbonded Post-Tensioned Tendons[D]. Beijing University of Technology. 2008, 06.

[8] F. Qiu, J. Qian and Z. Chen: Anti-seismic Test Method[M]. Beijing: Science Press, 2000.

©(2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.1008

Study of Urban Land-use Structure Situation on the Ground of Information Entropy

MI Cuilan [1, a], LIU Yingli [2,b]

[1] College of Science，Hebei United University，Tangshan, Hebei, China

[2] Department of Foreign Languages, Hebei United University，Tangshan, Hebei, China

[a] micuilan@163.com, [b]liuyingli2006@hotmail.com

Keywords: Urban land-use structure; Information entropy; Markov model; Principal component analysis

Abstract: In order to reveal the change law of land structure in recent years and the future land use structure in urban Tangshan, an analysis of the rules is conducted by applying such information theory as information entropy, equilibrium degree and dominance degree to nine types of land in urban Tangshan from 1998 to 2007. Then principal component analysis and SPSS are used to analyze the driving forces and the impact of various factors on land use structure. Finally, land use structure in 2012 and 2017 is predicted through Markov models.

Introduction

Along with the further discussion of the mechanism of human activities in global changes, the study of land use and its changes is experiencing a transition from global to region and from nature to humanity. Regional land use change has become a crucial part in humanistic factors leading to global environmental changes and is emphasized by many researchers[1][2]. In recent years, a multitude of fruitful researches are conducted on dynamic changes of regional land use structure from different angles, revealing the temporal-spatial differentiation characteristics of regional land use, providing reference for reasonable distribution, effective utilization and regional planning. However, the previous studies paid more attention to land use changes in quantity and quality, while information theory is seldom used to reveal the entropy feature and evolving rules of regional land use structure. This paper analyzes the land use situation in urban Tangshan from 1998 to 2007, in which information entropy is used to illustrate the orderliness of Tangshan land use structure. It also analyzes the major driving forces in land use structure changes, therefore it provides data reference for regional land use structure optimal allocation.

Land composition in Tangshan and its information entropy

With dissipative structure characteristics, urban land use system is a open, complicated system made up of nature, human, society, economy and technology. The spatial structure characteristics of urban land use can be expressed by information entropy:

$$H = -\sum_{i=1}^{n} P_i \ln(P_i). \tag{1}$$

In that formula, $p_i = A_i / A$, $A = \sum_{i=1}^{m} A_i$. A refers to the total area of a city's construction land. A_i is the area of different types of land use. According to the classification system of urban land use in "the standard of urban land classification and construction land plan" released by Ministry of Construction in 1991, urban land use can be classified into 9 types: residential land, public facilities, industrial land, storage land, traffic land, roads and squares, municipal land, grass land and special land.

Entropy reveals the number of regional land use types and the homogeneous degree of their distribution. The bigger the entropy is， the more land use types; the smaller the difference between

land use types, the more balanced. Entropy peaks when different land use area equals to each other, ie $P_1 = P_2 = ... = P_n = 1/n$, indicating an equilibrium state in regional land use.

Based on the information entropy formula，equilibrium degree E and dominance degree D in urban land use structure is,

$$E = -\sum_{i=1}^{m} \frac{P_i \ln P_i}{\ln m}, \quad D = 1 - E. \tag{2}$$

Information entropy reveals the diversity and complexity of urban land use; while equilibrium degree and dominance degree describe the area differences among land use types and the structural layout of various function types.

Land use structure in Tangshan manifests as higher percentage of industrial land，moderate percentage of roads and squares and lower percentage of residential and grass land[5].
Composition of urban construction land in Tangshan：Residential land , Public facilities, Industrial land, Storage land, Traffic land, Roads squares, Municipal land, Grass land, Special land.The change in entropy and equilibrium degree of construction land in urban Tangshan can be obtained with the help of formulae (1)and(2) based on the nine types of land in urban Tangshan from 1998 to 2007.

Table 1 Entropy and equilibrium degree change of construction land in urban Tangshan

year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Information entropy	1.955	1.977	1.981	1.976	2.011	2.02	2.027	2.011	2.006	2.004
Equilibrium degree	0.889	0.899	0.902	0.899	0.919	0.919	0.923	0.915	0.913	0.912
Dominance degree	0.111	0.101	0.098	0.101	0.081	0.081	0.077	0.085	0.087	0.088

From table 1, structure change of urban land use in Tangshan experienced two obvious phases: drastic change stage（1998-2004） sees an apparent difference in the increase of various functional land in urban construction and exploitation. The most dramatic changes are found in construction land and public facilities, while grass land, traffic land, storage land experience the smallest change. There is an obvious increase in construction land, public facilities and residential land during stable growth stage(2005-2007）.

From table 1，the information entropy of Tangshan land use first increased and then decreased, with a general increase of 0.12% , at he same time H averaged 1.995. The change in information entropy acknowledges the gradual adjustment process of order-disorder -order in urban land use system. With the release and implementation of the new “land law”, methods of land management are emphasized and strengthened, land use structure is becoming in order；after 1997, the disorder degree of urban land use structure increased due to the upsurge of the " development zone " and the " real estate ”and peaked in 2000; after 2000, with the strengthening of illegal land use punishment and the regulation of land use, the change in urban land use structure slowed down, information entropy reduced and order degree increased.

Analysis of driving forces in land use structure changes

1. The selection of evaluation index. Considering the principal of dominance，coverage and representativeness in selecting indexes and the actual situation in Tangshan, 11 indexes are chosen as target index to analyze land use structure in Tangshan: they are X_1 urban population(ten thousand)，X_2 urban social fixed assets investment(ten thousand yuan), X_3 per capita living space (m2), X_4tertiary industry proportion (%),X_5 gross industrial output (ten thousand yuan), X_6 estate development (ten thousand yuan), X_7 urban gdp(ten thousand yuan), X_8 urban finance(ten thousand yuan), X_9 urban per capita disposable income (yuan), X_{10} actual use of foreign fund (ten thousand dollar), X_{11}total retail sales (ten thousand yuan). Covering urban population, urban social structure, urban economic level, urban industrial structure,etc, these indexes are highly representative and dominant 。

2. The extraction and features of driving factors.

With the help of statistical software SPSS[6], principal component analysis is used to analyze the driving factors of land use structure changes in Tangshan from 1998 to 2007 by making X_1-X_{11} independent variables and y dependent variables. Major procedures are:

1) the extraction and standardization of original data
2) calculation of correlation coefficient matrix (table 2)
3) calculation of eigenvalues (table 3)
4) calculation of component Matrix (table 4)

the correlation degree is different among all the 11 driving factors which are affecting land use structure changes in Tangshan. The correlation coefficient is shown in table 2 .

Table 2 Correlation coefficient matrix of driving factors

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
X1	1										
X2	0.659	1									
X3	0.875	0.898	1								
X4	0.345	0.76	0.624	1							
X5	0.8	0.915	0.936	0.543	1						
X6	0.816	0.887	0.916	0.542	0.972	1					
X7	0.847	0.729	0.88	0.319	0.926	0.907	1				
X8	0.583	0.97	0.846	0.652	0.906	0.878	0.734	1			
X9	-0.359	-0.192	-0.298	-0.471	-0.257	-0.355	-0.258	-0.031	1		
X10	0.513	0.911	0.744	0.522	0.836	0.829	0.647	0.969	0.089	1	
X11	0.823	0.965	0.972	0.703	0.96	0.938	0.849	0.916	-0.292	0.83	1

There is a great correlation degree in X_2:X_8, X_5:X_6 , X_8:X_{10} and X_{11}:X_5 with correlation coefficient respectively 0.970,0.972,0.969 and 0.972, indicating the necessity of principal component analysis on driving factors. In addition, 2 driving factors are extracted based on the principle of characteristic root>1.0. From table 4, eigenvalue of F1 is 8.293 and contribution rate is 75.394. Urban social fixed assets investment(X_2), per capita living space (X_3), gross industrial output (X_5), estate development (X_6), urban finance(X_8) and total retail sales (X_{11}) have the highest positive load on the first principal component (F_1). There is a positive correlation between them and F_1. Among the seven variables highly related with F_1, total retail sales (X_{11}) holds the highest load, therefore F_1 can be viewed as the representative of retail factors and is named retail factor.

Eigenvalue and contribution rate of F_2 is 1.283 and 11.663. Urban per capita disposable income (X_9) has the highest, greatest positive load on F_2 and has a positive correlation withF_2, therefore F_2 can be viewed as the representative of urban per capita disposable income and named urban per capita disposable income. The accumulative contribution rate of the 1st and 2nd principal component reaches 87.057 % and was completely accorded with the demands of analysis.

In sum, influencing factors of Tangshan land structure change can be divided into 2 groups: 1) the first principal component plays a leading role because of the comprehensive and concentrated information it reveals. There is a higher positive correlation among such factors as urban social fixed assets investment, per capita living space, gross industrial output, estate development, urban finance and total retail sales reflecting a close relationship between urban land structure change and the city's welfare, social structure adjustment and economic development level. In the year 2007 urban social fixed assets investment in Tangshan is 60113.12 million yuan, per capita living space reaches 22.44 m2, gross industrial output is 229621.29 million yuan and investment in estate development is 4012.66 million yuan. Financial income and total retail sales experience an obvious increase than those in the previous year, promoting the demands of construction land resulting in the rapid increase in the amount of construction land and extending the urban area. 2) The second principal component has a correlation with such factors as urban per capita disposable income which reflects the fact that

construction land structure is under the influence of per capita disposable income after urban industrial structure adjustment. (see table 3, table 4).

Table 3 Eigenvalue and principal component contribution rate

Principal component	eigenvalue	Contribution rate	accumulative contribution rate
F_1	8.293	75.394	75.394
F_2	1.283	11.663	87.057

Table4 Loading matrix of principal component

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
F_1	0.814	0.96	0.965	0.664	0.978	0.969	0.876	0.93	-0.294	0.856	0.993
F_2	-0.236	0.124	-0.071	-0.266	0.038	-0.046	-0.05	0.303	0.921	0.436	-0.02

3. Multiple linear regression model of driving factors. By making land use structure Y dependent variable, principal component factors (F_1,F_2) independent variables, multiple linear regression model of urban construction area change in Tangshan from 1998 to 2007 and the driving factors can be obtained through multiple linear regression analysis.

$$Y = 0.413F_1 + 0.838F_2. \tag{3}$$

From the determination coefficient of the regression equation $R^2 = 0.624$, the two factors can explain 62.4% of land use change, indicating their great influence on land use in Tangshan. The results of the regression test indicate an dissatisfactory linear fit outcome, yet they share the same trend. The regression equation can still express their approximate interactive relations .When the two components increase a single unit, land use structure changes 0.413 and 0.838 standard unit.

Markov prediction of urban land use structure

Land use change can be seen as a random process，affecting by many different random factors. Discretize time as t=0,1,2...，for every t, suppose r(t) is the transfer rate of land use type of a set period of time，random variable X_i represents the state of the system; the probability of X_i=i，i=1,2,...k is $a_i(t)$ namely state probability; The probability from X_i=t to X_{i+1}=j is p_{ii} ie transfer probability. If the value of X_{i+1} is only determined by the value of X_i and transfer probability，with no relation with the value of X_{i-1} ,X_{i-2} .....then the random transfer process of discrete state by time is called Markov chain:

$$a_i(t+1) = \sum_{j=1}^{k} r(t) a_j(t) p_{ji}(t), \quad i = 1,2,...,k. \tag{4}$$

Satisfy

$$\sum_{j=1}^{k} a_i(t) = 1, \quad p_{ij}(t) \geq 0, \quad \sum_{j=1}^{k} p_{ij} = 1, \quad i,j = 0,1,2,...,k; \quad t = 0,1,2,... \tag{5}$$

Land use situation in 2000 and 2005, future land use change trend through Markov model is illustrated in table 5.

Table5 Actual value in 2002 and 2007 and predictive value in 2002 and 2017 of the corresponding factors

Land use type	actual area in 2002	actual area in 2007	predictive area in 2012	predictive area in2017
residential land	890.2	782.2	857.3354	820.3641
public facilities	545.8	378.9	548.378	560.4578
industrial land	817.4	819.4	633.6785	603.2146
storage land	134.8	167.7	136.8467	124.4568
traffic land	318.4	428.5	500.4682	498.6213
roads and squares	319.6	211.1	382.4796	390.2456
municipal land	129	130.6	154.6564	167.2321
grass land	191	223.1	265.4312	346.4486
special land	527.6	732.6	452.3614	440.1248

The results show that land use types will experience greater changes in the next several years. Residential land area will continue to shrink，but with a decreased range. Traffic land, grass land roads and squares will experience a greater increase. All the other types of land will either increase or decrease, with an overall tendency of moderate speed.

From table 5, land use structure is shown in the following table,

Table 6 Land use structure in Tangshan in 2012 and 2017 (%)

Land type	residential land	public facilities	industrial land	storage land	traffic land	roads & squares	municipal land	grass land	special land
2012	22.13	14.16	16.36	3.5	11.9	9.9	4	6.85	11.2
2017	21.18	14.47	15.57	3.2	12.9	10.07	4.32	7.93	10.36

1) Urban land use structure in Tangshan will be more orderly in 2010.

2) There is drastic change in urban land use structure from 1998-2007 with a general feature of increased land for public facility, roads and squares. Land use structure changes faster in the first 5 years than in the latter 5 years.

3) Urban land use in Tangshan changes due to the change of driving forces which are caused by such factors as population, economy and policy. Therefore, the key to reveal the source of land use change is to grasp driving factors system. In the process of driving forces analysis, special attention should be paid to the interaction of the inner system and the outer system. Avoid analyzing only from the aspect of time change.

4) The major factors affecting land use change are population increase, economic development and policy alteration. Among the factors, there is a negative correlation between the increase of population and land use. Policy change poses a dramatic impact on the tendency of land use change. Economic development and land interact with each other. In the city of Tangshan, population increase, economic development and the increase of residents' living level will definitely result in the change in land use styles，while technical level enhancement reinforces land use and promotes land use efficiency.

References

[1] Chen Baiming. Studies on Land Use and Land Cover Change in China and Man's Driving Force upon it. Natural Resources, 1997

[2] Wang Xiulan, Bao Yuhai. Study on the Methods of Land Use Dynamic Change Research . Progress in Geography, 1999

[3] Tan Minghong, Li Xiubin, Lv Changhe. An Analysis of Driving Forces of Urban Land Expansion in China[J], Economic Geography, 2003

[4] Liu Jian, Huang Xian-jin, Zhao Cai-yan . An Analysis of Structure Effect of Urban Land Use Based on Datum Envelopment Analysis (DEA) Model[J]. ACTA AGRICULTURAE UNIVERSITATIS JIANGXIENSIS, 2005

[5] TANGSHAN STATISTICAL YEARBOOK, 2003-2007

[6] Guo Zhigang, Method for Social-Economic Statistic--How to Use SPSS. Beijing, China Renmin University Press, 1999

doi:10.4028/www.scientific.net/AMM.50-51.1013

Statistical analysis of reliability field data in CNC system

LI Ting[1,a], ZHONG Yuning[1.2,b], YOU Dazhang[1.3,c]

[1] HuBei University of Technology, Mechanical Engineering, P.R. China.WuHan 430068,

[2] HuBei key lab of Modern Manufacture Quality Engineering, P.R. China.430068,

[3] National Numerical Control System Engineering Research Center, Huazhong of University Science and Technology, P.R. China. Wuhan 430074,

[a]lacylmc@163.com,[b]ld_zhong@sina.com,[c]youdazhang@sina.com

Key words: CNC system, reliability field data, statistical analysis

Abstract. Based on reliability field data of failure in the CNC system (Computer numerical control), failure spots, failure patterns and failure mechanism of the key parts in the CNC system was analyzed. According to the field data of reliability, the key subsystems of frequent failure were discovered through statistical analysis and criticality analysis. The fault tree was established, then the weak part of reliability was point out, and the basis for its reliability improvement was provided.

Introductions

Into the 21st century, with the rapid development of electronic technology and computer technology and automatic control technology, as the growing demand of the modern industrial production automation, servo control technology in many high-tech fields were used more and more widely. The CNC system and industrial robotics has been widely used and the servo system was the primary component of them. The high performance and technical requirements of modern CNC system were proposed. Therefore, strengthening of high-performance, high-quality servo system, analyzing fault factors which affecting the quality of reliability, improving requirements of CNC machine tools and drivers control, so it has extremely important practical significance and practical value.

General overview

Failure Statistical analysis of the CNC system is very important to improve reliability and the quality of important link of the CNC system .Through statistics analysis, the weak part would be exposed in the servo system that in design, manufacturing, application and maintenance and management phase etc. By searching for the failure reasons, effective maintenance plan and improvement measures were put forward to improve the quality and reliability.

For the whole system, detailed analysis of the failure spots, fault model etc was carried out to identify key parts of machine and its failure mode, the overall macro-control of the CNC system failure conditions. For the subsystem, study of the failure patterns and cause in-depth, summarize reliability model, the key reasons affect the reliability of the quality of the system was found out, the direction of reliability improvement was determined.

Failure Statistical Analysis of the whole system

Statistical analysis of failure spots and failure patterns of key part

For the needs of research, ,the CNC system was divided into several major subsystems or components, statistics the frequency of failure spots and failure patterns of key part respectively, the corresponding fault patterns was summed up .The frequency of the CNC system's failure spots in Table 1and its frequency polygon in Figure 1

WuHan Science and Technology Bureau.No:200951830559

Table 1 Frequency table of failure spots of key part

number	Failure spot	frequencies	frequency
A1	Motor drive system unit	13	37.1
A2	Mechanical transmission parts	7	20.1
A3	Mechanical implementation	6	17.1
A4	parts	4	11.4
A5	Detection feedback link	3	8.6
A6	Drive components	2	5.7
	Others	35	100
	Summation		

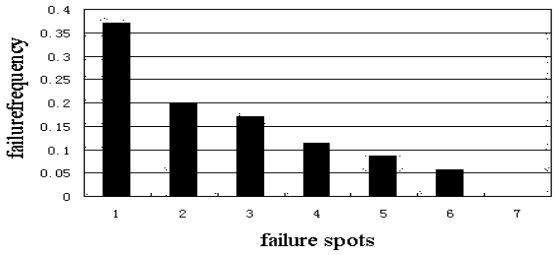


Figure 1 Frequency polygon of failure spots

The frequency of the CNC system's failure patterns in Table2.

Table 2 Frequency polygon of failure patterns

Number	Failure patterns	frequencies	frequency
FM1	Component damage	6	17.1
FM2	Dislocation\translocation inaccurate	5	14.3
FM3	Circuitry\cable abruption	5	14.3
FM4	Labor precision out of limit	4	11.4
FM5	Position error excessive	4	11.4
FM6	Component loose	3	8.6
FM7	Return to zero inaccurate	3	8.6
FM8	No display screen	2	5.7
FM9	Fuse broken	2	5.7
FM10	Uncontrolled	1	2.9
	Summation	35	100

The frequency table and frequency polygon of failure key parts and failure patterns in CNC system could be seeing:

(1) Motor drive control unit (37.1%) is the most frequent position, its failure rate than other parts or subsystems. Major failure performance: The primary device damage, the abruption, the wire fall off and so on; Mechanical transmission parts (20.1%) is also more, major failure performance: dislocation or translocation inaccurate, X axis clearance is too big, etc.

(2)The most frequent failure mode in the CNC system is the component damage（17.1%), dislocation or translocation inaccurate (14.3%), Circuitry\cable abruption (14.3%), etc.

Critically analysis

Critically analysis is quantitative analysis on failure effect, by studying the comprehensive influence according to the serious level of each failure mode and criticality or the occurrence probability. We can integrate evaluation the criticality of each position in the round, reaction influence of function and the surrounding environment and safety of whole machine when subsystem or components break down.

Through analyzing the harmfulness influence, we can obtain the key pattern of reliability in CNC system, find out the weak link of product, to make improvement design of reliability directly.

The failure mode j cause the harmfulness of failure mode on Component i, recorded as CRij. Its calculation formula is:

$$CR_{ij} = \alpha_{ij}\beta_{ij}\lambda_{ij} \quad (1)$$

In this formula: αij is frequency ratio of component i which break down by failure mode j. Its calculation formula is:

$$\alpha_{ij} = \frac{n_j}{n_i} \quad (2)$$

In this formula: ni is the total frequency of which the component i failure pattern occurs completely; nj is frequency of which the component i break down by failure mode j.

In this formula: β_{ij} is the conditional probability of the loss of function. When the probability is unable to determine, it can be divided into four uniforms, as shown in Table 3

Table 3 β_{ij} 's value and its corresponding significance

β_{ij}	explanation
1.00	That certainly can lead to component failure
0.50	That could lead to parts failure
0.10	The possible led to component failure is relatively small
0.00	That lead to component failure impossibility

In this formula: λ is the basic failure rate of component I, and it was obtained by average failure rate of the field data. Its calculation formula is:

$$\overline{\lambda_i} = \frac{N_i}{T} \quad (3)$$

In this formula: Ni is the failure times of component i during the examination. T is the commutative times of component in the total working hours during the examination.

The CNC system is divided into five main subsystems, and their criticality respectively for motor driving control system CR1, mechanical transmission system CR2 and mechanical implementation system CR3, drive component CR4 detection feedback system CR5. The accumulation working time of 35 sets series CNC machine for 10303 hours. According to the formula (3),it can obtain below:

{λi }={0.00126; 0.00068; 0.00058; 0.00039; 0.00029}

According to relevant statistics and formulas, it could obtain the following data as shown in Table 4

Table 4 FMECA tables of servo system

subsystem	Failure mode	α	β	λ_i	CR_i
Motor drive system	Component damage	0.385	1	0.00126	0.0009702
	Circuitry\cable abruption	0.231	1		
	Fuse broken	0.154	0.5		
	Position error excessive	0.154	0.5		
	uncontrolled	0.077	0.1		
Mechanical transmission parts	Dislocation\translocation inaccurate	0.428	1	0.00068	0.0004077
	Component loose	0.286	0.5		
	Labor precision out of limit	0.286	0.1		
Mechanical implementation parts	Dislocation\translocation inaccurate	0.286	1	0.00058	0.0003401
	Return to zero inaccurate	0.286	0.5		
	Position error excessive	0.286	0.5		
	Labor precision out of limit	0.143	0.1		
Detection feedback system	Component damage	0.250	1	0.00039	0.0002438
	Circuitry\cable abruption	0.500	0.5		
	Component loose	0.250	0.5		
Drive components	Component damage	0.333	1	0.00029	0.0001545
	Return to zero inaccurate	0.333	0.1		
	No display screen	0.333	0.5		

From the above table data can be seen, the motor drive system, the highest criticality, and it is key components affecting reliability of the series domestic CNC machine, followed by mechanical transmission system and the implementation system. Therefore, for the domestic servo system, motor drive system is the weakest link. We should focus on improving the design and analysis of reliability.

Failure Analysis of key subsystems

Failure mechanism analysis

Through the above analysis, we can find out that motor drive control system is the subsystem of the most failure frequent occurrence, and the most criticality effect. For the failure mechanism of the system, we could analyze two aspects, servo motor body and its drive control.

(1) Servo motor the servo motor's function is implement transformation between the electrical energy and the mechanical energy. Its running is related to the load and environmental factors, also many factors involved, such as electrical systems, insulation system, mechanical system, ventilation system . Any system with bad or between failures will lead to motor break down. Therefore its failure mechanism is complicated relatively. In general, the motor failure includes mechanical failure and electrical failure: the main failure of mechanical aspects is bearings overheating, mechanical vibration, the sound operation abnormal; the main failure of electrical aspects is winding circuit, short circuit, poor contact.

(2)Drive control Motor and drive control are closely interactive. According to the related documents about failure mode analysis, the Failure in drive control includes current over-current, module failure, over voltage, power management burn, integrated circuits short, etc.

Fault Tree Analysis—FTA

Fault Tree Analysis—FTA, is the effective means for statistical analysis of reliability and safety analysis. FTA principle, to analysis hardware fault, environmental factors and human factors, draw the logic diagram of event causation, in order to determine the top event failure occurs and the probability of various possible causes. Using FTA technology can find out all failure composition reasons result top event happen, identify a potential failure, then it is beneficial to make improvements in the design stage and improve product quality. By the way of this logic, FTA technology has become a kind of engineering and management in the extremely useful tool. According to the above analysis of motor drive system structure composition and failure mechanism analysis .The fault tree of motor drive system model was shown in Figure2.

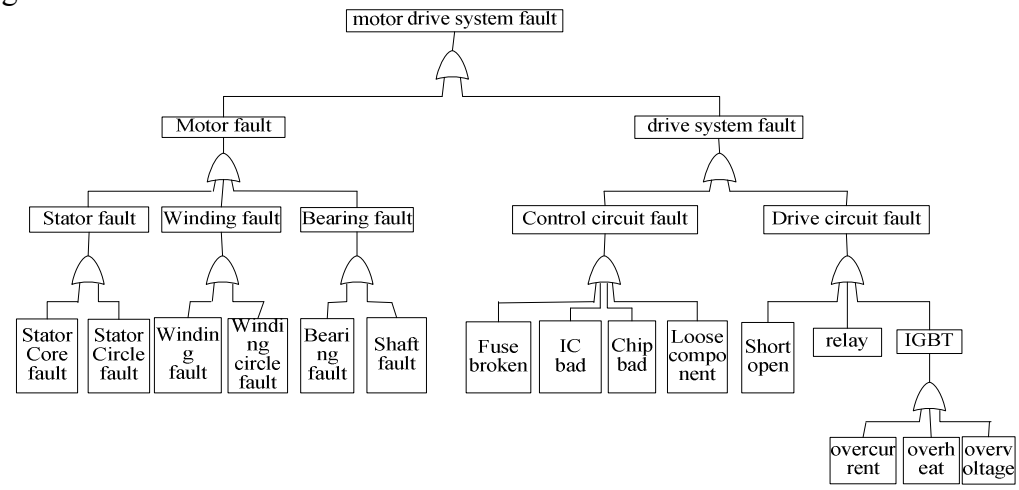


Figure 2 FTA of motor drive system

Reliability Analyses

Reliability model

Reliability models are reliability block diagram and mathematics models that built to predict and evaluate reliability of products. Servo system consist of motor body and drive control circuit, if the failure occurred in any of them, the whole servo system will work unmorally, so the logical reliable relationship between motor and drive control circuit is series relations actually. It is shown in Figure 3.

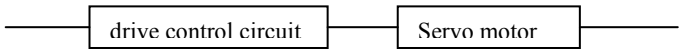


Figure 3 Reliability block diagram of servo system

The drive control circuit is usually consist of power tubes, resistors, capacitors, manifold blocks and other components, basically any of these components fail, then the whole drive control circuit will not work properly, that is, the drive control circuit, the reliability logic relations of these components for the series. The servo motor was composed of rotator winding, rotor winding, bearing and so on , any unit breaks down, will cause the motor to break down, therefore, the logical reliable relation between various parts in motor body for series relations.

Reliability improvement

From the above analysis, reliability model of motor drive system is the series model composed by bearing, winding insulation and drive control circuit, as the weak link, also the important component of reliability improvement.

(1)Bearing Affect the reliability of the main stress bearing load to the bearing, speed and temperature. Increase bearing load makes fatigue strength increased, speed up the failure of the bearing; Increase speed makes the surface wear of rolling bearing increases; Temperatures prompting bearing grease acceleration, bearing outer oxidation, volatile so bearing fatigue limit reduced. Therefore is mainly reduces its load, the rotational speed and the highest operating temperature.

(2)Winding The most important influence factor of the reliability of the insulation is temperature. Thus keep the current application temperature and winding down.

(3) Drive control circuit The main consideration for the drive control circuit machine circuits, transistors, resistors, capacitors, and several other components. As the internal parameters for the integrated circuit usually allowed range is small, we should focus on ways to improve its cooling package to lower the device's junction temperature, to minimize the input level, output current and operating frequency; Temperature is the most important transistor stress, so keep the power consumption and transistor junction temperature down, and keep the voltage down to prevent voltage breakdown. For fixed resistors and potentiometer, to impact its reliability mainly on voltage, power and the stress environment temperature; For thermostats, impact its reliability mainly on the power and the environment temperature; For resistors and capacitors, to impact its reliability manly on voltage and environmental temperature stress. Therefore, in its high frequency application situation, its loss of voltage volume scope should further increase, and it is especially sensitive regarding the electrolytic capacitor.

Conclusions

Based on statistics analysis reliability of the CNC system, utilizing the field data, the failure spots, failure patterns and failure mechanism of the CNC system were analyzed and then the critical and fault tree were analyzed, so the weak link of servo system were obtained which were bearing, winding insulation and drive control circuit. The reliability of the weak part of the CNC system and the reliability model were analyzed, and then some improvement measures for its reliability were put forward.

Now the work on the reliability of the CNC system is far enough, the texts of the reliability are rarely seen. Hope that it can provide a reference for further research of the reliability of the CNC system, and laid the foundation intensifying of reliability enhancement test in the future.

References

[1] Zhang Ye,Kang JinSong. Research of the reliability on motor drive system [J]. *Mechatronics*. 2008.12, p64-68

[2] Zhang HaiBo,Jia YaZhou,Zhou GuangWen. Failure Mode, Effects and Criticality Analysis (FMECA) of CNC System [J].*Mechanical engineering*.2004, 15(6), p 491-494

[3] Qiao WeiWei. Modeling and entropy weight fuzzy comprehensive evaluation of reliability in CNC system [D]. *Jilin University,* 2006

[4] Xiao Jun. Reliability technology analysis and research of CNC Machine Tools [D].*shanghai jiaotong University,* 2007

[5] MackenKJP,Bellenmll.Reliability. Assessment of Motor Drives[C].*Power Electronics Specialists Conference, 2006. PESC"06.37th IEEE* 8-22 June 2006,pl-7.

[6] WangMin,ChenDiansheng,LuTianshan.Failure Mode, Effect and Criticality Analysis of CNC Lathes [J]. *International Conference on Progress of Machining Technology.* 2004.12, p486-490

[7]QiaoWei-Wei,JiaYa-Zhou,ZhangHai-Boetc.Failure analysis and reliability enhancement measures of computer numerical control system [J]. *Journal of Jilin University.* 2006, 36(2).p69-72

[8] Alakesh Manna.A study on fault diagnosis and maintenance of CNC-WEDM based on binary relational analysis and expert system[J].*The International Journal of Advanced Manufacturing Technology*. 2006, 29(5), p23-27

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.1019

Study on the CT real-time scanning tests and deterioration evolution process of recycled concrete

Chen Ai-jiu[1,a], Zhang Qing[2,3,b], Wang Jing[1,c], Ge Zhan-fang[1,d]

[1]School of Civil Engineering and Communication,North China University of Water Conservancy and Electric Power; Zhengzhou,China;

[2]State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, China

[3]Department of Engineering Mechanics, Hohai University, Nanjing, 210098, China
[a]email:caj1967@163.com, [b]email:lxzhangqing@hhu.edu.cn, [c]emaill:wj@ncwcu.edu.cn, [d]email:gezhanfang@ncwcu.edu.cn

Key words: recycled concrete; CT numbers ; deterioration

Abstract. Based on CT imaging theory and self-designed loading device, the CT real-time scanning test on recycled concrete under uniaxial static compression has been accomplished, which can be used to specifically study the whole process about microcrack initiation, perforation, coalescence and macrocrack generation in meso-scale. By real-time CT scanning of cylinder recycled concrete specimen, the propagation process of the internal cracks can be effectively tracked during the whole loading process. By extracting and analyzing the CT numbers of CT images under different loading period, the relational expression between CT numbers and deterioration variable was established, which can better reflect the effect of mixing amount of recycled coarse aggregate on deterioration degree of recycled concrete. The results showed that with the increasing mixing amount of recycled coarse aggregate, the deterioration degree of the specimens increased slightly yet uniaxial compression strength of the specimens decreased slightly.

Introductions

Oral Buyukozturk [1] used medical CT machine to scan the small concrete structures and the clear cross-section images of aggregate and microcracks had been acquired, but there was a great gap between technical index of image resolution and imaging velocity at that time and nowadays; Dang F N et al[2-4]. used medical X-ray CT machine to scan the cross-section of rock and concrete materials。By real-time CT scanning of concrete specimen with different mixing amount of recycled coarse aggregate, the CT images under different loading period were extracted and the propagation process of the internal cracks can be effectively tracked. After analyzing the change of CT numbers for the same specimen, the relational expression between CT numbers and deterioration variable was established. Finally, several conclusions were drawn from the test results.

Primary materials and mix proportions for tests

Primary materials for tests

After crushing and processing the waste reinforced concrete, the recycled coarse aggregates can be acquired with the grain size ranging from 5mm to 20mm. The material used in the test was grade P.O32.5 cement, natural river sand with fineness modulus 2.90, natural coarse aggregate of gravel, polypropylene fiber of bundle monofilament with good dispersibility, drank tap water of

Zhengzhou city for concrete mixing and curing process.

Mix proportions for tests

Design strength grade of recycled concrete specimens used in the tests was C30 and mix proportions of each component were shown in Tab. 1. Size of cylinder specimens was 50mm(diameter)×100mm(height) and the specimens were loaded after 28 day's standard curing age.

Table 1 Mix proportions of recycled concrete for CT tests

serial number	mixing amount of recycled aggregate(%)	mixing amount of polypropylene fiber(%)	the content of every kind of components in 1m3 concret(kg)					
			water	cement	sand	nature gravel	recycled aggregate	polypropylene fiber
R0	0	1.0	165	413	626	1180	0	1.0
R50	50	1.0	165	413	626	590	590	1.0
R100	100	1.0	165	413	626	0	1180	1.0

CT tests

Test conditions

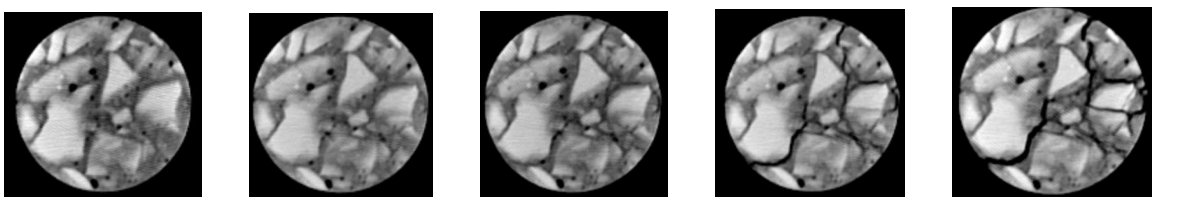
The tests were carried out in the CT room and medical spiral x-ray scanner of Siemens type Huan Yue Duo was used. The scanning resolution was 0.35 mm×0.35 mm×1mm; density resolution was 3‰ of water absorption coefficient. The scanning parameters were voltage 137kV, current 220mA and amplification factor 6.0 Uiaxial static compression tests of specimens R0, R50 and R100 were carried out on servo-controlled testing machine and self-designed loading equipment before CT tests aiming at determining compressive strength grade and corresponding stresses and deformation characteristics of specimens in advance.

Test process

Move the CT worktable and set the specimen at specified position. Start the loading device and load uniformly and continuously with a rate of 0.3MPa/s; when approaching failure and specimen deform rapidly, control the throttle and decrease the loading rate until the specimen failure ultimately and record the failure load. Each specimen need five or six scanning procedure with scanning interval 5mm and scanning thickness 3mm.

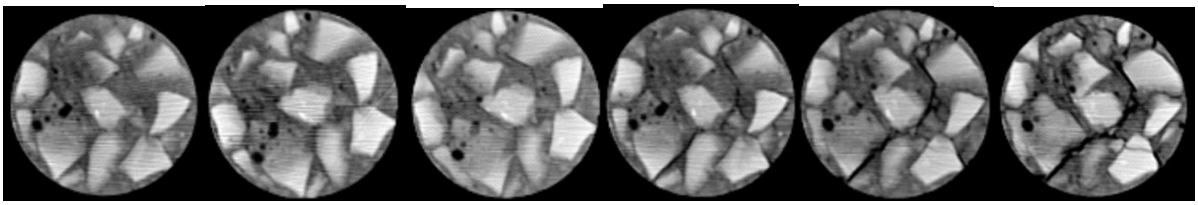
Results and analysis of spiral CT tests

CT scanning images was 512×512, image resolution was 0.15mm and each pixel point in the image was attaching with CT numbers stored with 12-bit number that had 4096 orders. The different scanning images of twelfth section of specimen R0, R50 and R100 with different stress stage were shown in Fig.1~Fig.3.



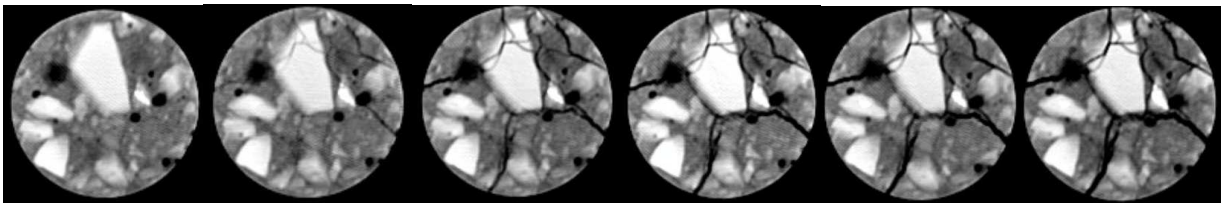
(a) σ =0MPa (b) σ =37.6MPa (c) σ =37.0MPa (d) σ =36.5MPa (e) σ =36.0MPa

Fig. 1. Scanning images of specimen R0 during different stress process



(a) σ =0MPa (b) σ =35.9MPa (c) σ =35.6MPa (d) σ =35.1MPa (e) σ =34.8MPa (f) σ =34.4MPa

Fig. 2. Scanning images of specimen R50 during different stress process



（a） σ =0MPa (b) σ =34.6MPa (c) σ =34.2MPa (d) σ =33.7MPa (e) σ =33.2MPa (f) σ =32.8MPa

Fig. 3. Scanning images of specimen R100 during different stress process

As shown in Fig.1 and Fig. 2, because of contraction and volume change caused by damp-heat effect, initial microcracks had already existed in the concrete specimen before loading. Microcracks extended continuously at the weak interface between cement mortar and aggregate with increasing load. For the image shown in Fig.3, many microcracks formed and some aggregates were intersected by the microcracks, this was mainly because mixing amount of recycled aggregate was 100% and some aggregates had been damaged during the crushing process. When the specimen went into failure stage, the number and width of interior microcracks were all increase quickly and continuous macocracks formed which was consistent with the macro failure behavior of concrete.

In the Fig.1~Fig.3, the brighter of the zone, the bigger density and CT numbers of the material and vice versa. For the same centre of the circle, the CT numbers of damage region and failure region at each loading stage with same radius were analyzed by statistics, average CT numbers and standard deviation of each image can be acquired, as shown in Tab.2~Tab. 4

Table 2 CT numbers of damage zone in R0 specimen with different stress

Stress(MPa)	0	37.6	37.0	36.5	36.0
CT numbers	2080	2078	2075	2039	1978
standard deviation	339.7	351.2	364.0	377.0	457.9

Table 3 CT numbers of damage zone in R50 specimen with different stress

Stress(MPa)	0	35.9	35.6	35.1	34.8	34.4
CT numbers	1981	1983	1966	1933	1910	1884
(standard deviation)	324.7	344.7	350.7	357.3	372.9	404.6

Table 4 CT numbers of damage zone in R100 specimen with different stress

Stress(MPa)	0	34.6	34.2	33.7	33.2	32.8
CT numbers	1853	1821	1802	1782	1742	1721
standard deviation	378.2	380.6	387.0	396.7	405.0	442.0

From Tab.2~Tab. 4, we can see that with the increase of mixing amount of recycled coarse aggregate, the CT numbers decreased which mean that interior deterioration degree of the specimen increased slightly. The relation between CT numbers and different stress can be plotted as Fig.4, which showed that there was a linear relationship between CT numbers and stress before the peak strength which mean that the specimen occured uniform compression deformation and the deformation was mainly elastic; the curve deviated from straight line after the peak strength and the change of CT numbers increased rapidly which mean that there were local macrocracks and the deformation was mainly plastic, and the observation from the images was consistent with analysis result of average CT numbers. With the increase of mixing amount of recycled coarse aggregate, the change of CT numbers varied considerably which mean that the uniaxial compression strength of recycled concrete decreased correspondingly.

Analysis of deterioration evolution

There were many kinds of methods to define different deterioration variables for the same failure process of specimens. In this paper, change of CT numbers recorded in the CT test was used as deterioration variable to describe the damage degree of specimens. Based on the image processing, damage region of interest (ROI) can be defined and measured aiming at describing RIO of specimens divisionally by the same standard and CT numbers in each zoning can be determined and analyzed. The microcracks number and length can not be analyzed quantitatively, yet combined with change of CT numbers, microcracks areas of ROI can be calculated which can be used to describe the deterioration variable exactly and acquire corresponding evolution equation. The statistic deterioration variable D reflecting the local deterioration degree was defined as:

$$D = \frac{\text{decrease of CT numbers}}{\text{initial CT numbers}} = \frac{H_{rmo} - H_{rmi}}{H_{rmo}} \qquad (1)$$

Where H_{rm0} was the average CT numbers of ROI under initial condition; H_{rmi} was the average CT numbers of ROI under a certain stress stage.

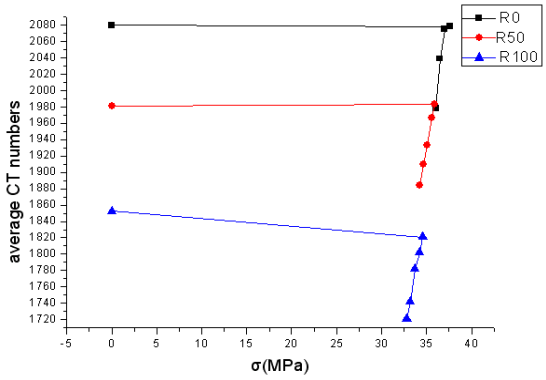


Fig. 4. The relation between average CT numbers and stress

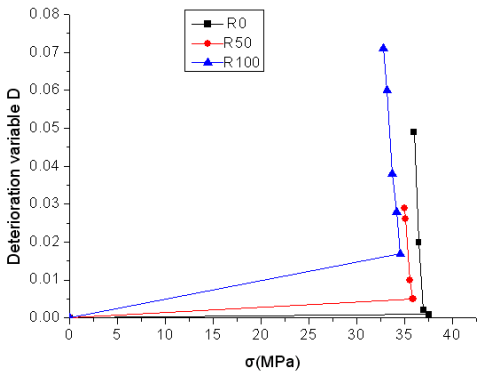


Fig. 5. The relation between deterioration variable of specimen section and stress during loading process

Microcracks and damage region gradually appeared in the interior of specimen with increasing load, corresponding CT numbers decreased gradually. Setting the unloading region as the basic region, the change of CT numbers can be substituted in Formulae (1) and deterioration variable D can be determined. The relation between deterioration variable of specimen section and stress

during loading process was shown in Fig. 5.

In Fig.5, the deterioration variable was very small prior to peak strength of concrete specimens which mean that the number and length of microcracks were very small; when approaching to peak strength, there was a steep rising stage in the curve which mean that mocrocracks extensions rate increased rapidly. The deterioration of specimen cumulated and evolved with continuous loading and CT numbers of each section decreased. We can also see that with the increase of mixing amount of recycled coarse aggregate, the deterioration variable of recycled concrete increased faster than that of ordinary concrete and uniaxial compressive strength of recycled concrete decreased faster than that of ordinary concrete.

Conclusions

(1)Combing self-designed loading device and CT testing system, the CT real-time scanning test on recycled concrete under uniaxial static compression had been accomplished aiming at tracking the propagation process of the internal cracks during the whole loading process, which can provide an effective way for studying on the damage mechanism of recycled concrete.

(2)By extracting and analyzing the CT numbers of CT images under different loading period, the relational expression between CT numbers and deterioration variable was established which can be better show the effect of mixing amount of recycled coarse aggregate on deterioration degree of recycled concrete.

(3)The real-time CT images showed that the concrete failure began at the interface of cement mortar and aggregate: micro-cracks initiated, perforated to the interior of mortar and formed marocrack finally.

(4) The results showed that with the increase of mixing amount of recycled coarse aggregate, the deterioration variables of the specimens increased which mean that the density of recycled concrete was lower that the ordinary concrete and uniaxial compression strength of the specimens decreased.

References:

[1]Oral Buyukozturk. Imaging of concrete structures[J].*NDT&E International,* 1998,31(4):233-243.

[2]Dang Faning，Yin Xiaotao，Ding Weihua，Chen Houqun. Breakage mechanism of sandstone based on uniaxial compression CT test [J]. Chinese *Journal of Rock Mechanics and Engineering*,2005,24(22):4003-4009.

[3]Chen Houqun,Ding Weihua, Pu Yibin. Real time observation on meso fracture process of concrete using X-ray CT under uniaxial compressive condition[J]. *Journal of Hydraulic Engineering*,2006,37(9):1044-1050.

[4]Li Fen, Li Zhida, Gong Junan, Zhao Xianpeng. Research on Micro-Structure Evolution of Asphalt Mixture Based on Estimation of Fractal Demensions of CT Image [J]. *Natural Science Journal of Xiangtan University*， 2006 ，28(2)：52-55.

© (2010) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.1024

Design and Development of Simple and Instant Communication Tool

Xilong Qu[1,2] ， Lei Xiao[2]

[1] School of Computer Science & Engineering,South China University of Technology,Guangzhou, China

[2]School of Computer & Communication, Hunan Institute of Engineering, Xiangtan, China

quxilong@126.com (Xilong Qu)

Key words: instant message; multithreaded; socket program; JSP

Abstract. Instant Message is currently the most popular way to communicate on the Internet, so various Instant Messaging software are continuously appearing; Service providers offer more and more communication service nowadays. Java programming language (Java 2 Standard Edition) is one of popular programming languages. With its features of cross-platform, security, multi-threaded capability, Socket Programming, data flow concepts and so on, Java has its own distinctive and functional advantages. Therefore, based on the research and analysis of LinuxQQ, with the Java language and J2EE, this paper designs an instant messenger—JICQ (Java for I seek you), and then analyzes its architecture, modules, as well as its key technologies. During the design and modeling process of instant messenger system, the UML and method of object-oriented analysis and design are used. Furthermore, the modeling tool of Rose is included. Based on the system j2se1.5, j2ee1.4, the Eclipse development tool in the development process, and refactoring development method, the system design is optimized. Due to these factors, the system is safe, efficient and practical in different operating system platforms.

Introduction

Technology based on the sun company java language, the language has many advantages, for example, multi-thread, Network flows concept, arrested deal with exception, security and speed and performance, can write with a time, crossing platform merit and running anywhere.

Multiple threads is such a mechanism, which allows programs executing concurrently more instruction flow, each instruction flow is called a thread, and one independent, threads also called processes that lightweight. The multiple threads technology allows one application starts many light of program that makes many applications to competition the CPU resources, thereby improving efficiency of the system. For communication of internet tools, multiple threads is required, not only to efficiency but also can achieve deal with many requests at the same time,

Network flow ： the internet communications and the reading of Java are understanding as a conception, and so on, network data transmission like reading the document on a machine, in favor of to understand and write programs ， in java class library is full of IO flow of resources and internet programming resources, the software developers can be very convenient to write the web application program.

EXCEPTION：Java deal with the exception by object-oriented way. in a method of operation, if the exceptions, then this method generates an object representative the exception and put it to run-time systems, the system look for the corresponding code to handle this exception. We call throwing an exception for creating an object and putting it to the Run-time the system process.

Run-time system finds in the call stack ways, from method generating exceptions now to back, until find to contain the corresponding exception method, the process called capture (catch) an exception. Therefore, java exceptions are very convenient.

Cross-platform ： the very important feature of java language is irrelevant to the platform. Using java virtual machine is the key to carry out the characteristic. if the high-level language wants to run in different platforms, which needs to be compiled into object code at least. However, the java language does not to be recompiled when runs on different platforms after inducting the java virtual machine. Java language to shield impertinent information for concrete platform uses virtual machine, the java language compiler to generate java virtual machine runs on the object code () representation, in a variety of platforms are running. Java virtual machine in the representation is the representation of specific explanation in the machine instructions.

Development Environment

Borland JBuilder is the first Java development environment across platform in globe that can be used to build with the industry standard java applications systems , not only develop the table level procedure but also develop web, ejb , xml and databases applications. Two-way, visual design tool that we can build a variety of applications, and deploying to many kinds of application servers and use jbuilderwe can quickly and conveniently build a java applications, and it provides the function in the editor similar to the vs. net, it has an absolute issue, we don't have to arrangement the java layout manager by hand, the whole can drive fast using jbuilder.

Eclipse is an opened source code, based on java extensible development platform. for itself, it is only a framework and a set of service, used the components of plug-in to develop environment. fortunately, eclipse attaches an additional plug-in set of standard including java development tools(ava development tools, jdt). Although most people are very pleased to use eclipse as java ide, but the eclipse target more than it, the eclipse also includes plug-in development environment (plug in and development environment, pde), the components aims to the software developers hope to expand eclipse, because it allows them to build a seamless integration with the eclipse environment. Using eclipse plug-in to build a dynamic jsp process will become more comfortable, fast, easy and convenient.

Rose, a set of great value system modeling tools, it can provides a convenient system modeling functions such as example chart, behavior chart, the movements and the weather chart etc. It can be speeded up the system design and development effectiveness， providing user and the developer clear chart, their communication will be more convenient and effective.

SQL SERVER2000 Database is Microsoft's powerful database management system, have many database management common functions, such as storage process, data query and update etc. it is a very ideal backend database system.

The JSWDK Tomcat's Sun microsystem JavaServer (Web) in the short Kit Servlet operation environment (Servlet vessels), is a free open source of Serlvet containers, it is a core Jakarta project of Apache foundation, they are developed by Sun micro system and some other companies and individuals commonly.

System design

The system refines module as function, mainly including registered logged on, buddy talking, buddy management, crowd talking four modules, there are many other small parts, the four modules

is the core modules to achieve the system, the coupling among modules is very low, the registered module achieve by b/s structure, without install the client.

The registered module is most basic module of the whole system, the module primarily through b/s to realize, the technology adopts three (n) structure of the j2ee, it has the following characteristics:

- effectively reduce the building and maintenance costs and simplify the management;
- adapt scale and complexity application needs;
- can be used to constantly change and the new business requisitions,
- visit from different database
- effectively improve the system concurrent processing capability;
- effectively improve the system security

Performance of the modules with technology is the servlet and jsp and the database use sql database is 2000. The modules including the following functions ： account registration, passwords retrieve, passwords alter and personal information for enquiries. The user can make the number of registered, passwords retrieve and other features interact with the web server through the module.

Complimentary close

Along with the rapid development of Internet, the Internet communication has become one way of communication by many people, this system is designed to realize real-time communication on the Internet, it is one of the most popular Java programming languages to realize real-time communication tools, including the rudiment of many aspects of the technology used is a Java or difficulties. You can go through this system for the application number, you can login system after some simple operation and Settings, such as adding friends and friends group, etc, and then you can communicate with friends, such as real-time transmission, and the immediate response, etc. This system can run on the Internet or LAN, users can easily for information exchange.

Through the implementation of this system, the instant communications has good application prospects, function will more and more perfect, the service providers improve more services, there will more users, the kinds of instant communications software itself is also being expanded, and to an instant communications software is also perfect. Instant communication development and application may be the internet and the whole world's a miracle, I sincerely hope that more scientific workers and experts join in the field of research.

Acknowledgment

Authors gratefully acknowledge the Projects Supported by Scientific Research Fund of Hunan Provincial Education Department(08A009) for supporting this research. Project supported by Provincial Natural Science Foundation of Hunan（10JJ6099）supports the research. Project supported by Provincial Science & Technology plan project of Hunan(2010GK3048) supports the research.This research is supported by the construct program of the key discipline in Hunan province.

Reference

[1]Mao Yuxin, Wu Zhaohui, Chen Huajun, Xu Zhao. Context-based web ontology service for TCM information sharing. Proceedings - 2005 IEEE International Conference on Web Services, July,2005. United states,pp:699-705

[2]Wang Shuying, Zhao Huijuan,Sun Linfu. Research on Resource Integration Framework of Regional Networked Manufacturing ASP Service Platform. China Mechanical Engineering. 2005, Vol.24, No.5, pp:1729-1732

[3]Zhang Yang1, Guan Xue-Tao1, Huang Tao1, Cheng Xu1. A heterogeneous auto-offloading framework based on web browser for resource-constrained devices. Proceedings of the 2009 4th International Conference on Internet and Web Applications and Services. May, 2009 ,Venice, Mestre, Italy, pp:1-22

[4]Paci Federica1, Ouzzani Mourad2, Mecella Massimo. Verification of access control requirements in web services choreography. Proceedings of 2008 IEEE International Conference on Services Computing. July, 2008, Honolulu, HI, United states, pp: 5-12

[5]Gao QB, Jin ZC, Ye XF, et al. Prediction of nuclear receptors with optimal pseudo amino acid composition[J]. Analytical Biochemistry, 2009, 387(1): 54-59.

[6] Müller KR, Mika S, Rätsch G, et al. An Introduction to Kernel-Based Learning Algorithms[J]. IEEE Transactions on Neural Networks, 2001, 12(2): 181-201.

Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.50-51.1028

Improvement design research of a CNC lathe-bed structure

DUAN Yaodong[1,a] ZHANG Guangpeng[1,b] GUO Chun[1,c]

PENG Lili[1,d] Huang Yumei[1,e]

Faculty of Mechanical and Precision Instrument Engineering, Xi'an University of Technology,

Xi'an, 710048, China

duanyaodong@sina.com, gpzhang@xaut.edu.cn, guochun0417@163.com,

penglilijyxy@163.com, hymxaut@163.com

Keywords: Structure modifying, Lathe-bed, Dynamic characteristics, Static characteristics, Finite element analysis

Abstract. The lathe-bed structure has a great influence on the overall performance of a lathe. The CNC lathe-bed model is established based on the finite element method, then the defects of the lathe-bed structure were found out by dynamic and static analysis. The reasonable layout of stiffened plate structure in the lathe-bed is put forward without altering the shape size and interface dimension of the lathe-bed, and three better types of modification configuration of the lathe-bed were proposed in this paper.

Introduction

With the development of high speed and high precision of CNC lathe, the requirements of the static and dynamic performance of the lathe structural components have become more and more significant. Therefore, the study of reasonable structural design method of a machine tool has become an important issue in the field of machine tool design. Nowadays, the methods of stiffened plate configurations in machine tool structure design are mainly based on finite element method[1,2], such as combining
the parametric design with finite element analysis[3] and bionics method[4].

Although all these methods could effectively increase the stiffness of machine tool structure, a class of problems has not been effectively resolved in the field of machine tools structure design, i.e. how to efficiently and quickly modify the machine tool structure without altering the shape size and interface dimension of the structure. For a machine tool structure, good technological properties is necessary, not only to consider structural stiffness, but also easy to manufacture. Therefore excessive constrained conditions for the structure optimum design have emerged. It is difficult to adopt topology optimization to optimize the machine tool structure. Exploring an effective and practical design method on modifying machine tool structure is essential.

Based on finite element code ABAQUS, static and dynamic model of a high precision CNC lathe-bed were established. The defects of the lathe-bed structure were found out by dynamic and static analysis. The reasonable layout of stiffened plate structure in the lathe-bed is put forward without altering the shape size and interface dimension of the lathe-bed, and three better types of modification configuration of the lathe-bed were proposed in this paper.

Finite element model of the lathe-bed

Structure simplified methods

First the CAD model of the lathe-bed was built by Solidworks. Then the model is input into ABAQUS by the form of Para solid. The actual structure of the lathe-bed was simplified so as to establish a reasonable effective finite element model and improve the calculation efficiency. The followings were the basic principles for simplified the structure:

(1) Ignoring small features in the CAD model, such as chamfer, circular bead, and some small holes and lug bosses which size are less than 20mm.

(2) Changing the smal1 taper and small curvature surfaces to straight line and flat surface.

(3) Removing some small structures, such as the uplift in the guide way basement and some small lug bosses on the side bed for installation, etc.

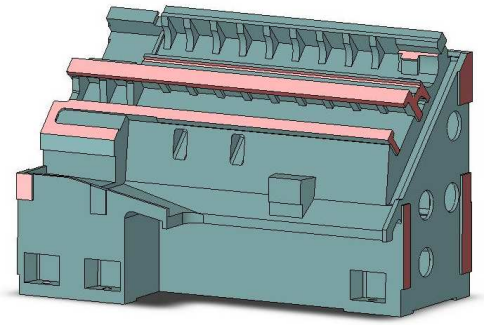


Fig. 1 a CNC lathe bed outside view

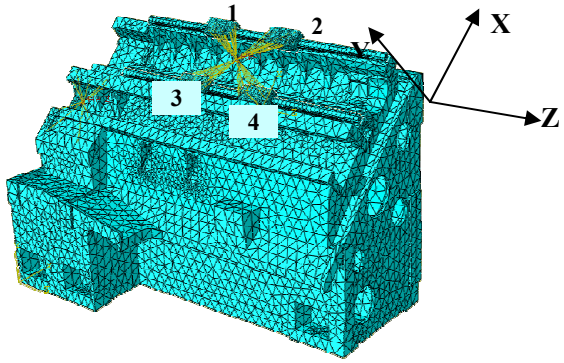


Fig. 2 the finite element model of bed

The static FEM model of the lathe-bed

The main purpose of static analysis is to get static deformation of the lathe-bed under cutting forces. The load forces on the lathe-bed include two parts: (1) The gravity of other parts on the bed, for example, spindle box, vertical and horizontal slide system, turret, tailstock, vertical drive screw, motor, hydraulic system, and electronic control system, etc. Gravity of these components is equivalent to distributed loads or torques and loaded on the corresponding surface of the lathe-bed. (2) The cutting forces on the lathe-bed.

In addition, the lathe-bed was connected with foundation by bolts. Since the stiffness of the join between the lathe-bed and foundation is difficult determined, the join constraints were simplified as fixed constraints. And this constraint has no effect on comparing different bed options.

Figure 1 is a CNC lathe-bed outside view, the length of the bed is 1590mm, and width 860mm, height 1076mm. The material elastic
modulus of the bed is 1.1e5MPa, density is 7200kg/m3, and poisson's ratio is 0.28. The finite element model used modified tetrahedron element (C3D10M), and size of grid is 45mm. Figure 2 is the finite element model of the lathe-bed.

The dynamic model of the lathe-bed

As the same reason above, the stiffness of bolt joins between the lathe-bed and foundation can not be determined accurately. If using different bolt joins stiffness, it is difficult to find the modifying location in the bed by the FEM analysis results. In this paper, free modal finite element analysis was conducted on all configurations of the lathe-bed.

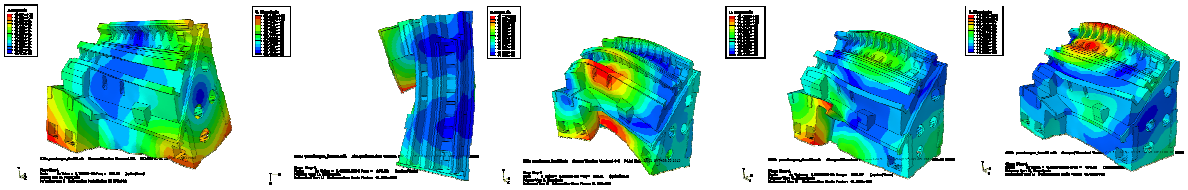
Performance analysis of the bed structure

Modal analysis

For a machine tool structure, the lower order modes have greater impact on actual manufacturing process. This paper analyzed the first five modal of the lathe-bed using the Lanczos method, and the first five natural frequencies and their modes were extracted.

The following information can be obtained from the modes in figure 3: the first-order mode is rotation motion along the Z direction, the second-order mode is transverse horizontal bending, and

the third-order transverse vertical
bending. The forth and firth-order modes reveal a weak stiffness in the guideways. Therefore, from the view of structural modification, improving the bending stiffness and torsional stiffness are the most efficient methods to enhance the performance of the lathe-bed structure.



(a)1st mode (b)2nd mode (c)3rd mode (d)4th mode (e)5th mode

Fig. 3, Bed Structural natural modes predicted from FEM analysis

Bed structure modification

In order to improve torsional stiffness of a machine structure, the material should be distributed in the surroundings to increase torsional moment of inertia section. Longitudinal stiffened plate layout directly affects torsional stiffness of the lathe-bed. Figure 4 display the section of the lathe-bed, the longitudinal stiffened plate layout used network structure, and B cavity is rectangular, C cavity is square and A cavity is triangular. Rectangular cavity has poor torsional capacity and triangular cavity A is larger. The middle slideway D lacks of longitudinal stiffened plate resulting in relatively lower bending stiffness.

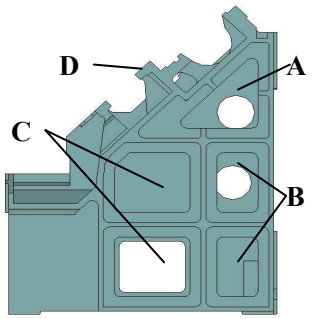


Fig. 4 the cross section of original bed

In view of the above analyses, the lathe-bed has weak torsional and bending stiffness. Three types of modification configurations were proposed without altering the shape size and interface dimension of the lathe-bed in this paper:

(1)Arc-shaped stiffened plate configuration (fig 5.a)

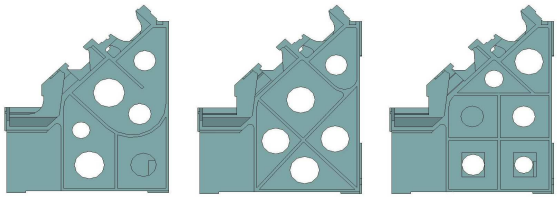
Change the stiffened plate on the triangle A to approximate arc-shaped to enhance overall torsional stiffness of the lathe-bed. At the same time, a longitudinal stiffened plate under the middle slideway D was added to improve the bending stiffness.

(2) Cross-vertical stiffened plate configuration (fig 5.b)

A longitudinal stiffened plate under the middle slideway D was extended to back wall of the lathe-bed to improve bending stiffness and formed a closed area in the top of the lathe-bed. In addition, a longitudinal cross-vertical plate was accepted to enhance overall torsional stiffness of the lathe-bed.

(3) Square mesh stiffened plate configuration (fig 5.c)

A longitudinal stiffened plate under the middle slideway D was added to improve the bending stiffness of the slideway. B and C cavity were changed into small square cavity, which not only improved the torsional stiffness of the lathe-bed structure, but also increased the bending stiffness.



(a) Option 1 (b) Option 2 (c) Option3

Fig. 5 the geometry of three machine configuration

The three types of modification configurations have the same number of transverse plate with that of the original lathe-bed. Parametric finite element analyses were conducted on these three models to determine the best parameter of the lathe-bed with minimum weight, maximum static stiffness and maximum first natural frequency as optimization objectives. In option one the parameters are diameter of circle in transverse plate, radius of Arc-shaped stiffened plate, length of longitudinal stiffened plates under the middle slideway D, size of openings in longitudinal stiffened plates, distance between middle plate and front wall of the lathe-bed. In option two and three the parameter is diameter of circle in transverse plate.

Table 1 shows the first five natural frequencies and mass of the different lathe-bed options. It can be seen that the first five natural frequencies of modified configurations are much higher than that of the original bed. The three modified lathe-beds have a higher stiffness than that of the original bed.

Table 1, the first five natural frequencies of four designs and their mass

	Frequency(Hz)			
Mode.no	Original	Option1	Option2	Option3
1	213.1	329.6	304.6	318.3
2	292.7	395.4	395.9	420.7
3	374.8	517.2	529.2	527.0
4	390.1	585.3	595.9	590.7
5	408.5	603.3	610.0	622.0
Mass/kg	1256.3	1404.1	1462.5	1467.0

Table 2, the displacement on the point 1, 2, 3, 4[μm]

View point	1				2			
direction	X	Y	Z	total	X	Y	Z	total
original	-5.13	-0.50	-1.17	5.29	-4.94	-0.34	-1.31	5.13
Option 1	-3.24	-0.61	-0.70	3.37	-3.36	-0.19	-0.77	3.45
Option 2	-3.23	-0.57	-0.65	3.34	-3.27	-0.16	-0.75	3.43
Option 3	-3.19	-0.53	-0.62	3.29	-3.25	-0.14	-0.69	3.32
View point	3				4			
direction	X	Y	Z	total	X	Y	Z	total
original	-4.44	-4.95	-1.15	6.75	-4.44	-4.95	-1.15	6.34
Option 1	-2.93	-2.38	-0.86	3.84	-2.69	-2.22	-0.75	3.56
Option 2	-2.91	-2.22	-0.85	3.75	-2.67	-1.92	-0.72	3.36
Option 3	-2.88	-2.09	-0.81	3.64	--2.63	-1.89	-0.69	3.31

Option 1 has the highest first-order natural frequency (329.6Hz), 55% higher than that of the original design, indicating the Option 1 configuration was more conducive to raising the torsional stiffness of the lathe-bed structure. Option 3 has the highest second-order natural frequency (420.7Hz), 44% higher than the original design, indicating the configuration of option 3 is more conducive to raising the bending stiffness of the lathe-bed structure in the horizontal direction. Option 2 and 3 both have the highest third-order natural frequency, 41% higher than the original design, indicating the configurations of option 2 and 3 are more conducive to raising the bending stiffness of the lathe-bed structure in the vertical direction. Because of a longitudinal stiffened plate under the middle slideway D was added in the three types of configurations, the forth and fifth-order natural frequencies are 50% higher than the original bed structure.

In order to compare static stiffness of the three types of configurations, static analysis with the same cutting force was conducted to get displacement on the slide block in the lathe-bed. Table 2 indicated the displacement on the point 1,2,3,4 and figure 2 showed the location of point

1,2,3,4.Deformation of tool rest is effected by the displacement on the slide block, which were all smaller than the original one, indicating that the static stiffness of modified options were greatly improved.

Conclusions

The design defects of the lathe-bed structure were found out by dynamic and static analysis based on finite element code ABAQUS. Three better types of modification configurations were proposed without altering the shape size and interface dimension of the lathe-bed. The simulation results reveal that static and dynamic performances of modified lathe-bed structures were greatly improved than the original structure. Lathe-bed structure modification methods under constraints proposed in this paper were beneficial to machine tool structure design.

Acknowledgments

The work is supported by National S & T Major project (2009ZX04014032) and National Basic Research Program of China (2009CB724406) and Project of NSFC (50875208).

References

[1].MAO Haijun, etc. Influence of Arrangement Patterns of Strengthened Bars on Dynamics of Machine Tool, Manufacturing Technology & Machine Tool, 2001(4):8-9.

[2] Dai Gil Lee, Jung Do Suh, Hak Sung Kim, Jong Min Kim，Design and manufacture of composite high speed machine tool structures，Composites Science and Technology 64 (2004) 1523–1530

[3] D.I. Kim, S.C. Jung, J.E. Lee, S.H. Chang，Parametric study on design of composite–foam–resin concrete sandwich structures for precision machine tool structures. Composite Structures 75 (2006) 408–414

[4] Ling Zhao, Wu-yi Chen, Jian-feng Ma, Yong-bin Yang，Structural Bionic Design and Experimental Verification of a Machine Tool Column，Journal of Bionic Engineering Suppl. (2008) 46–52

Keywords Index

D

E

F

G

H

I

N

O

P

Q

R

S

T

Authors Index

H

J

K

L

T

V

W

X

Y

Z