



COMPUTER MODELLING
AND
NEW TECHNOLOGIES

2014
VOLUME 18 NO 3

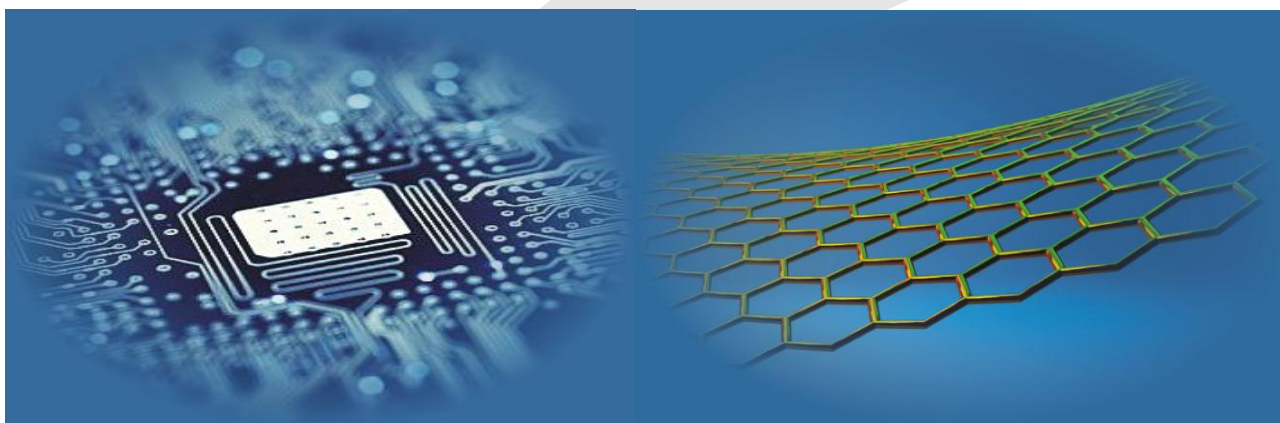
ISSN 1407-5806 ISSN 1407-5814 on-line

Transport and Telecommunication Institute
and
Latvian Transport Development and Education Association

Computer Modelling and New Technologies

2014 Volume 18 No 3

ISSN 1407-5806 ISSN 1407-5814 (*On-line: www.tsi.lv*)



Riga – 2014

EDITORIAL BOARD

Prof. Igor Kabashkin	Chairman of the Board , <i>Transport & Telecommunication Institute, Latvia</i>
Prof. Yuri Shunin	Editor-in-Chief , <i>Information Systems Management Institute, Latvia</i>
Prof. Adolfas Baublys	<i>Vilnius Gediminas Technical University, Lithuania</i>
Dr. Brent Bowen	<i>Embry-Riddle Aeronautical University, United States of America</i>
Prof. Olgierd Dumbrajs	<i>University of Latvia, Solid State Physics Institute, Latvia</i>
Prof. Sergey Maksimenko	<i>Institute for Nuclear Problem, Belarus State University, Belarus</i>
Prof. Vladimir Litovchenko	<i>V. Lashkaryov Institute of Semiconductor Physics of National Academy of Science of Ukraine, Ukraine</i>
Prof. Pavel D'yachkov	<i>Kurnakov Institute for General and Inorganic Chemistry, Russian Academy of Sciences, Russian Federation</i>
Prof. Stefano Bellucci	<i>Frascati National Laboratories – National Institute of Nuclear Physics, Italy</i>
Prof. Arnold Kiv	<i>Ben-Gurion University of the Negev, Israel</i>
Prof. Alytis Gruodis	<i>Vilnius University, Lithuania</i>
Prof. Michael Schenk	<i>Fraunhofer Institute for Factory Operation and Automation IFF, Germany</i>
Prof. Dietmar Fink	<i>University of Mexico, United Mexican States</i>
Prof. Ravil Muhamedyev	<i>International IT University, Kazakhstan</i>
Prof. Eugene Kopytov	<i>Transport & Telecommunication Institute, Latvia</i>
Prof. Kurt Schwartz	<i>Gesellschaft für Schwerionenforschung mbH, Darmstadt, Germany</i>
Prof. Eva Rysiakiewicz-Pasek	<i>Institute of Physics, Wroclaw University of Technology, Poland</i>
Literary Editor	Prof. Tamara Lobanova-Shunina, <i>Riga Technical University, Latvia</i>
Technical Editor , secretary of Editorial Board	MSc Comp Nataly Burluckaya, <i>Information Systems Management Institute, Latvia</i>

Journal topics:	Host Organization	Supporting Organizations
mathematical and computer modelling computer and information technologies natural and engineering sciences innovative technologies operation research and decision making nanoscience and nanotechnologies operation research and decision making innovative education	Transport and Telecommunication Institute	Latvian Transport Development and Education Association Latvian Academy of Sciences Latvian Operations Research Society
Articles should be submitted in English . All articles are reviewed.		

EDITORIAL CORRESPONDENCE	COMPUTER MODELLING AND NEW TECHNOLOGIES, 2014, Vol. 18, No.3 ISSN 1407-5806, ISSN 1407-5814 (on-line: www.tsi.lv)
Transport and Telecommunication Institute 1 Lomonosova, Bld 4, LV-1019, Riga, Latvia Phone: (+371) 67100593 Fax: (+371) 67100535 E-mail: cm&nt@tsi.lv www.tsi.lv	Scientific and research journal The journal is being published since 1996 The papers published in Journal 'Computer Modelling and New Technologies' are included in: INSPEC (since 2010) , www.theiet.org/resources/inspec/ VINITI (since 2011) , http://www2.viniti.ru/ CAS Database http://www.cas.org/ EI Compindex



Content

Editors' Remarks	5
-------------------------	---

Mathematical and Computer Modelling

Yurii Kozachenko, Anatolii Pashko	Accuracy of Simulations of the Gaussian random processes with continuous spectrum	7
Shuren Wang, Junqing Su, Paul Hagan	Energy dissipation characteristics of sandstone cutting under mechanical impact load	13
Zhenzhen Jia, Feng Tao	Gas explosion characteristics and its control technologies in closed fire zone	21
G Wang, H Y Wang, Q M Huang, C Q Su	Numerical modelling of rock cross-cut coal uncovering based on ANSYS	26
Hua Zhang, Xiaoping Jiang, Chenghua Li	UKF-based underground intrusion localization algorithm for optical-fibre sensing perimeter protection	33
Junxia Zhang, Hongxing Zhao	A novel compressed air solar energy photo-thermal generating electricity system	39
Zhi Hui Wen, Jian Ping Wei, Hong Tu Zhang, Shao Hua Dai	An experimental study on coal gas desorption laws with different particle size	44
Zhigang Yin, Guohui Gao	Application of fractal theory to dam deformation forecast	49
Yinglin Li, Lianhe Yang, Suying Chen, Lei Xu	Three dimensional simulation of weft knitted fabric based on surface model	52
Chang Zhi Zhu, Quan Chen Gao	Three-dimensional computer simulation of soil nailing support in deep foundation pit	58
Bing Bai, Xiuquan Deng, Zhiqiong Guo, Dehua Gao	Organizations learning mechanism in the cyber society ecology system: an agent-based simulation	63
Xing-bin Ma, Cui-pin Jiang	The approach of fixed asset management based on the shortest path	68
Xuhong Guo, Xing Huang, Shaohua Wu	Correlation between austempering parameters and hardness of austempered ductile iron based on artificial neural network	72
Chih-Feng Chao, Ming-Huwi Horng	Firefly algorithm for training the radial basis function network in ultrasonic supraspinatus image classification	77
Shiqing Yang, Lanfang Miao	Material simulation based on Phong illumination model	84
Li Jun Tan, Bo Fang, Ming Ming Li, Ye Tang, Wen Hu Huang	Parametric identification for GHM and application of viscoelastic damper	88
Longmiao Chen, Qiang Fu, Gui Lin	Study on the sealing properties of the sealing structure for the rotating chamber of a certain cased telescoped ammunition gun	93

Information Technologies

Xin Pan, Hongbin Sun	A self-adaptive selective method of remote sensing image classification algorithms	98
Hong Jiang, Wenlei Sun, Yongfang Shi, Yanhua Huang	A study and implementation on the data reduction based on the curvature of point clouds	104
Xin Zhang	Multi-objective improved algorithm for flow allocations in hazardous chemicals logistics preference paths	111
Wei Wang, Xiao-dan Huang	Spatially aware in implicit human robot interaction	115
Hongqing Hou, Qian Miao, Chuanqiang Yu, Qinhe Gao	Study on T-S fuzzy sliding mode control based on a new reaching law	124
Lei Zhao, Guangxue Chen	A printer reverse characterization model based on BP neural network	133
LiLi Chu, LiLi Qin	Communication technology in the application of the smart grid	144
Ling Zhou, De Feng Zhang	A distributed multicast routing algorithm based on bone node set for mobile IP	151
Bing Xu	The study of campus network traffic monitoring platform	159
Fu Zhang, Yakun Zhang, Binbin Yue, Guoying Zhang	The gait analysis on the sloping walking of goat	164

Shang Zhang, Tingyan Xing	RSSI Enhanced indoor LBS platform design	170
Zhaohua Liu, Yuxia Yang	Study on semi-global matching algorithm extended for multi baseline matching and parallel processing method based on GPU	174
Rui-Lin Lin	Industrial product innovative design of toilet sensor timer	179
Hongkai Li, Zhendong Dai	Mechanism design and flow estimation method of a hydraulic actuated robot	183
Maoyi Tian, Rufeil Liu, Xiushan Lu	Ground point filtering method of vehicle-borne laser point cloud in urban street	188
Beibut Amirgaliyev, Magzhan Kairanbay, Chingiz Kenshimov, Didar Yedilkhan	Development of automatic number plate recognition system	193
Song Fei, Cui Zhe	Study on HDFS improvement scheme based on the GE code and dynamic replication strategy	198
Zheng Wang, Zhenjiang Miao	Design a media art installation based on fuzzy controlling system	204

Operation research and decision making

Shilong Li, Hongyan Tian	Decision-making model of the urban regeneration construction project based on environment improvement	209
Q X Zhang, Y P Wu, G Zh Ou, X G Fan, J H Zhou	Displacement prediction of liangshuijing landslide based on time series additive model	215
Liu Wanli	Study on the decision value of analysts' recommendations	224
Bao Xing	A dual capacity sourcing model of disruption management for an injured power system	231
Jianxin Bi, Lianghai Lei	Analysis on the dynamic effects of the aggregate supply, aggregate demand and macroeconomic policies of China based on SVAR model	237
Qiong Gu, Wu Zheng, Xianming Wang	Study on Xiangyang's population and aging trend prediction based on discrete population development equation model	249
Dong-ping Li, Kong Xiangsheng	A Study on Fast Assessment of Medium and Small Earthquake	254
Wan He	Deep neural network based load forecast	258
Junxiang Tu	Retrieving product information of collaborative enterprises based on Bayesian network	263
Li Wu, Xinyuan Wang	Geoinformatics-based study on the regionalization of ecological function in the Chaohu Lake Basin, East China	267
Yue Wei Ai, Yan He, Zhi Jian Wang, Yang Wang	A new method of digital manufacturing of orthoses	271
Wei Sun, Yang Yu	Evolutionary game and simulation of organizational information security investment	276

Innovative Education

Laiquan Liu, Li Lei, Yanrui Lei	The application of fuzzy association rules in the employment data mining of a higher vocational college	283
Jia Geng	Information technology-based promotion of educational resource sharing	288

Nanoscience and Nanotechnology

Xia Ma, Mian Mian Zhang, Shi Wen Chen	Optimized preparation of γ -polyglutamic acid/chitosan nanocapsule	293
---------------------------------------	---	-----

Authors' Index

299

Cumulative Index

300

Editors' Remarks

A General Summary

by Rudyard Kipling

We are very slightly changed
From the semi-apes who ranged
India's Prehistoric clay;
He that drew the longest bow
Ran his brother down, you know,
As we run men down to-day.

"Dowb," the first of all his race,
Met the Mammoth face to face
On the lake or in the cave:
Stole the steadiest canoe,
Ate the quarry others slew,
Died - and took the finest grave.

When they scratched the reindeer-bone,
Some one made the sketch his own,
Filched it from the artist -- then,
Even in those early days,
Won a simple Viceroy's praise
Through the toil of other men.
Ere they hewed the Sphinx's visage
Favouritism governed kissage,
Even as it does in this age.

Who shall doubt "the secret hid
Under Cheops' pyramid"
Was that the contractor did
Cheops out of several millions?
Or that Joseph's sudden rise
To comptroller of Supplies
Was a fraud of monstrous size
On King Pharaoh's swart Civilians?

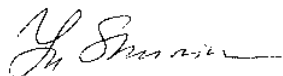
Thus, the artless songs I sing
Do not deal with anything
New or never said before.
As it was in the beginning
Is to-day official sinning,
And shall be for evermore!

Rudyard Kipling (1809-1849) *

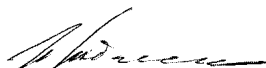
This 18th volume No.3 presents actual papers on main topics of Journal specialization, namely, **Nanoscience and Nanotechnologies, Mathematical and Computer Modelling, Information and Computer Technologies, Operation Research and Decision Making and Innovative Education.**

Our journal policy is directed on the fundamental and applied sciences researches, which are the basement of a full-scale modelling in practice. This edition is the continuation of our publishing activities. We hope our journal will be interesting for research community, and we are open for collaboration both in research and publishing. We hope that journal's contributors will consider the collaboration with the Editorial Board as useful and constructive.

EDITORS



Yuri Shunin



Igor Kabashkin

* **Joseph Rudyard Kipling** (30 December 1865 – 18 January 1936) was an English short-story writer, poet, and novelist. He is chiefly remembered for his tales and poems of British soldiers in India and his tales for children. He was born in Bombay, in the Bombay Presidency of British India, and was taken by his family to England when he was five years old. Kipling is best known for his works of fiction, including *The Jungle Book* (a collection of stories, which includes and his poems, including "Mandalay" (1890), "Gunga Din" (1890), "The Gods of the Copybook Headings" (1919), "The White Man's Burden" (1899), and "If—" (1910). He is regarded as a major "innovator in the art of the short story"; his children's books are enduring classics of children's literature; and his best works are said to exhibit "a versatile and luminous narrative gift".



Accuracy of Simulations of the Gaussian random processes with continuous spectrum

Yurii Kozachenko, Anatolii Pashko*

Taras Shevchenko National University of Kyiv, Prospect Hlushkov, 4D, 03187 Kyiv, Ukraine

Received 1 March, 2014, www.tsi.lv

Abstract

This paper investigates algorithms for the construction of sub-Gaussian models for the Gaussian stationary random processes with continuous spectrum. Estimates for random processes with analytical correlation functions retrieved and improved existing ones. Algorithms for simulation of random processes with given accuracy and reliability in various function spaces were constructed.

Keywords: Gaussian process, simulation, sub-Gaussian model, model accuracy, model reliability

1 Introduction

In this paper, we continue to examine algorithms for constructing sub-Gaussian models for Gaussian stationary random processes and fields [1-5]. To construct models of stochastic processes we use their spectral image in the form of stochastic integrals. In [6-9] we studied the modelling techniques and conditions of weak convergence models. Some works suggested to use the assessment points of difference for the process and model to assess the accuracy of the simulation. Papers [3-4] studied the rate of convergence of sub-Gaussian patterns in different functional spaces. It is possible to construct algorithms for simulation of random processes and fields with given accuracy and reliability for various functional spaces.

Strictly sub-Gaussian random spectral expansions of random processes was proposed to examine as the model. Sub-Gaussian properties of random processes researches are given in [10-11].

When constructing models of Gaussian random processes with given accuracy and reliability in various functional spaces we need to assess growth of $F(\infty) - F(\Lambda)$, where $\Lambda > 0$, $F(\Lambda)$ - the spectral function of the process. If the spectral function is known, it is not difficult. Another situation where the only correlation function is known and the spectral function cannot be found explicitly.

Improved estimation for stationary Gaussian processes with continuous spectrum with correlation functions with desired properties was obtained in [1-5].

2 Basic concepts and definitions

Let (Ω, \mathcal{B}, P) - be a standard probability space.

Definition 1. A random variable ξ is called sub-Gaussian, if $E\xi = 0$ and exists $a \geq 0$ such that $E \exp\{\lambda\xi\} \prec \exp\left\{\frac{\lambda^2 a^2}{2}\right\}$ for all $\lambda \in R^1$.

A space of sub-Gaussian variables $Sub(\Omega)$ is Banach relative to the following norm

$$\tau(\xi) = \sup_{\lambda \neq 0} \left[\frac{2 \ln E \exp\{\lambda\xi\}}{\lambda^2} \right]^{\frac{1}{2}}.$$

If $\tau(\xi) = E\xi^2$ - called strictly sub-Gaussian random variable.

Let $\xi(t)$ - be a real Gaussian stationary random process with $E\xi(t) = 0$, $R(\tau)$ - correlation function of $\xi(t)$, $F(\lambda)$ - the spectral function of the process $\xi(t)$,

$$R(\tau) = \int_0^{\infty} \cos(\lambda\tau) dF(\lambda).$$

A random process be

$$\text{represented as } \xi(t) = \int_0^{\infty} \cos(\lambda t) d\xi_1(\lambda) + \int_0^{\infty} \sin(\lambda t) d\xi_2(\lambda),$$

where $\xi_1(t)$ and $\xi_2(t)$ - the centered and uncorrelated random processes with uncorrelated increments such as $0 \leq \lambda_1 < \lambda_2$ and $E(\xi_1(\lambda_2) - \xi_1(\lambda_1))^2 = E(\xi_2(\lambda_2) - \xi_2(\lambda_1))^2 = F(\lambda_2) - F(\lambda_1)$.

Let D_Λ - be some partition of the interval $[0, \Lambda]$, $D_\Lambda : 0 = \lambda_0 < \lambda_1 < \dots < \lambda_n = \Lambda$. The model of random process $\xi(t)$ can be obtained as

$$S_n(t, \Lambda) = \sum_{i=0}^{n-1} \left[\cos(\lambda_i t) (\xi_1(\lambda_{i+1}) - \xi_1(\lambda_i)) + \sin(\lambda_i t) (\xi_2(\lambda_{i+1}) - \xi_2(\lambda_i)) \right].$$

* Corresponding author e-mail: pashkoua@mail.ru

The model of process $\xi(t)$ can be obtained by modelling the sum $\sum_{i=0}^{n-1} [\cos(\lambda_i t) \eta_{1i} + \sin(\lambda_i t) \eta_{2i}]$, where $\{\eta_{1i}, \eta_{2i}\}$ - centered uncorrelated Gaussian random variables with $E(\eta_{1i})^2 = E(\eta_{2i})^2 = F(\lambda_{i+1}) - F(\lambda_i)$. Given the precision of work with real numbers and algorithms errors for simulation Gaussian random variables, lets consider $\{\eta_{1i}, \eta_{2i}\}$ - be a sequence of strictly sub-Gaussian uncorrelated random variables. Let $\xi(t)$ and all $S_n(t, \Lambda)$ belongs to certain functional Banach space $A(T)$ with norm of $\|\cdot\|$. Let the two numbers be as follow $\delta > 0$ and $0 < \alpha < 1$.

Definition 2. Model $S_n(t, \Lambda)$ approximates the process $\xi(t)$ with reliability $1 - \varepsilon$ and accuracy δ in the norm of space $A(T)$, if the following inequality holds $P\{\|\xi(t) - S_n(t, \Lambda)\| > \delta\} \leq \varepsilon$.

3 Main results

Theorem 1. Model $S_n(t, \Lambda)$ approximates the process $\xi(t)$ with reliability $1 - \varepsilon$ and accuracy δ in the norm of process $L_2(T)$, if for the numbers Λ and n the following inequalities hold $B_{n,\Lambda} < \delta^2$ and $\exp\left\{\frac{1}{2}\right\} \frac{\delta}{\sqrt{B_{n,\Lambda}}} \exp\left\{-\frac{\delta^2}{2B_{n,\Lambda}}\right\} \leq \varepsilon$, where

$$B_{n,\Lambda} = \int_T E(\xi(t) - S_n(t, \Lambda))^2 d\mu(t).$$

The proof of the theorem can be found in [3-5].

When for the correlation function $R(\tau)$ and spectral function $F(\lambda)$ of process $\xi(t)$ performs $R(0) = F(\infty) = 1$ (in the general case we can assume $R(\tau) = \frac{R(\tau)}{R(0)}$ and $F(\lambda) = \frac{F(\lambda)}{F(\infty)}$), then $R(\tau)$ can be considered as the characteristic function of some random variable with symmetric distribution function $G(\lambda)$ and when $\lambda > 0$ we have $F(\lambda) = G(\lambda) - G(-\lambda)$.

The following assertion holds.

Theorem 2. Let's assume that for the correlation function $R(\tau)$ and for some t_0 when $|t| < t_0$ the following condition holds $1 - \frac{R(t)}{R(\infty)} \leq \sum_{j=1}^K \psi_j(t)$, where $\psi_j(t)$, $j=1, 2, \dots, K$ - monotone nondecreasing functions, $\psi_j(0) = 0$, such as, when $|uv| < t_0$, $\psi_j(uv) \leq \psi_j^1(u)\psi_j^2(u)$, where $\psi_j^i(u) \geq 0, i=1, 2, j=1, 2, \dots, K$ - monotone

nondecreasing functions, $\psi_j^i(0) = 0$, then following inequalities hold

- 1) if $h > \frac{2}{t_0}$, then $1 - F(h) \leq \frac{2}{\pi} \sum_{j=1}^K H_{0,\psi_j} \left(\frac{h}{2}\right)$,
- 2) if $h > \frac{3}{t_0}$, then $1 - F(h) \leq \frac{8}{3\pi} \sum_{j=1}^K H_{1,\psi_j} \left(\frac{h}{3}\right)$,
- 3) if $h > \frac{4}{t_0}$, then $1 - F(h) \leq \frac{3}{\pi} \sum_{j=1}^K H_{2,\psi_j} \left(\frac{h}{4}\right)$,

$$H_{n,\psi_j}(h) = \psi_j^1\left(\frac{1}{h}\right) \left(\int_0^1 \psi_j^2(u) du + \int_1^{h_0} \frac{\psi_j^2(u)}{u^{n+2}} du \right) + \frac{2}{(n+1)(ht_0)^{n+1}}$$

The theorem is a general result of Lemma 4 and Theorem 1 from [2].

Corollary 1. Let assume that for the correlation function $R(\tau)$ and some t_0 when $|t| < t_0$ the following condition holds $1 - \frac{R(t)}{R(\infty)} \leq C_\alpha |t|^\alpha, 0 \leq \alpha \leq 2$. Then for

$1 - \frac{F(\lambda)}{F(\infty)}$ we have the following estimates

1) if $h > \frac{2}{t_0}, 0 \leq \alpha < 1$,
 then $1 - \frac{F(h)}{F(\infty)} \leq \frac{2}{\pi} H'_{0,\alpha} \left(\frac{h}{2}\right)$, (1)

2) if $h > \frac{3}{t_0}, 0 \leq \alpha < 2$,
 then $1 - \frac{F(h)}{F(\infty)} \leq \frac{8}{3\pi} H'_{1,\alpha} \left(\frac{h}{3}\right)$, (2)

3) if $h > \frac{4}{t_0}, 0 \leq \alpha \leq 2$,
 then $1 - \frac{F(h)}{F(\infty)} \leq \frac{3}{\pi} H'_{2,\alpha} \left(\frac{h}{4}\right)$, (3)

$$H'_{n,\alpha}(h) = \frac{C_\alpha}{h^\alpha} \left(\frac{1}{\alpha+1} + \frac{1}{n+1-\alpha} \right) + \frac{1}{(ht_0)^{n+1}} \left(\frac{2}{n+1} - \frac{C_\alpha t_0^\alpha}{n+1-\alpha} \right)$$

The proof follows directly from Theorem 2. These estimates depend on α , so, having small α (1) is a better estimate, but when $\alpha > 1$ then estimate (3) is better. Estimate (2) should be compared in each particular case. Obtained results can be used for evaluation $F(\infty) - F(\Lambda)$.

Suppose that the correlation function $R(\tau)$ has derivatives of order $4k, k=1, 2, \dots$. Since

$R^{(4k)}(\tau) = (-1)^{2k} \int_0^\infty \lambda^{4k} \cos(\lambda\tau) dF(\lambda)$, then $R^{(4k)}(\tau)$ – the correlation function of a random process with a spectral function $F_k(\lambda) = \int_0^\lambda u^{4k} dF(u)$.

There is an inequality $F_k(\infty) - F_k(\Lambda) = \int_\Lambda^\infty u^{4k} dF(u) \geq \Lambda^{4k} (F(\infty) - F(\Lambda))$, then $F(\infty) - F(\Lambda) \leq \frac{1}{\Lambda^{4k}} (F_k(\infty) - F_k(\Lambda))$.

Corollary 2. Suppose $k=1,2,\dots$ for the correlation function $R^{(4k)}(\tau)$ and for some t_0 when $|t| < t_0$ the following condition holds $1 - \frac{R^{(4k)}(t)}{R^{(4k)}(\infty)} \leq \sum_{j=1}^K \psi_{j,k}(t)$, where $\psi_{j,k}(t), j=1,2,\dots,K$ satisfies the conditions of the theorem 2. Then for $1 - \frac{F_k(\lambda)}{F_k(\infty)}$ we have

- 1) if $h > \frac{2}{t_0}$, then $F_k(\infty) - F_k(h) \leq \frac{2}{\pi} F_k(\infty) \sum_{j=1}^K H_{0,\psi_j} \left(\frac{h}{2} \right)$,
 - 2) if $h > \frac{3}{t_0}$, then $F_k(\infty) - F_k(h) \leq \frac{8}{3\pi} F_k(\infty) \sum_{j=1}^K H_{1,\psi_j} \left(\frac{h}{3} \right)$,
 - 3) if $h > \frac{4}{t_0}$, then $F_k(\infty) - F_k(h) \leq \frac{3}{\pi} F_k(\infty) \sum_{j=1}^K H_{2,\psi_j} \left(\frac{h}{4} \right)$, where
- $$H_{n,\psi_j}(h) = \psi_j^n \left(\frac{1}{h} \right) \left(\int_0^1 \psi_j^2(u) du + \int_1^{ht_0} \frac{\psi_j^2(u)}{u^{n+2}} du \right) + \frac{2}{(n+1)(ht_0)^{n+1}}$$

The proof follows directly from Theorem 2.

More accurate estimates can be obtained when $R(\tau)$ is analytical. Let ζ be a positive probability unit, $F(x)$ – its distribution function and $\phi(\tau) = \int_0^\infty \cos(u\tau) dF(u)$ – the characteristic function.

Lemma 1. If $\phi(\tau)$ analytic in some district of zero and there are numbers $L > 0$ and $\alpha > 0$, such as for all $k \geq 1$ the following condition holds $|\phi^{(2k)}(0)| \leq (Lk)^\alpha$,

then in case of $h > (Le)^{\frac{\alpha}{2}}$ the following inequality holds

$$1 - F(h) = P\{\zeta \geq h\} \leq \exp\left\{-\frac{h^{\frac{2}{\alpha}}}{2Le}\right\}.$$

Proof. From the Lyapunov inequality follows that when $2k - 2 \leq s \leq 2k, k \geq 1$ we have

$$(E\zeta^s)^{\frac{1}{s}} \leq (E\zeta^{2k})^{\frac{1}{2k}}. \quad \text{Whereas,}$$

$E\zeta^{2k} = |\phi^{(2k)}(0)| = \int_0^\infty x^{2k} dF(x)$ then when $s \geq 1$ there are relation

$$E\zeta^s \leq (E\zeta^{2k})^{\frac{1}{2k}} \leq (Lk)^{\frac{\alpha s}{2}} \leq (Ls)^{\frac{\alpha s}{2}} \left(\frac{k}{s}\right)^{\frac{\alpha s}{2}} \leq (Ls)^{\frac{\alpha s}{2}}.$$

So, when $s \geq 1$ we have $E\zeta^s \leq (Ls)^{\frac{\alpha s}{2}}$. From this inequality and Chebyshev inequality implies that $h > 0, s > 1$ the following irregularities take place

$$P\{\zeta > h\} \leq \frac{E\zeta^s}{h^s} \leq \left(\frac{Ls}{h^{\frac{2}{\alpha}}}\right)^{\frac{\alpha s}{2}}. \quad \text{Minimizing the right side by } s, \text{ in this case } s = \frac{h^{\frac{2}{\alpha}}}{Le}, \text{ we obtain the required inequality.}$$

Let $F(u), u \geq 0$ – be a spectral function of a stationary process. Since function $\frac{F(u)}{F(\infty)}$ can be considered as a function of the distribution of positive random variable ζ , then we have the following theorem.

Theorem 3. Let the correlation function $R(\tau)$ of a stationary random process $\xi(t)$ be analytic in some district of zero and there are numbers $L > 0$ and $m > 0$, such as for all $k \geq 1$ we have the following estimate $|R^{(2k)}(0)| \leq (Lk)^{mk}$, then when $h > (Le)^{\frac{m}{2}}$ the following

$$\text{inequality holds } 1 - F(h) \leq \exp\left\{-\frac{h^{\frac{2}{m}} m}{2Le}\right\}.$$

4 Examples

Consider a stationary random process $\xi(t)$ with correlation function $R(\tau) = A \exp\{-C|\tau|\}$, where $A > 0, C > 0$. Spectral function $F(\lambda) = \frac{A}{\pi} \arctg\left(\frac{\lambda}{C}\right)$

$$\text{and } F(\infty) - F(\Lambda) = \frac{A}{\pi} \left(\frac{\pi}{2} - \arctg\left(\frac{\Lambda}{C}\right) \right).$$

Let for stationary random process $\xi(t)$ correlation function $R(\tau) = A \exp\{-C|\tau|\} \cos(\beta\tau)$, where $A > 0$,

$C > 0$ and $\beta > 0$. Spectral function

$$f(\lambda) = \frac{A}{\pi} \frac{C(C^2 + \beta^2 + \lambda^2)}{(\lambda^2 - \beta^2 - C^2)^2 + 4C^2\lambda^2}. \text{ For } F(\infty) - F(\Lambda)$$

the following holds

$$F(\infty) - F(\Lambda) = \frac{AC}{2\pi} \int_{\Lambda}^{\infty} \left(\frac{1}{C^2 + (\beta + x)^2} + \frac{1}{C^2 + (\beta - x)^2} \right) dx = \frac{A}{2\pi} \left(\pi - \arctg\left(\frac{\Lambda + \beta}{C}\right) - \arctg\left(\frac{\Lambda - \beta}{C}\right) \right)$$

Consider a stationary random process $\xi(t)$ with correlation function $R(\tau) = A_{\alpha} \exp\{-C|\tau|^{\alpha}\} \cos(\beta\tau)$, where $A_{\alpha} > 0$, $C > 0$, $\beta > 0$ and $1 < \alpha \leq 2$. Let us estimate $F(\infty) - F(\Lambda)$ for the corresponding spectral function. Set $R(0) = F(\infty) = A_{\alpha} = 1$.

Let $1 < \alpha < 2$, since when $x > 0$ the following inequalities $\exp\{-x\} \geq 1 - x$ and $\cos(x) \geq 1 - \frac{x^2}{2}$, and

when $|\tau| \leq t_0 = \min\left(C^{\frac{1}{\alpha}}, \sqrt{2}\beta^{-1}\right)$ we have

$$1 - R(\tau) \leq 1 - \left(1 - C|\tau|^{\alpha}\right) \left(1 - \frac{\beta^2\tau^2}{2}\right) = C|\tau|^{\alpha} + \frac{\beta^2\tau^2}{2} - C|\tau|^{\alpha} \left(\frac{\beta^2\tau^2}{2}\right) \leq C|\tau|^{\alpha} + \frac{\beta^2\tau^2}{2}.$$

Thus, according to corollary 1, when $h > \frac{4}{t_0}$ we have

$$1 - F(h) \leq \frac{3}{\pi} \left(H'_{2,\alpha}\left(\frac{h}{4}\right) + H'_{2,2}\left(\frac{h}{4}\right) \right), \text{ where}$$

$$H'_{2,2}(a) = \frac{3\beta^2}{4a^2} \text{ and } H'_{2,\alpha}(a) = \frac{C}{a^{\alpha}} \left(\frac{1}{\alpha+1} + \frac{1}{3-\alpha} \right).$$

That is, when $h > \frac{4}{t_0}$ we have

$$F(\infty) - F(h) \leq \frac{3A_{\alpha}}{\pi} \left(H'_{2,\alpha}\left(\frac{h}{4}\right) + H'_{2,2}\left(\frac{h}{4}\right) \right) = \frac{3A_{\alpha}}{\pi} \left(\frac{4^{\alpha}C}{h^{\alpha}} \left(\frac{1}{\alpha+1} + \frac{1}{3-\alpha} \right) + \frac{12\beta^2}{h^2} \right).$$

Let $\alpha = 2$. Note that $R(\tau) = \exp\{-C|\tau|^2\} \cos(\beta\tau)$ - is the product of two characteristic functions: $\psi_1(\tau) = \exp\{-C|\tau|^2\}$ - characteristic function of a random variable ζ_1 , which has a normal distribution from $E\zeta_1 = 0$, $E\zeta_1^2 = \frac{1}{2C}$, and $\psi_2(\tau) = \cos(\beta\tau)$ -

characteristic function of a random variable ζ_2 with a distribution law $P\{\zeta_2 = \beta\} = \frac{1}{2}$ and $P\{\zeta_2 = -\beta\} = \frac{1}{2}$.

So $R(\tau)$ - is the characteristic function of a random variable $\zeta = \zeta_1 + \zeta_2$, where ζ_1 and ζ_2 are independent. That is, if $\lambda > 0$

$$1 - F(h) = P\{|\zeta| > h\} = P\{|\zeta_1 + \zeta_2| > h\} = \frac{P\{|\zeta_1 + \beta| > h\}}{2} + \frac{P\{|\zeta_1 - \beta| > h\}}{2}.$$

If $h > \beta$, it is easy to see

$$1 - F(h) = \frac{1}{2} \left(P\{|\zeta_1| > h - \beta\} + P\{|\zeta_1| > h + \beta\} \right) = \left(\frac{C}{\pi} \right)^{\frac{1}{2}} \left(\int_{\lambda-h}^{\infty} \exp\{-t^2 C\} dt + \int_{\lambda+h}^{\infty} \exp\{-t^2 C\} dt \right) = \left(\frac{1}{\pi} \right)^{\frac{1}{2}} \left(\int_{(h-\beta)\sqrt{C}}^{\infty} \exp\{-u^2\} du + \int_{(h+\beta)\sqrt{C}}^{\infty} \exp\{-u^2\} du \right).$$

$$1 - F(h) \leq \frac{1}{\sqrt{\pi}} \left(\frac{2}{(h-\beta)\sqrt{C}} \exp\{-C(h-\beta)^2\} + \frac{2}{(h+\beta)\sqrt{C}} \exp\{-C(h+\beta)^2\} \right).$$

So, That is, when $h > \beta$ we have the following inequality

$$F(\infty) - F(h) \leq \frac{2A_{\alpha}}{\sqrt{\pi C}} \left(\frac{2}{(h-\beta)\sqrt{C}} \exp\{-C(h-\beta)^2\} + \frac{2}{(h+\beta)\sqrt{C}} \exp\{-C(h+\beta)^2\} \right).$$

5 Estimation of model parameters

Let the random process $\xi(t)$ be defined on the interval $[0, T]$, $T > 0$. Model of random process $\xi(t)$ built as

$$S_n(t, \Lambda) = \sum_{i=0}^{n-1} [\cos(\lambda_i t) \eta_{1i} + \sin(\lambda_i t) \eta_{2i}], \text{ where}$$

$D_{\Lambda} : 0 = \lambda_0 < \lambda_1 < \dots < \lambda_n = \Lambda$ - uniform partition of the interval $[0, \Lambda]$, and $\{\eta_{1i}, \eta_{2i}\}$ - strictly sub-Gaussian independent random variables with $E(\eta_{1i})^2 = E(\eta_{2i})^2 = F(\lambda_{i+1}) - F(\lambda_i)$.

Let's find such Λ and M , that the selected model approximates the process $\xi(t)$ with given accuracy δ and reliability $1 - \varepsilon$ in $L_2(T)$.

By the theorem 1 for the unknown Λ and M the following inequalities must hold $B_{M,\Lambda} \leq \delta^2$ and

$B_{M,\Lambda} \leq \frac{\delta^2}{s_\varepsilon}$, where s_ε - the root of the equation $\exp\left\{\frac{1}{2} - \frac{s^2}{2}\right\} s = \varepsilon$.

For random process $\xi(t)$ with correlation function $R(\tau) = A \exp\{-C|\tau|\}$, where $A > 0$, $C > 0$ and uniform partition of the interval $[0, \Lambda]$ the following holds $B_{n,\Lambda} \leq \frac{T^3 \Lambda^2}{3M^2} F(\Lambda) + T(F(\infty) - F(\Lambda))$. Hence,

$$B_{M,\Lambda} \leq \frac{T^3 \Lambda^2}{3M^2} F(\Lambda) + \frac{TA}{\pi} \left(\frac{\pi}{2} - \arctg\left(\frac{\Lambda}{C}\right) \right) \quad \text{and}$$

$$E(\eta_{li})^2 = E(\eta_{2i})^2 = \frac{A}{\pi} \left(\arctg\left(\frac{\lambda_{i+1}}{C}\right) - \arctg\left(\frac{\lambda_i}{C}\right) \right).$$

For random process $\xi(t)$ with correlation function $R(\tau) = A \exp\{-C|\tau|\} \cos(\beta\tau)$, where $A > 0$, $C > 0$, $\beta > 0$ and uniform partition of the interval $[0, \Lambda]$ the following holds

$$B_{M,\Lambda} \leq \frac{T^3 \Lambda^2}{3M^2} F(\Lambda) + \frac{TA}{2\pi} \left(\pi - \arctg\left(\frac{\Lambda + \beta}{C}\right) - \arctg\left(\frac{\Lambda - \beta}{C}\right) \right)$$

and $E(\eta_{li})^2 = E(\eta_{2i})^2 = \frac{A}{2\pi} \left(\arctg\left(\frac{\lambda_{i+1} + \beta}{C}\right) + \arctg\left(\frac{\lambda_{i+1} - \beta}{C}\right) - \arctg\left(\frac{\lambda_i + \beta}{C}\right) - \arctg\left(\frac{\lambda_i - \beta}{C}\right) \right).$

For stationary random process $\xi(t)$ with correlation function $R(\tau) = A_\alpha \exp\{-C|\tau|^\alpha\} \cos(\beta\tau)$, where $A_\alpha > 0$, $C > 0$, $\beta > 0$, $1 < \alpha \leq 2$ and

$$E(\eta_{li})^2 = F(u_{i+1}) - F(u_i) = \frac{2}{\pi} \int_{u_i}^{u_{i+1}} \int_0^\infty \cos(v\tau) R(\tau) d\tau dv =$$

$$= \frac{2A_\alpha}{\pi} \int_0^\infty \left(\sin\left(\frac{(i+1)\Lambda\tau}{M}\right) - \sin\left(\frac{i\Lambda\tau}{M}\right) \right) \times$$

$$\exp\left\{-C|\tau|^\alpha\right\} \frac{\cos(\beta\tau)}{\tau} d\tau$$

Consider Λ , when $\Lambda \geq \Lambda_0 = \max\left(1, \frac{1}{t_0}\right)$, where

$$t_0 = \min\left(C^{-\frac{1}{\alpha}}, \sqrt{2}\beta^{-1}\right). \quad \text{Then we have}$$

$$F(\infty) - F(\Lambda) \leq \frac{A_\alpha L_\alpha}{\Lambda^\alpha}, \quad \text{where}$$

$$L_\alpha = \frac{3}{\pi} \left(\frac{C4^{\alpha+1}}{(\alpha+1)(3-\alpha)} + 12\beta^2 \right).$$

Therefore, $B_{M,\Lambda} \leq \frac{T^3 \Lambda^2}{3M^2} F(\infty) + \frac{TA_\alpha L_\alpha}{\Lambda^\alpha}$.

Minimizing the right side by Λ , so

$$\Lambda_M = \left(\frac{3L_\alpha \alpha M^2}{2T^2} \right)^{\frac{1}{2+\alpha}}, \quad \text{we get the inequality}$$

$$B_{M,\Lambda} \leq \frac{T^{\frac{3\alpha+2}{2\alpha}}}{M^{\frac{2+\alpha}{2\alpha}}} \left(\frac{2A_\alpha}{3} \right)^{\frac{\alpha}{2+\alpha}} (L_\alpha A_\alpha \alpha)^{\frac{2}{2+\alpha}} (\alpha + 2).$$

So, for M the inequality must holds

$$M \geq \frac{T^{\frac{3\alpha+2}{2\alpha}} A_\alpha^{\frac{1}{2} + \frac{1}{\alpha}} \left(\frac{2}{3}\right)^{\frac{1}{2}} (L_\alpha \alpha)^{\frac{1}{\alpha}} (\alpha + 2)^{\frac{\alpha+2}{2\alpha}}}{(\delta^2 \min(1, s_\varepsilon^{-1}))^{\frac{\alpha+2}{2\alpha}}} = N_1.$$

Lemma 2. Model $S_n(t, \Lambda)$ approximates the process $\xi(t)$ with reliability $1 - \varepsilon$ and accuracy δ in the norm $L_2([0, T])$, if $\Lambda = \Lambda_M$, and for M the following inequality holds $M \geq \max(N_1, N_2, N_3)$, where

$$N_2 = \Lambda_0^{\frac{\alpha+2}{2}} \left(\frac{2}{3L_\alpha \alpha} \right)^{\frac{1}{2}} T, \quad N_3 = \left(\frac{3}{2} L_\alpha \alpha \right)^{\frac{1}{\alpha}} T,$$

$$\Lambda_0 = \max\left(1, \frac{1}{t_0}\right), \quad t_0 = \min\left(C^{-\frac{1}{\alpha}}, \sqrt{2}\beta^{-1}\right).$$

Same results can be achieved in the space $L_p([0, T])$, $p > 2$.

Lemma 3. Model $S_n(t, \Lambda)$ approximates the process $\xi(t)$ with reliability $1 - \varepsilon$ and accuracy δ in the norm $L_p([0, T])$, $p > 2$, if $\Lambda = \Lambda_M$, and for M the following inequality holds $M \geq \max(N_4, N_5)$, where

$$N_4 = \frac{\sqrt{2} T^{\frac{\alpha+2+p\alpha}{2\alpha p}} A_\alpha^{\frac{1}{2} + \frac{1}{\alpha}} (L_\alpha \alpha)^{\frac{1}{\alpha}} (\alpha + 2)^{\frac{\alpha+2}{2\alpha}}}{(\delta^2 \min((p+1)^{-1}, s_\varepsilon^{-2}))^{\frac{\alpha+2}{2\alpha}}},$$

$$N_5 = \Lambda_0^{\frac{\alpha+2}{2}} \left(\frac{2}{L_\alpha \alpha} \right)^{\frac{1}{2}} T.$$

Proof. Let $\mu(T) = T$. For the unknown Λ and M

the following inequalities must hold $T^p D_{M,\Lambda} \leq \frac{\delta^2}{s_\varepsilon^2}$ and

$$T^p D_{M,\Lambda} \leq \frac{\delta^2}{p+1}. \quad \text{Consider } \Lambda, \quad \text{as}$$

$\Lambda \geq \Lambda_0 = \max\left(1, \frac{1}{t_0}\right)$, the following inequality must

hold $D_{M,\Lambda} \leq \frac{T^2 \Lambda^2}{M^2} F(\infty) + \frac{A_\alpha L_\alpha}{\Lambda^\alpha}$. Minimizing the

right side by Λ , i.e., if put $\Lambda = \Lambda_M = \left(\frac{L_\alpha \alpha M^2}{2T^2} \right)^{\frac{1}{2+\alpha}}$,

then we get inequality

$D_{M,\Lambda} \leq \frac{T^{\frac{\alpha+2}{2\alpha}}}{M^{\frac{2\alpha}{2+\alpha}}} (2A_\alpha)^{\frac{\alpha}{2+\alpha}} (L_\alpha A_\alpha \alpha)^{\frac{2}{2+\alpha}} (\alpha+2)$. So, for M the

following inequality must hold

$$M \geq \frac{\sqrt{2} T^{\frac{\alpha+2+p\alpha}{2\alpha}} A_\alpha^{\frac{1}{2+\alpha}} (L_\alpha \alpha)^{\frac{1}{2+\alpha}} (\alpha+2)^{\frac{\alpha+2}{2\alpha}}}{(\delta^2 \min((p+1)^{-1}, s_\alpha^{-2}))^{\frac{\alpha+2}{2\alpha}}} = N_4.$$

Given that $\Lambda_M \geq \Lambda_0$, we get the proof of the lemma.

In the real simulations, for example, for a given reliability $1 - \varepsilon$ and accuracy δ at $A = A_\alpha = 1$, $C = 1$ and $T = 1$ we have the following results (table 1) in $L_2([0, T])$.

TABLE 1 Estimates for the parameters M and Λ for the given accuracy and reliability in $L_2([0, T])$

δ	ε	β	α	M	Λ
0.1	0.05	0	1	17000	400
0.05				137000	1600
0.01				20000000	35000
0.1	0.05	1	1	24000	420
0.05				199000	150
0.1	0.05	1	1.1	175200	6925
0.05				1236000	24420
0.01				115300000	455500
0.1	0.05	1	1.5	27060	1010
0.05				136400	2535
0.01				5830000	21680
0.1	0.05	1	2	8080	281
0.05				32320	562
0.01				808000	2810

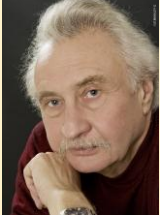

6 Conclusions

The paper shows estimates for increases of spectral functions of random processes. The results are used for the parameter estimations of sub-Gaussian models in the simulation of Gaussian random processes. Found assessment allow us to build a model with given accuracy and reliability in various function spaces, in particular $L_2([0, T])$, $L_p([0, T])$, $p > 2$.

References

- [1] Kozachenko Yu V, Pashko A A 1988 Simulation of Gaussian stationary stochastic processes represented by stochastic integrals *Theory and applications of statistical simulation* Novosibirsk 10-24
- [2] Kozachenko Yu V, Kozachenko L F 1992 On the modelling of Gaussian stationary processes with absolutely continuous spectrum *Probability Theory and Math. Statistics* 47 47-54
- [3] Kozachenko Yu V, Pashko A A 1998 The accuracy of the simulation of random processes in the Orlicz spaces I *Probability Theory and Math. Statistics* 58 75-90
- [4] Kozachenko Yu V, Pashko A A 1998 The accuracy of the simulation of random processes in the Orlicz spaces II *Probability Theory and Math. Statistics* 59 45-60
- [5] Pashko A A 2006 Numerical simulation of sub-Gaussian random fields *Bulletin of Taras Shevchenko National University in Kyiv. Physics and Mathematics* 1 35-39
- [6] Ermakov S M, Mikhailov G A 1982 *Statistical simulations* Nauka: Moscow p 296
- [7] Prigarin S M 1988 Simulation of Gaussian stationary stochastic processes represented by stochastic integrals *Theory and applications of statistical simulation* Novosibirsk 31-39
- [8] Prigarin S M 2001 *Spectral Models of Random Fields in Monte Carlo Methods* Utrecht: VSP p 195
- [9] Prigarin S M 2005 *Numerical Modelling of Random Processes and Fields* Novosibirsk p 259
- [10] Buldygin V V, Kozachenko Yu V 2000 *Metric characterization of random variables and random processes* Amer. Math. Soc., Providence, RI p 260
- [11] Kozachenko Yu V, Pashko A A 1999 *Simulation of random processes* Kyiv: T. Shevchenko University p 224

Authors

	<p>Yurii Kozachenko, born on December 1, 1940, Kyiv, Ukraine</p> <p>Current position, grades: professor, doctor of sci. in physics and mathematics University studies: Taras Shevchenko National University of Kyiv Scientific interest: Analytical properties of stochastic processes, distribution estimation of functionals from random processes, random processes in Orlicz spaces, pre-Gaussian and sub-Gaussian random processes, Cauchy problem for mathematical physics equations with random initial conditions, simulation of random processes, statistics of random processes, wavelet expansions of random processes Publications: 220</p>
	<p>Anatolii Pashko, born on February 28, 1962, Kuznetsovsk, Ukraine</p> <p>Current position, grades: assistant professor University studies: Taras Shevchenko National University of Kyiv Scientific interest: Stochastic processes theory, simulation of stochastic processes and fields Publications: 100</p>

Energy dissipation characteristics of sandstone cutting under mechanical impact load

Wang Shuren^{1*}, Su Junqing¹, Hagan Paul²

¹*School of Civil Engineering and Mechanics, Yanshan University, Qinhuangdao, 066004, China*

²*School of Mining Engineering, University of New South Wales, Sydney, 2052, Australia*

Received 28 march 2014, www.tsi.lv

Abstract

Based on the sandstone experiment by using the linear rock cutting machine at the school of mining engineering, University of New South Wales (UNSW), the computational model was built by using particle flow code (PFC), and the energy dissipation characteristics was studied considering these parameters such as cutting speed, cutting depth, rock strength, rock temperature and rock confining pressure. The results showed that the specific energy and acoustic emission of sandstone were proportional to the cutting speed and rock strength, there was an optimal depth of rock cutting with the cutting depth increasing under the confining pressure condition, and the specific energy and acoustic emission of sandstone changed obviously due to the influence of the rock temperature and rock confining pressure.

Keywords: Rock Cutting, Energy Dissipation, Specific Energy, Acoustic Emission, Particle Flow Code

1 Introduction

Rock-cutting usually caused the process of initiation crack, development, expansion, aggregation and transfixion of micro fissures in the rock under dynamic loading [1]. For the difficulty degree and energy dissipation of rock-cutting, they are not only the important indicators to evaluate the drilling, blasting, mining, beneficiation process, but also the key parameters to optimize the rock fragment and improve the production capacity.

At present, many domestic scholars have done lots of work in evaluating the difficulty of rock -cutting and energy dissipation, and achieved fruitful research results [2-3]. For example, X.B. Li et al. had conducted the impact crushing tests on four kinds of typical rocks under different loading waveforms, and found that the amount of energy absorbed by the rock was the most under the bell-wave loading [4]. M.Q. You et al. conducted the conventional triaxial loading test, and proved that the amount of energy absorbed by the broken rock was increased in a linear manner with increasing confining pressure [5]. C.J. Xia et al. had carried out the impact experiments of man-made rock with different porosities with Split Hopkinson Pressure Bar (SHPB), the characteristic of energy dissipation and the influence of rock porosity on energy dissipation were investigated during the impact process, and the relationship between the energy dissipation of critical failure and rock porosity was also analysed [6]. Q. Su et al. researched the mechanism of sand energy dissipation under the cyclic loading using particle flow code (PFC), and proved that

the smaller the sand porosity, the more system energy dissipation, and the energy dissipation improved with the confined pressure increasing [7]. For the research of specific energy of rock breaking, the main representative theories were put forward by foreign scholars such as the new-surface theory (P.R. Rittingerde), the similarity theory (G. Kick) and the crack theory (F.C. Bond) [8]. R.M. Goktan et al. researched the relationship between the rock brittleness index and rock-cutting efficiency, and found that the rock-cutting efficiency improved with the rock brittleness increasing under other conditions remained unchanging [9]. H. Copur et al. developed a set of new empirical indices and a new interpretation method of macro-scale indentation tests, which were useful for predicting the cutting efficiency and mechanical properties of rocks [10]. B. Adebayo made an assessment of the cuttability of granite and limestone, and found that the cutting rate increased in a linear manner with the growth of rebound hardness value, uniaxial compressive strength and silica content [11]. K. Dey et al. had proposed a new index, which could reasonably predict the production efficiency of the surface coal mining [12]. D. Tumac et al. tried to evaluate the cuttability of rocks by using the rock hardness and the uniaxial compressive strength index [13], and the related studies [14-15], etc.

In summary, though many research achievements have been obtained for evaluating the difficulty of rock breaking and energy dissipation, there is still a lot of work to be researched in depth due to the non-uniformity of rock material, the difference of geological conditions and the complexity of engineering conditions. Therefore, the numerical calculation model of rock-cutting was built

* *Corresponding author* e-mail: w_sr88@163.com

based on the test by using the linear rock-cutting equipment at the University of New South Wales (UNSW) in Australia. The numerical simulation of rock-cutting energy dissipation was conducted under different conditions, which was of important theoretical significance and practical value in guiding the similar engineering to reduce the cost of drilling and blasting and improve the production efficiency.

2 Rock-cutting experiment under the impact load

2.1 TESTING EQUIPMENT AND SAMPLES

The cutting test was conducted by using the linear rock cutting machine in the laboratory of School of Mining Engineering (UNSW). As shown in Fig. 1, the dimensions of the cutting machine were 1829 mm long, 1118 mm wide and 1321 mm high, respectively.

All the sandstone samples were obtained from the field adjacent to hynds seam in Hunter Valley, New South Wales. For each cutting test, the dynamic monitor fixed on the cutter monitored the voltage change data along the horizontal cutting direction and vertical cutting direction (Fig. 1), the mechanical impact curves of cutting force and normal force could be shown on the computer screen in real time through the conversion software program. For each sandstone sample, the cutting test should be operated for four times orderly, as showed in Fig. 1(c).

2.2 TEST METHODS AND PROCEDURES

The process of the test was as follows:

Step 1: Each cylindrical core sample with diameter $\varnothing 76$ mm should be fixed in the planer with flat clamp horizontally.

Step 2: Installing the new standard cutting tool, which was 12.5 mm wide chisel with front rake angle 0° , back clearance angle 5° , and cemented tungsten carbide.

Step 3: Adjusting the cutter depth of 5 mm through raising or lowering the lathe height. A flat tray was prepared for collecting cutting debris, and the testing machine must be covered with transparent plastic sheeting to prevent debris scattered everywhere.

Step 4: Starting the testing machine to cut the sandstone sample at a constant speed of 150 mm/s. During the test, the sandstone sample should be rotated counter clockwise by 180° after the first cutting being finished, then the second cutting can be operated along the same direction as the first cutting. After the second

cutting it should be rotated counter clockwise by 90° according to the position of the first cutting and the third cutting should be conducted. Finally, the sandstone sample should be rotated counter clockwise by 180° according to the third cutting and the forth cutting should be completed [16].

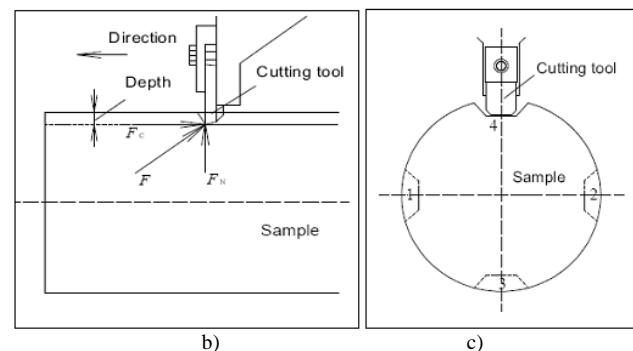


FIGURE 1 Linear rock-cutting machine and cutting sequence diagram: a) Testing machine and test system, b) Side view of sample, c) Front view of sample.

2.3 EXPERIMENT RESULTS AND ANALYSIS

2.3.1 Fragment characteristics of cutting debris

Fig. 2 showed the four sandstone samples and the collected rock-cutting debris, respectively. For each rock sample and for each cutting, the cutting debris must be collected, bagged, numbered and weighted. From the test results we found that samples No. 377542 and No. 377590 were short, their recovery ratios were low, their cleavage and cranny were developed, so the integrity was poor. On the contrary, samples No. 377554 and No. 370526 were long, their recovery ratios were high, their cleavage and cranny were not developed, and the integrity was good.



FIGURE 2 Sandstone samples and rock-cutting debris: a) No. 377542, b) No. 377554, c) No. 370526, d) No. 377590

2.3.2 Specific energy analysis of cutting debris

In the experiment, the specific energy (*SE*) is the energy or work required to cut unit volume of the rock, then *SE* is calculated as follows:

$$SE = \frac{F_c}{V} \tag{1}$$

Where F_c is the mean cutting force (kN), and V is the volume of excavated material per unit length of cutting (m^3/km). The unit of *SE* is usually MJ/m^3 .

$$V = LS \tag{2}$$

Where L is the distance the pick travels (m), and S is the unit cross-sectional area along the pick travels (m^2).

In a given rock the specific energy is used as a measure of the efficiency of a cutting system with lower values indicating higher efficiency, and what's more, in the context of cuttability assessment, it can be used to both compare the cuttability of different rocks and indicate approximately the potential excavation rate for a particular machine type in a given rock [17]. Table 1 showed the mean cutting force, the mean normal force and the specific energy of rock-cutting on the basis of the experimental data statistics by using the linear rock-cutting equipment.

As seen from Figs. 3 and 4, with the curves of the mean values changed from high to low, it indicated that the strength of the four types of sandstones changed from high to low. That is, the specific energy of rock-cutting increased with the growth of rock strength, and vice versa.

TABLE 1 Parameter statistics of the rock-cutting experiment

Sample	Cutting sequence	Cutting force (kN)	Normal force (kN)	Distance (mm)	Specific quantity (m^3/km)	Specific energy (MJ/m^3)
a) 377542	RTC14	1.56	1.62	180	0.0534	29.28
	RTC15	1.56	1.54	326	0.0738	21.17
	RTC16	1.78	1.87	486	0.0793	22.47
	RTC17	2.51	2.68	611	0.0615	40.78
	Mean	1.85	1.93	401	0.067	28.43
b) 377554	RTC30	1.77	1.62	180	0.0711	24.96
	RTC31	1.82	1.63	198	0.0666	27.29
	RTC32	1.80	1.63	208	0.0671	26.79
	RTC33	1.28	1.25	190	0.0665	26.79
	Mean	1.67	1.53	194	0.068	26.46
c) 370526	RTC34	1.31	1.09	218	0.0640	20.47
	RTC35	1.21	1.01	437	0.0566	21.32
	RTC36	1.37	1.06	647	0.0720	19.00
	RTC37	1.35	1.01	863	0.0682	19.77
	Mean	1.31	1.04	541	0.065	20.14
d) 377590	RTC53	0.75	0.58	162	0.0940	7.95
	RTC54	1.15	0.78	322	0.1489	7.75
	RTC55	0.79	0.62	458	0.0517	15.32
	RTC56	0.81	0.63	593	0.0897	8.98
	Mean	0.88	0.65	384	0.096	10.00

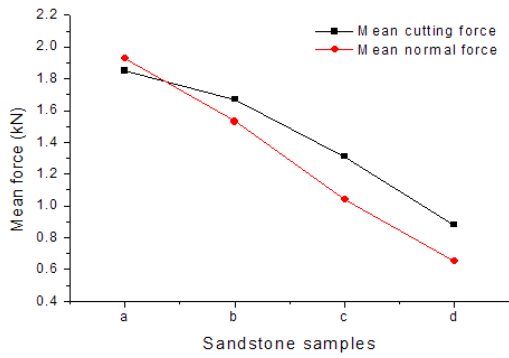


FIGURE 3 Curves of mean force changing

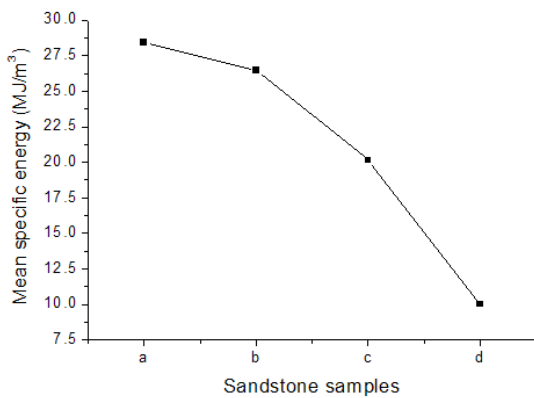


FIGURE 4 Curves of mean specific energy changing

3 Rock-cutting numerical test under the impact load

3.1 BUILDING COMPUTATIONAL MODEL

As shown in Figs. 5 and 6, the numerical calculation model was built by using PFC according to the engineering mechanical model. The horizontal velocity of the calculation model was zero, the top surface was free or the confining pressure boundary. The cutter simulated by using two mutually perpendicular wall sections was moved to the right at a certain velocity, the back-rake angle and the cutting depth to simulate the cutting process.

The specific process of modelling was as follows:

Step 1: The specimen was rectangular (width × height = 100 mm × 50 mm) and confined by three frictionless walls on the bottom, the left and the right sides, which was generated by the radius expansion method.

Step 2: The radii of all particles were changed uniformly to achieve a specified isotropic stress so as to reduce the magnitude of locked-in stresses that would develop after the subsequent bond-installation. In this paper the isotropic stress was set to 0.1 MPa.

Step 3: The floating particles that have less than three contacts were eliminated. Then the parallel bonds were installed throughout the assembly among all particles. The cutter was represented by two wall segments (both of length $l = 80$ mm) that were mutually perpendicular with

a back-rake angle of 20° . The cutter was moved horizontally across the rock at a velocity of 0.5 m/s and at a cutting depth of 5 mm.

Step 4: In order to simulate the compressive stress on the rock surface, a pressure-application algorithm is developed for PFC by continually identifying a connected chain of particles on the rock surface, and applying the pressure to those particles as the cutting process proceeds.

Step 5: In the process of rock-cutting simulation, the calculation model could trace energy data through the built-in energy function in PFC, and record the crack data through the built-in crack function so as to monitor the variation of rock temperature, specific energy and acoustic emission.

Step 6: The variation law of specific energy and acoustic emission would be analysed statistically with the change of the cutting speed, the cutting depth, the rock temperature and the confined pressure.

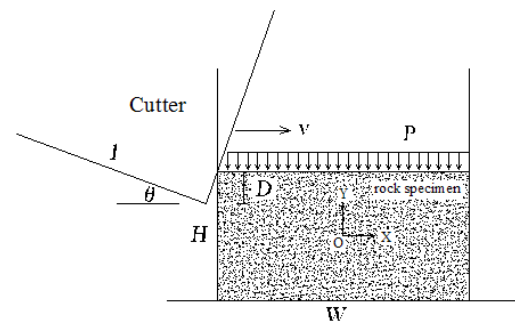


FIGURE 5 Engineering mechanic model of rock-cutting

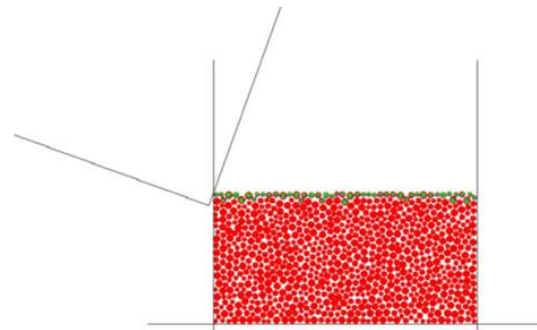


FIGURE 6 Computational model of rock-cutting

3.2 PARAMETERS CALIBRATION OF THE ROCK SPECIMEN

The rock specimen was treated as the porous material that consisted of particles and the cemented bodies. Before the numerical test, the micro mechanical parameters should be adjusted repeatedly until these parameters were consistent with the physical experiment results.

The micro mechanical required to be adjusted are as follows: ρ_{bulk} is bulk density, R_{max}/R_{min} is ball size ratio, R_ratio is ball size ratio, μ is ball friction coefficient, $\bar{\lambda}$ is parallel-bond radius multiplier, E_c is ball-ball contact modulus, \bar{E}_c is parallel-bond modulus, k_n/k_s is

ball stiffness ratio, \bar{k}_n/\bar{k}_s is parallel-bond stiffness ratio, $\bar{\sigma}_c$ is parallel-bond normal strength, $\bar{\tau}_c$ is parallel-bond

shear strength. The micro-parameters of the model were listed in the Table 2.

TABLE 2 Micro mechanical parameters of the model in PFC

ρ_{bulk} (kg/m ³)	R_{max}/R_{min}	R_ratio	μ	$\bar{\lambda}$	E_c (Pa)	\bar{E}_c (Pa)	k_n/k_s	\bar{k}_n/\bar{k}_s	$\bar{\sigma}_c$ (Pa)	$\bar{\tau}_c$ (Pa)
2620	1.66	1.66	0.5	1.0	83e9	83e9	3.8	3.8	91e6	91e6

3.3 ROCK-CUTTING ENERGY CHARACTERISTICS

3.3.1 Specific energy variation with the cutting velocity

As shown in Fig. 7, the specific energy of rock-cutting was proportional to the cutting velocity. The specific energy under the confined pressure was greater than that without the confined pressure at the same cutting velocity. What was more, as a result of the effect of the confined pressure, the curve slope was greater, that is, the specific energy of rock-cutting increased faster with the cutting velocity increasing under the confined pressure condition.

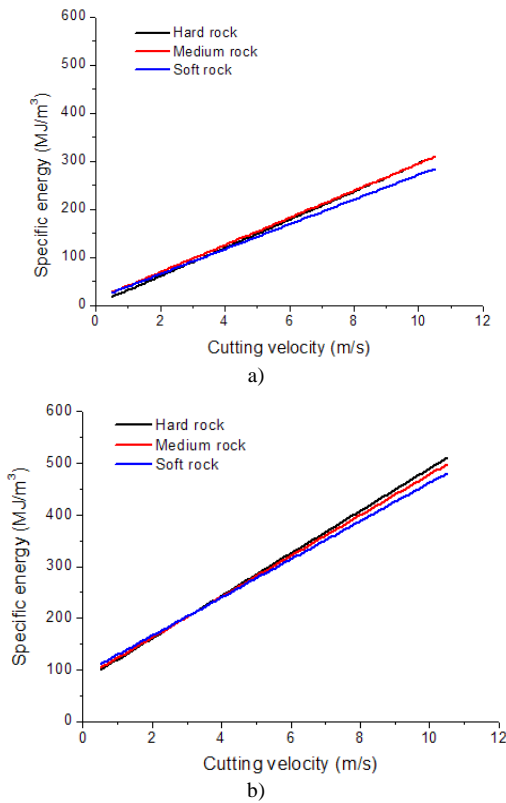


FIGURE 7 Specific energy-cutting velocity variation curves: a) without the confined pressure, b) with the confined pressure P = 10 MPa

Fig. 7(a) showed that the specific energy of the hard rock increased the fastest, followed by the medium hard rock, the last was the soft rock with the cutting velocity increasing without the confined pressure. Fig. 7(b) showed that the slope of curve, the value of specific energy and the amount of energy consumption increased obviously under the confined pressure being set to 10 MPa.

3.3.2 Specific energy variation with the cutting depth

As shown in Fig. 8(a), with the cutting depth increasing, the specific energy curves of three kinds of sandstones presented the slow-growth trend without the confined pressure, while the medium hard rock began to decline when the cutting depth was at 20 mm. As shown in Fig. 8(b), under the confined pressure, with the cutting depth increasing, the specific energy curves decreased at first and then increased. There was a minimum value of specific energy, which was called the optimal rock-cutting depth. The optimal rock-cutting depth of hard rock was about 18 mm, that of the medium hard rock was about 24 mm, and that of the soft rock was about 21 mm.

Compared Fig. 8(a) and Fig. 8(b), we could find that the cutting depth should not be too deep for hard rock and soft rock, while the cutting depth of the medium hard rock could increase appropriately.

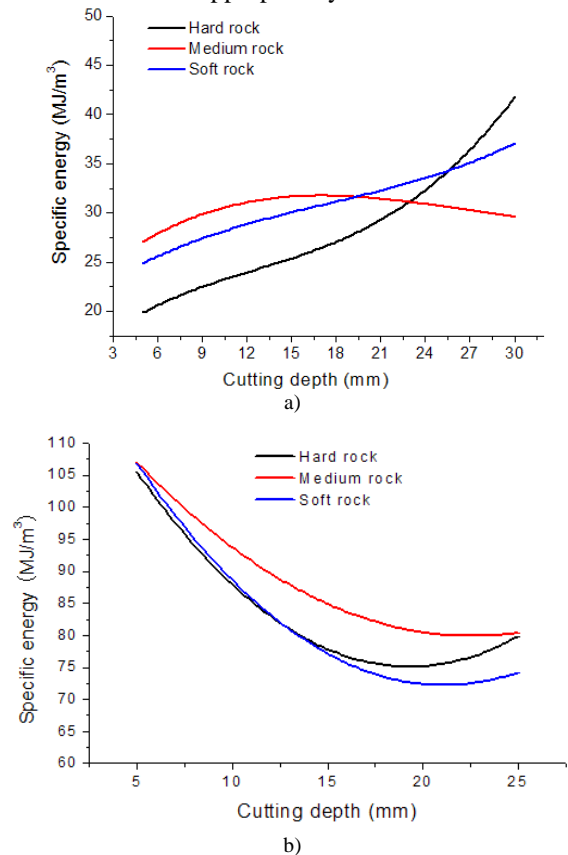


FIGURE 8 Specific energy-cutting depth variation curves: a) without the confined pressure, b) with the confined pressure P = 10 MPa

3.3.3 Specific energy variation with the rock temperature

As shown in Fig. 9, with the rock temperature increasing, the specific energy curves of three kinds of sandstones demonstrated S-shaped nonlinear change at the low value without the confined pressure, and the soft rock changed obviously than the others. After applying the confined pressure, the specific energy curves of hard rock and soft rock demonstrated S-shaped nonlinear change at the high value, while the medium hard rock changed greatly than the others.

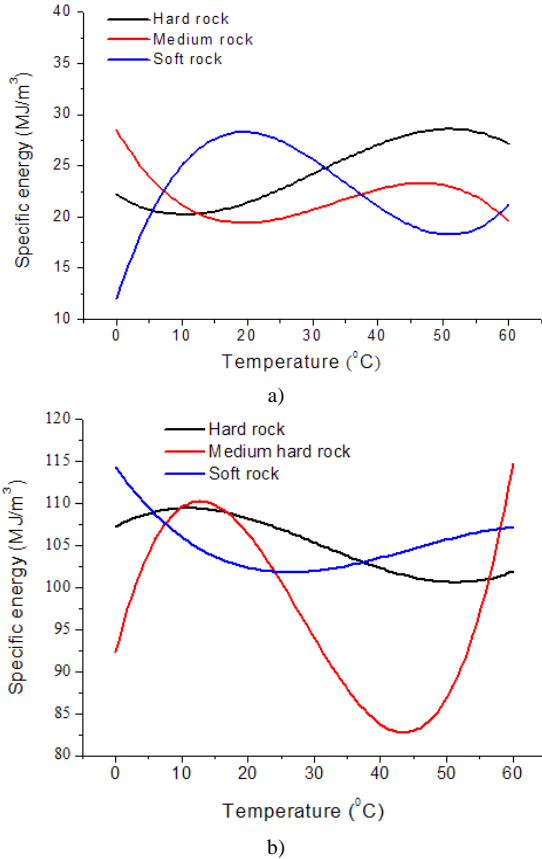


FIGURE 9 Specific energy-temperature variation curves: a) without the confined pressure, b) with the confined pressure P = 10 MPa

Compared Fig. 9(a) and Fig. 9(b), we find that the specific energy under the confined pressure at 10 MPa was about five times larger than that without the confined pressure.

3.4 ROCK-CUTTING AE CHARACTERISTICS

3.4.1 Acoustic emission variation with cutting velocity

The acoustic emission (AE) is commonly defined as transient elastic waves within a material caused by the release of localized stress energy. The rock AE is directly related to the generation of internal micro cracks (damage). A micro crack occurs along with a release of strain energy, that is, there is an AE. Therefore, the rock AE can be simulated by recording number of micro cracks. So the crack function defines a link damage

between particles as a crack generation in PFC. As a result, the rock AE can be simulated by the crack function in PFC.

As shown in Fig. 10, the AE was directly proportional to the cutting velocity. Without the confined pressure, there was an obvious difference among the growth rate of the AE three kinds of sandstones, while under the confined pressure, the growth rate of the AE were almost the same.

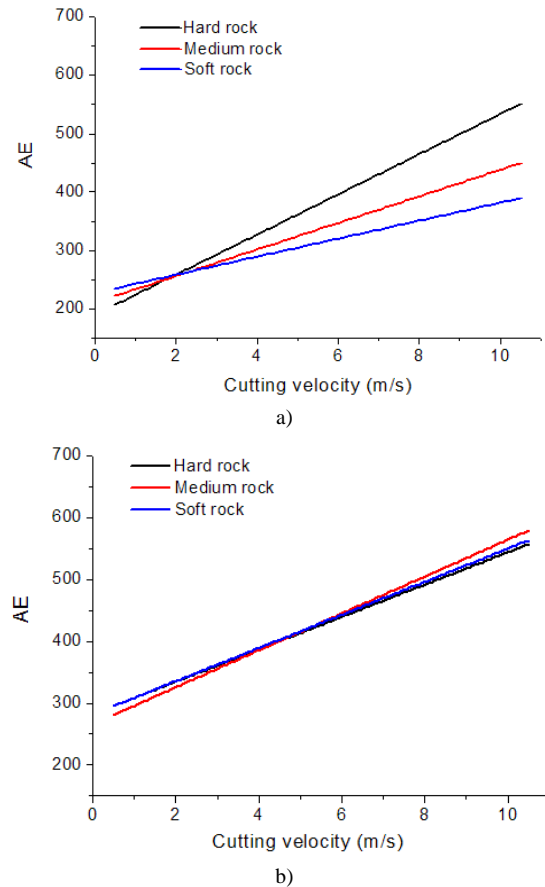


FIGURE 10 Specific energy-AE variation curves: a) without the confined pressure, b) with the confined pressure P = 10 MPa.

As shown in Fig. 10(a), with the cutting velocity increasing, the AE growth rate of the hard rock increased the fastest, followed by the medium hard rock and the last was the soft rock. When the cutting velocity was less than 2.0 m/s, the growth rate of AE was soft rock>medium hard rock>hard rock; When the cutting velocity was more than 2.0 m/s, the growth rate of AE was hard rock>medium hard rock>soft rock. As shown in Fig. 11(b), after being set the confined pressure at 10 MPa, the growth rate of AE of three kinds of sandstones were almost the same.

3.4.2 Acoustic emission variation with the cutting depth

As shown in Fig. 11, with the cutting depth increasing, the confined pressure only had a slight influence on the AE growth rate of the medium hard rock, and it had no effect on that of the hard rock and the soft rock. In a

word, no matter whether the confined pressure existed or not, the hard rock had the same AE growth rate as that of the soft rock.

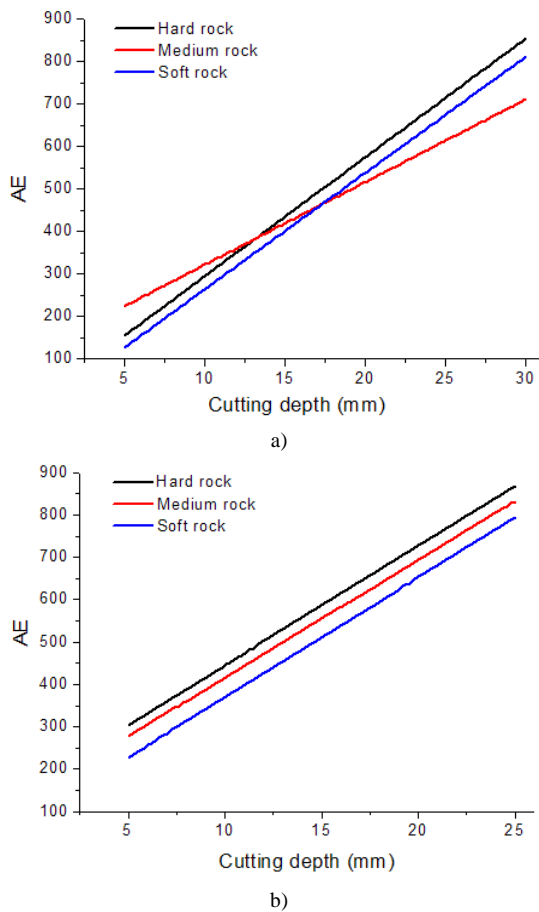


FIGURE 11 Cutting depth-AE variation curves: a) without the confined pressure, b) with the confined pressure P = 10 MPa

As shown in Fig. 11(a), without the confined pressure, the number of AE was proportional to the cutting depth. The AE growth rate of the hard rock was equal to the soft rock and greater than the medium hard rock. The curves in Fig. 11(b) showed that three kinds of sandstones AE were proportional to the cutting depth, and had the same slope, namely the growth rates were equal under the confined pressure. The AE of three kinds of sandstones had the following relationship: the hard rock was the greatest, followed by the medium hard rock and the last was the soft rock.

3.4.3 Acoustic emission variation with the rock temperature

As shown in Fig. 12(a), with the rock temperature increasing, the AE of the hard rock and the medium hard rock had little changes, while the soft rock increased at first and then decreased without the confined pressure within the temperature range of 0 °C to 60 °C, and reached the maximum value at 30 °C.

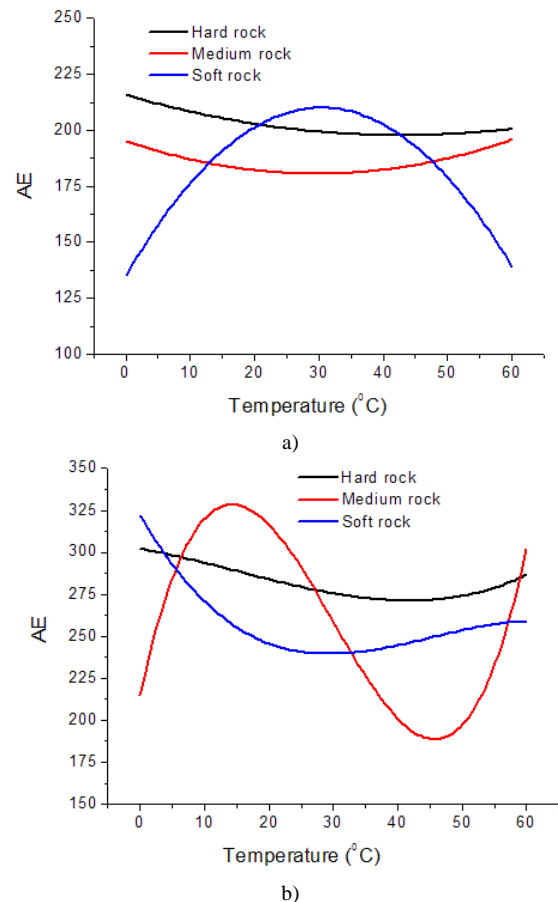


FIGURE 12 Rock temperature-AE variation curves: a) without the confined pressure, b) with the confined pressure P = 10 MPa

The curves in Fig. 12(b) showed that the AE of the hard rock and the soft rock presented a slowly downward trend and remained the same above 30 °C. And the medium hard rock was greatly influenced by the temperature factors, of which the curve presented an inverted S-shaped and obtained the maximum value at about 15 °C and reached the minimum value at about 45 °C.

4 Conclusion

The specific energy of rock-cutting was proportional to the cutting velocity. With the cutting depth increasing, the specific energy curves of three kinds of sandstones presented the slow-growth trend without the confined pressure, while the medium hard rock began to decline when the cutting depth at a certain value. Under the confined pressure, with the cutting depth increasing, the specific energy curves decreased at first and then increased and there was an optimal depth of rock cutting.

With the rock temperature increasing, the specific energy curves of three kinds of sandstones demonstrated S-shaped nonlinear change at the low value without the confined pressure, and the soft rock changed obviously than the others. After applying the confined pressure, the specific energy curves of hard rock and soft rock demonstrated S-shaped nonlinear change at the high

value, while the medium hard rock changed greatly than the others.




AE was directly proportional to the cutting velocity and the cutting depth. With the rock temperature increasing, the AE curves of three kinds of sandstones demonstrated nonlinear change.

References

- [1] Lin D N, Chen S R 2005 Experimental study on damage evolution law of rock under cyclical impact loadings *Chinese Journal of Rock Mechanics and Engineering* **24**(22) 4094-4098 (in Chinese)
- [2] Hu L Q, Li X B, Zhao F J 2002 Study on energy consumption in fracture and damage of rock induced by impact loadings *Chinese Journal of Rock Mechanics and Engineering* **S2** 2304-2308 (in Chinese)
- [3] Guo L J, Yang Y H, Hua Y H 2013 Test and analysis on distortion and damage of rock under impact loading *Journal of Water Resources and Architectural Engineering* **11**(6) 31-34 (in Chinese)
- [4] Li X B, Lai H H, Gu D S 1992 Energy consumption of ore rocks under different loading waveforms *The Chinese Journal of Nonferrous Metals* **2**(4) 10-14 (in Chinese)
- [5] You Q M, Hua A Z 2002 Energy analysis on failure process of rock specimens *Chinese Journal of Rock Mechanics and Engineering* **21**(6) 778-781 (in Chinese)
- [6] Xia C J, Xie H P, Ju Y, Zhou H W 2006 Experimental study of energy dissipation of porous rock under impact loading *Engineering Mechanics* **23**(9) 1-5 (in Chinese)
- [7] L. Kong, Q. Su, Y.C. Wang, R. Peng 2012 Mesoscopic mechanics and energy dissipation analysis of sand deformation under slow cyclic loading. *Chinese Journal of Underground Space and Engineering* **8**(2), 268-273 (in Chinese)
- [8] Guo C Y, Xie X F, Wu X H, Yao W J, Zhang C Y 2008 The relationship among rock crushing energy, the Protodyakonov coefficient and rock strength *Journal of Chongqing Jianzhu University* **30**(6) 28-31 (in Chinese)
- [9] Goktan R M, Yilmaz N G 2005 A new methodology for the analysis of the relationship between rock brittleness index and drag pick cutting efficiency *Journal-south African Institute of Mining and Metallurgy* **105**(10) 727-733
- [10] Copur H, Bilgin N, Tuncdemir H, Balci C 2003 A set of indices based on indentation tests for assessment of rock cutting performance and rock properties *Journal-south African Institute of Mining and Metallurgy* **103** (9) 589-599
- [11] Adebayo B 2008 Evaluation of cuttability of selected rocks in south-western Nigeria *AU Journal of Technology* **12** (2) 126-129
- [12] Dey K, Ghose A K 2011 Review of Cuttability Indices and A new rockmass classification approach for selection of surface miners *Rock mechanics and rock engineering* **44**(5) 601-611
- [13] Tumac D, Bilgin N, Feridunoglu C, Ergin H 2007 Estimation of rock cuttability from shore hardness and compressive strength properties *Rock Mech. Rock Eng* **40**(5) 477-490
- [14] Krolczyk G, Legutko S, Stoić A 2013 Influence of cutting parameters and conditions onto surface hardness of Duplex Stainless Steel after turning process *Tehnički vjesnik – Technical Gazette* **20**(6) 1077-1080
- [15] Guettaf A, Chabane F, Arif A, Benramache S 2013 Dynamic Modeling in a Switched Reluctance Motor SRM using Finite Elements *Journal of Power Technologies* **93**(3) 149-153
- [16] Wang S R, Hagan P, Cheng Y 2012 Fractal Characteristics of Sandstone Cutting Fracture Under Mechanical Shock Loading Conditions *Applied Mechanics and Materials* **226-228** 1789-1794
- [17] Roxborough F F 1987 *The role of some basic rock properties in assessing cuttability. Seminar on Tunnels-Wholly Engineered Structures AFCC/IEAust Canberra, April, pp.1-21*

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (51074140; 51310105020) and the Natural Science Foundation of Hebei Province of China (E2014203012), all these are gratefully acknowledged.

Authors	
	<p>Shuren Wang, born in December, 1968, Luannan County, Hebei Province, P.R. China</p> <p>Current position, grades: Professor of School of Civil Engineering and Mechanics, Yanshan University, China. University studies: received his Ph.D degree of Engineering Mechanics from University of Science and Technology Beijing, China. Scientific interest: His research interest fields include mining engineering, geotechnical engineering, rock mechanics and numerical simulation analysis, etc. Publications: more than 65 papers published in various journals. Experience: He has working experience of 22 years, has completed 30 scientific research projects.</p>
	<p>Junqing Su, born in August, 1987, Mengcun County, Hebei Province, P.R. China</p> <p>Current position, grades: Now he is a master student at School of Civil Engineering and Mechanics, Yanshan University, China. University studies: He received B.S. degree from School of Civil Engineering and Mechanics, Yanshan University, China. Scientific interest: His interested research is mining engineering, rock mechanics, and numerical simulation analysis, etc. Publications: 3 papers published in various journals. Experience: He has completed 2 scientific research projects.</p>
	<p>Paul Hagan, born in February, 1958, Newcastle City, New South Wales, Australia</p> <p>Current position, grades: Associate Professor of School of Mining Engineering, University of New South Wales, Australia University studies: received his Ph.D degree of Mining Engineering from School of Mining Engineering, University of New South Wales, Australia Scientific interest: His research interest fields include mining engineering, rock mechanics, etc. Publications: more than 60 papers published in various journals and academic conferences. Experience: He has over thirty years' experience including twelve years in the mining industry.</p>

Gas explosion characteristics and its control technologies in closed fire zone

Zhenzhen Jia^{1,2}, Feng Tao^{1,2*}

¹ School of Resources and Safety Engineering Central South University, Changsha City, Hunan Province, China, 410083

² School of Energy and Safety Engineering Hunan University of Science and Technology, Xiangtan City, Hunan Province, China, 411201

Received 6 October 2013, www.tsi.lv

Abstract

The closure measures of fire zones are taken after gas explosion in the working face, which can bring two problems: whether closure measures will lead to a secondary gas explosion in closed fire zones or not, and what will be the rough time interval between taking measures and gas explosion occurrence. To solve these problems, gas accumulation characteristics, oxygen concentration characteristics and fire sources in the fire closure process were analysed, and then the characteristics and rules of gas explosion were acquired, in addition, the pressure change and gas accumulation model in closure zones under three conditions (the temporary closure wall only in air inlet laneway, only in air return laneway, or both in air inlet laneway and air return laneway) were obtained. Finally, the measures and technologies to prevent and control gas explosion were introduced in the fire closure process of working face.

Keywords: Closed Fire Zone, Gas Explosion Characteristics, Gas Accumulation Model, Fire Sources, Control Technologies

1 Introduction

With the increase of coal mining scale and the extension of production level, gas emission quantity increases gradually and mine fires occur more and more frequently. In high gas area, once fire happens, regardless of an internal fire, an external fire or a secondary fire caused by gas explosion, the improper handling measures can result in gas explosion, which can bring the injuries and deaths of disaster relief personnel [1, 2]. When high-temperature and high-pressure smoke produced by gas explosion flows through the over-limit area of gas concentration, two scenarios may appear: one is that a secondary gas explosion or a fire is resulted in, thus directly threatening the workers' lives; the other is that a secondary gas explosion or a fire is not brought, but a large number of harmful gases (such as high concentration CO₂, high temperature water vapour, CO, H₂S) generated by gas explosion seriously threaten the workers' lives. The CO₂ with the concentration above 5 % or the CO with the concentration above 0.5% can result in workers to death; in addition, the high temperature water vapour can scald the internal organs.

Whether gas explosion can cause secondary disasters or not, the diffusion of harmful smoke is the main factor on threatening the lives of underground workers. A large number of gas explosion accidents show that the poisoning and suffocation caused by the harmful smoke are the main reasons for the casualties. In addition, the smoke produced by gas explosion can reduce the visibility and block the sight, thus affecting the safety evacuation of personnel and the success of disaster relief

work. Generally speaking, the decision-making and operations for disaster relief always follow gas explosion accidents. Therefore, for high gas coal mines with fire zones, the measures, such as closing fire zones and injecting inert gases, are often taken in order to prevent the disaster expansion, which can bring the changes of ventilation pressure and gas concentration in closed fire zones. Meanwhile, two problems can be brought [3]: the one is whether closure measures will lead to a secondary gas explosion in closed fire zones or not, the second is what will be the rough time interval between taking measures and gas explosion. Consequently, oxygen concentration characteristics, gas accumulation characteristics and fire characteristics in closed zones are analysed in this paper. The study results can reduce many uncertain factors in disaster relief process to a minimum range and increase science and correctness of disaster relief decision.

2 Gas explosion characteristics in closure process of fire zone

Gas explosion occurrence must be provided with the following three conditions at the same time [4]: the gas concentration is within the explosion limits (i.e. 5-16%), the oxygen concentration of the mixed gas is not less than 12%, and the ignition source has sufficient energy.

In the closure process of fire zone, if the oxygen concentration is more than 10%, the heat released by the oxidization reaction of coal and oxygen can maintain the constant coal temperature. Thus, the coal temperature will not decrease significantly (namely, the coal

* Corresponding author e-mail: jiazhenzhen1982@126.com

temperature can be assumed to be constant). As long as the coal temperature does not drop below the minimum temperature of gas explosion, there are fire sources in combustion zones. In closure process of fire zones in the working face, the fire sources are the main factors on causing gas explosion. According to the practical experience in coal mines, after the fire zone is closed, the methods, such as gel injection to fire zones, can usually be taken to make the temperature drop below the lower limit of gas explosion temperature. If the cooling measures are not taken, the fire sources in fire zones exist for a long time. The main fire sources resulting in gas explosion in the gob are the spontaneous combustion of residual coal and rock friction (collision) sparks during roof caving. Therefore, the changes of oxygen concentration and gas concentration in the fire zone were mainly analysed.

In the closed zones, the gob and other fractures, because of oxygen consumption or lack of oxygen supply conditions, the oxygen concentration may be less than 12%. In other laneways and workplaces, the oxygen concentration is not less than 12%. In the fire closure process, because the oxygen is consumed by the reaction between coal and oxygen, the oxygen concentration decreases gradually in the closed fire zone [5]. When the oxygen concentration drops below 12%, the gas loses its explosive capability. In the fire closure process, if the gas concentration increases to 5-16%, the oxygen concentration is still more than 12%, and the temperature is still high, then gas explosion may happen [6]. In order to make calculation easy, the reaction between coal and oxidation is assumed to be the simple gas-solid reaction after fire closure. According to the principle of chemical reaction kinetics, the reaction rate is proportional to the oxygen concentration. Consequently, as long as several groups of oxygen concentration data are measured in different time, the curve equation of oxygen concentration varying with time in the actual conditions can be obtained by using relational expression of oxygen concentration varying with time, and then the oxygen concentration at any given moment can be obtained after fire closure.

In the fire closure process, gas emission is from working face (coal laneway wall) and fallen coal (residual coal) [7], the gas concentration increases rapidly and the local gas accumulation is formed around the gas emission place [8]. According to the gas sources and accumulation time, the gas accumulation during fire closure has four types: the gas accumulation due to normal emission from the gob and blind laneway, the gas accumulation due to geological condition change, such as the connection between fault and gob, the gas accumulation due to the roof caving in the gob, and the gas accumulation due to gas outburst. In fact, in the fire closure process, gas explosion in the gob is mainly affected by four factors [9]: the properties of the exothermic oxidation of coal, the thickness and fragment of residual coal in the gob, the air leakage in the gob

after fire zone closure, and the original temperature of the surrounding rock in the gob. Due to the comprehensive effect of the above four factors, the spontaneous combustion process of residual coal in the gob shows dynamic change. Therefore, the occurrence of gas explosion in the closed fire zone mainly depends on oxygen concentration and gas concentration.

In the fire closure process, gas accumulation caused by the undesirable gas drainage results in gas explosion accidents mainly in the working face, secondly in the laneway. Therefore, reasonable gas drainage can effectively control gas explosion [10]. Gas explosion in working face is mostly because much gas emits into the working face for roof caving. Thus, the measures on controlling roof caving should be taken to prevent sudden gas emission and avoid gas explosion. In addition, strengthening the management and implementing operation instruction strictly can effectively reduce gas accumulation to avoid gas explosion accidents.

3 Gas accumulation model in fire closure zone

In the fire closure process, first, the auxiliary or temporarily closure is often built fast in the air inlet laneway and the air return laneway of zones, which will be closed. After the closure, zones are stability for a period of time, and then other disaster relief measures are taken. In the closure process, the closure measures can increase the local wind resistance of the air inlet laneway and air return laneway of closure zones, thus changing the atmospheric pressure distribution in the fire closure zones, which makes changes of the gas emission and migration rule in the fire closure zone.

According to facts, the spontaneous combustion in working face is most frequently, at the same time, the possibility that the gas concentration reaches the explosion limits due to gas accumulation is largest, thus the working face is analysed. Gas migration of a typical working face is described in Figure 1, if spontaneous combustion appears in the gob, the working face must be closed, there are three methods for the fire closure (or constructing temporary closure places): the temporary closure walls are only constructed in air inlet laneway, the temporary closure walls are only constructed in air return laneway, and the temporary closure walls are constructed in air inlet laneway and return laneway.

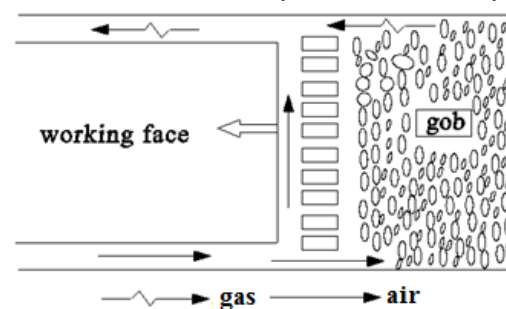


FIGURE 1 Gas migration of typical working face

3.1 GAS PRESSURE MODEL IN CLOSED ZONES

When the temporary closure walls are constructed in air inlet laneway and return laneway (or only in air inlet laneway, or only in air return laneway), at temporary closure places, the closure wall cuts off airflow in the closed zones, at the same time, it also changes and adjusts wind pressure distribution. The pressure changes at the temporary closure wall can result in the pressure change in the closed zone. The static air pressure increases in the air inlet side of temporary closure wall, while the static air pressure decreases in the air return side of temporary closure wall. The effect of the temporary closure wall on static air pressure in closed zone and laneway is determined by its gas tightness, which depends on the wall thickness, material properties and construction quality, etc.

Figure 2 describes the closure model of a working face with U type, M_1 and M_2 are the temporary closure walls in air inlet laneway and in air return laneway respectively, R_1 is the local wind resistance in M_1 , Q_1 is the air leakage in M_1 , P_{1n} and P_{1w} are the internal static air pressure and external static air pressure of M_1 respectively, R_2 is the local wind resistance in M_2 , Q_2 is the air leakage in M_2 , P_{2n} and P_{2w} are the internal static air pressure and external static air pressure of M_2 respectively, P_g is the static air pressure in the gob, Q_g is gas emission quantity from the gob to working face. As shown in Figure 2, when the temporary closure walls are constructed in the air inlet laneway and the air return laneway, with the increase of gas tightness of M_1 and M_2 , the air leakage reduces, so P_g and Q_g increase. When the temporary closure wall is constructed only in the air inlet laneway, with the increase of gas tightness of M_1 , the air leakage reduces, so Q_g increases, but P_g is constant. When the temporary closure wall is constructed only in the air return laneway, with the increase of gas tightness in M_2 , the air leakage reduces, so P_g and Q_g increase.

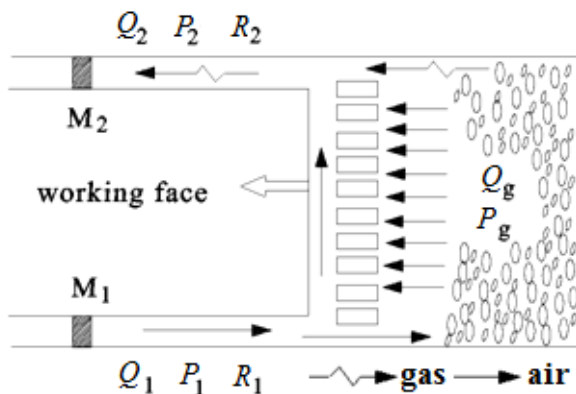


FIGURE 2 Closure model of U-type working face (The temporary closure walls constructed in air inlet laneway and return laneway)

3.2 CALCULATION MODEL OF GAS CONCENTRATION WITH TIME IN CLOSED ZONE

As shown in Figure 2, when the temporary closure walls are constructed in air inlet laneway and air return laneway, the assumptions are as follows. (1) When there is little or no wind in the working face, gas emission in air inlet laneway, air return laneway and working face are uniform [11]. (2) When the temporary closure walls are constructed in air inlet laneway and return laneway, the effect of ventilation pressure changes in working face on gas emission is negligible, namely, the gas emission is constant.

When the temporary closure walls are constructed in the air inlet laneway and air return laneway, the assumptions are as follows: the air leakage is Q (m^3/s), gas emission from coal laneway is b (m^3/s), the gas emission from working face is a (m^3/s), the laneway volume is V (m^3), the gas concentration at time t is x , the gas concentration at time $t+dt$ is $x+dx$, the initial gas concentration in the laneway is x_0 . Thus, when the time increment is dt , in the closed zone, the gas increment is $(a+2b)dt$ (m^3/s), the gas reduction due to air leakage is $Qxdt$ (m^3/s), the change of gas concentration is Vdx (m^3/s).

Taking the closed zone as the research object, according to the law of conservation of mass, the change of gas concentration in closed zone is equal to the difference between gas increment and gas reduction due to air leakage, namely:

$$Vdx = (a + 2b)dt - Qxdt \tag{1}$$

Then we can derive formula:

$$\frac{dx}{dt} + \frac{Q}{V}x = \frac{a + 2b}{V} \tag{2}$$

Solving for equation (2), we obtain:

$$x = Ce^{-\frac{Q}{V}t} + \frac{(a + 2b)}{Q} \tag{3}$$

Where C is a constant. Putting the boundary conditions ($t=0, x=x_0$) into equation (3), we get

$$C = x_0 - \frac{a + 2b}{Q}$$

Putting C into equation (3), we get

$$x = \left(x_0 - \frac{a + 2b}{Q} \right) e^{-\frac{Q}{V}t} + \frac{a + 2b}{Q} \tag{4}$$

The formula (4) is the calculation model of gas concentration varying with time when the temporary closure walls are constructed in air inlet and air return laneway. Similarly, when the temporary closure wall is constructed only in the air inlet laneway or in the air return laneway, the law of gas concentration varying with time can be obtained.

4 Control measures and technologies of fire closure zone with explosion risk

Based on the theoretical analysis, if the fire zone has explosion risk and must be closed, explosion-proof treatment in fire zone is necessary before fire closure. When the closure sequential is chosen according to the actual situation and the underground fire characteristics, the change of wind pressure and the effect of each laneway closure on the wind road of fire zone should be considered carefully. When many laneways need to be closed, the laneway without great impact on the system should be firstly closed.

1) When the fire zone with explosion risk must be closed, the problem on explosion proof should be first considered. Generally speaking, the closure should be constructed in the air inlet laneway firstly, the explosion proof methods (such as setting sandbags) should be used to prevent explosion, and the ventilation window should be reserved to ensure certain air into fire zone. After the explosion proof work is over, the ventilation window should be block off. If there is no problem within 24 hours, the closure in the air return laneway can be constructed.

2) When the fire zone is closing, the specific persons should be arranged to closely monitor the change of air compositions in fire zone. The gas sample should be obtained in the top, the middle and the bottom of the laneway. The best way is to carry out beam tube monitor by installing beam tubes once, which can avoid entering fire zones repeatedly.

3) The change of atmospheric pressure should be monitored during the fire closure process, the time that atmospheric pressure rises is the best time to build a firewall, which can help to prevent gas in the fire zone from flowing outward and eliminate the threat to disaster relief personnel.

4) The working face with high gas generally has the gas drainage system. After the working face is forced to close due to fire, to prevent gas accumulation and gas explosion, the gas drainage should continue until there is no explosion risk in the closed zone.

5) When closure conditions in fire zone are good and it can ensure no air leakage, the method that the temporary closure walls are constructed in the air inlet laneway and air return laneway can be considered, which can shorten time as possible and quickly cut off oxygen supply conditions. The injection of the gelatine, CO₂, N₂ can further enhance the safety of fire zone closure. The gelatine injection can reduce the temperature in fire zone below the low limit of gas explosion temperature [12].

6) In the fire closure process, the temperature and CO concentration in fire zone are very high, so workers cannot work near the fire zone. If conditions are permitted, the nitrogen injection from the ground into the

fire zone can be adopted to reduce oxygen concentration and coal temperature, thus ensuring the safety in working face.

5 Conclusions

1) When high-temperature and high-pressure smoke produced by gas explosion flows through the over-limit area of gas concentration, two scenarios may appear: one is that a secondary gas explosion or a fire is resulted in, thus directly threatening the workers' lives. The other is that a secondary gas explosion or a fire is not brought, but a large number of harmful gases (such as high concentration CO₂, high temperature water vapour, CO, H₂S) generated by gas explosion seriously threaten the workers' lives.

2) The gas explosion characteristics in closure process were obtained by analysing the time-space relationship of oxygen concentration, temperature, and gas concentration.

3) In the fire closure process, gas accumulation caused by the undesirable gas drainage results in gas explosion accidents mainly in the working face and laneway. Therefore, reasonable gas drainage can effectively control gas explosion. Gas explosion in working face is mostly because much gas emits into the working face for roof caving. Thus, the measures on controlling roof caving can be taken to prevent sudden gas emission and avoid gas explosion.

4) According to gas sources and accumulation time, strengthening the management and implementing operation instruction strictly can effectively reduce gas accumulation to avoid gas explosion.

5) When the temporary closure walls are constructed in the air inlet laneway and return laneway (or only in the air inlet laneway, or only in the air return laneway), at temporary closure places, the closure wall cuts off airflow in the closed zones, at the same time, it also changes and adjusts wind pressure distribution. The pressure changes at the temporary closure wall can result in the pressure changes in the closed zone. According to the law of conservation of mass, the law of gas concentration varying with time was obtained.

6) The measures and technologies such as laneway closure sequential, gas drainage in closed zones, measures of gel injection into the fire zone, and dilution of gas and oxygen by injecting inert gas (i.e. N₂) are introduced.

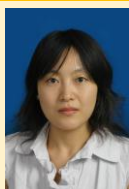
Acknowledgements

This work is supported by the National Natural Science Foundation of China (51004048, 51374003, 50834005, 50674047). Their supports are acknowledged with thanks.

References

- [1] Lin B Q, Ye Q 2012 *Mechanism and Control Technology for Gas Explosion in Coal Mines* China University of Science and Technology Press: Xuzhou (in Chinese)
- [2] Kobiera A, Kindracki J, Zydek P, Wolanski P A 2007 A New phenomenological model of gas explosion based on characteristics of flame surface *Journal of Loss Prevention in the Process Industries* 20(2) 271-280
- [3] Jiang B Y, Lin B Q, Shi S L, Zhu C J, Ning J 2012 Numerical simulation on the influences of initial temperature and initial pressure on attenuation characteristics and safety distance of gas explosion *Combustion Science and Technology* 184(2) 135-150
- [4] Lin B Q, Ye Q, Zhai C, Jian C G 2008 The propagation rule of methane explosion in bifurcation duct *Journal of China Coal Society* 33(2) 136-139 (in Chinese)
- [5] Domnina R, Venera B, Maria M, Codina M, Dumitru O 2009 Inerting effect of the combustion products on the confined deflagration of liquefied petroleum gas-air mixtures *Journal of Loss Prevention in the Process Industries* 22(4) 463-468
- [6] Wang C, Han W H, Ning J G, Yang Y Y 2012 High resolution numerical simulation of methane explosion in bend ducts *Safety science* 50(4) 709-717
- [7] Lunarzewski L W 1998 Gas emission prediction and recovery in underground coal mines *International Journal of Coal Geology* 35(1-4) 117-145
- [8] Ye Q, Lin B Q, Jiang W Z 2006 The study of methane outflow law in coal mining face *China Mining Magazine* 15(5) 38-41 (in Chinese)
- [9] Liu Z J, Li C Y, Zhang L L 2007 Analysis on characters of fire source of combustible gas exploration in work-out area *Coal Technology* 26(9) 67-69 (in Chinese)
- [10] Zhang X M 2012 *Gas Explosion Risk Analysis On Face of Closing-Fire-District in Coal Mine* XiAn University of Mining and Technology: Xi'an (in Chinese)
- [11] Valliappan S, Zhang W H 1996 Numerical modelling of methane gas migration in dry coal seams *International Journal for Numerical and Analytical Methods in Geomechanics* 20(8) 571-593
- [12] Zhang X H, Xu J C, Liu J Y, Wen H 1999 Analysis on the danger of gas explosion in coal mine after sealing the fire zone *Journal of Xi'an Mining Institute* 19(2) 110-113 (in Chinese)

Authors



Zhenzhen Jia, born in December, 1982, Xiangtan City, Hunan Province, P.R. China

Current position, grades: the Lecturer of School of Energy and Safety Engineering, Hunan University of Science and Technology, China.

University studies: received her B.Sc. from China University of Mining and Technology in China. She received her M.Sc. from China University of Mining and Technology in China.

Scientific interest: Her research interest fields include Gas Disasters Prevention in Coal Mines.

Publications: more than 10 papers published in various journals.

Experience: She has teaching experience of 7 years, has participated in three scientific research projects.



Feng Tao, born in April, 1957, Xiangtan City, Hunan Province, P.R. China

Current position, grades: the Professor of School of Energy and Safety Engineering, Hunan University of Science and Technology, China.

University studies: received his B.Sc. from Chongqing University in China. He received his M.Sc. from Chongqing University in China. He received his D.Sc. from Central South University in China.

Scientific interest: His research interest fields include Mining Engineering, Strata Control and Safety in Coal Mines.

Publications: more than 70 papers published in various journals.

Experience: He has teaching experience of 32 years, has completed twenty scientific research projects.

Numerical modelling of rock cross-cut coal uncovering based on ANSYS

G Wang^{1,2*}, H Y Wang^{1,2}, Q M Huang², C Q Su³

¹ State Key Laboratory of Ministry of Mining Disaster Prevention and Control, Shandong University of Science and Technology, Qingdao, Shandong, China, 266590

² College of Resources and Environmental Engineering, Shandong University of Science and Technology, Qingdao, Shandong, China, 266590

³ Lindong Mining Group Co. Ltd., Guiyang, Guizhou, China, 550056

Received 3 December 2013, www.tsi.lv

Abstract

Outbursts of coal and gas could be induced by rock cross-cut coal uncovering. ANSYS is used to numerically simulate the stress, strain and energy of surrounding rock of roadway during the process of rock cross-cut coal uncovering. Modelling results show that there is a banding tension stress zone in roof and floor of roadway after excavation. Principal and shear stress concentration are formed in the upper and under area of the anterior heading face, which is symmetrically distributed between medial axis of roadway, while stress, strain and strain energy of overlying strata above the coal seam approximately keep invariant. The occurrence of the stress concentration in upper area of the anterior heading face could contribute to the instability and failure of coal and rock mass. The area with weak destruction-resisting ability is the most easily to be the releasing port of outburst and should be regarded as the key region in outburst prevention.

Keywords: Rock Cross-cut Coal Uncovering, Numerical Modelling, ANSYS, Stress Distribution

1 Introduction

Outburst of coal and gas is a complicated dynamic disaster in underground coal mining and can occur when certain conditions of ground stress, coal physical and mechanical property and gas pressure are met [1]. In China, outbursts of coal and gas mainly occur in the development of roadway and advance of mining face. The outburst intensity of rock cross-cut coal uncovering is the largest, having average intensity of 586.1 t. Over 80% of large-scale outbursts are induced by rock cross-cut coal uncovering [2].

In the past few decades, some attempts have been made to numerically model the outburst of coal and gas, such as a gas desorption and flow model by Paterson [3], an airway gas flow model by Otuonye and Sheng and a fracture mechanics model by Odintsev [4-5]. More recently, Xu et al. modelled outbursts with a simple finite element model, which coupled gas flow and the deformation and failure of solid [6]. The model also incorporated small scale variability in modulus and strength of coal to model heterogeneity. A model was proposed by Wold et al. to model the coal failure by using plasticity theory with softening mechanism and fragmentation is modelled with continuum damage mechanics [7]. Xue et al. developed a coupled simulator which could model the deformation and failure of a coal seam, adsorption and desorption of gas and flow of gas

and water in coal to numerically simulate the initiation process of the outburst by linking and sequentially executing two numerical codes: FLAC^{3D} and COMET3 [8].

In terms of numerical modelling of rock cross-cut coal uncovering, rock failure process analysis code RFPA^{2D} was used to simulate the outburst in cross-cutting in steep coal seam containing gas [9]. Solid-gas coupling model RFPA^{2D}-Flow to numerically perform the instantaneous outburst induced by cross-cut driving and the whole process of micro-cracking, propagation, coalescence and ejection of coal and rock [10]. He used RFPA^{2D}-Flow to simulate the process of rock cross-cut coal uncovering of different gas pressure, coal strength and angle of coal seam [11]. The 6.20 outburst mechanical process of Hongling coal mine was simulated by using finite element analysis software of ANSYS and Flac^{3D} [12].

Rock cross-cut coal uncovering is dangerous and has high-degree difficulty in technology [13]. The changes of physical and mechanical properties in coal and rock mass caused by rock cross-cut coal uncovering are the main factors inducing outbursts. To perform numerical modelling of the stress, strain and energy of roadway surrounding rock during the process of rock cross-cut coal uncovering could be instructive for mechanism study and outburst prevention.

* Corresponding author e-mail: 15088427@qq.com

2 ANSYS modelling

ANSYS is a finite-element-method numerical software with the function of structure, fluid, electric field, magnetic field and electric field analysis. It is a commercial application code with powerful function and convenient operation and has the characteristics of data unification, powerful modelling, solving and nonlinear analysis function, automatic mesh generation, and friendly develop kit. Nowadays, it has been widely used in the field of aerospace, petrochemical industry, mechanical manufacturing, nuclear industry, railway, energy, defence and civil engineering.

To model the distribution of stress, strain and energy field during the process of rock cross-cut coal uncovering, in view of the influence of ground stress and original gas pressure, killing elements function module is used to simulate excavation. Structural statics module of ANSYS is used to perform the numerical modelling of rock cross-cut coal uncovering and the evolution laws of stress, strain and energy are revealed.

2.1 MODEL DEVELOPMENT

Simulation test of rock cross-cut coal uncovering based on 8# coal seam of Luling coal mine, Huaibei, China, mining area was conducted [14]. The model development and parameter determination of numerical modelling are based on the simulation test.

2.2 PHYSICAL MODEL AND UNIT DIVISION

Geometry size of physical model is shown in Figure 1. Three-dimensional model which is divided into 6 layers including coal seam and the columnar section of coal seam is shown in Figure 2, where the angle of coal seam is 18.44°. The 6 geological strata presented in Figure. 1 are generated based on physical model, as is shown in Figure 3.

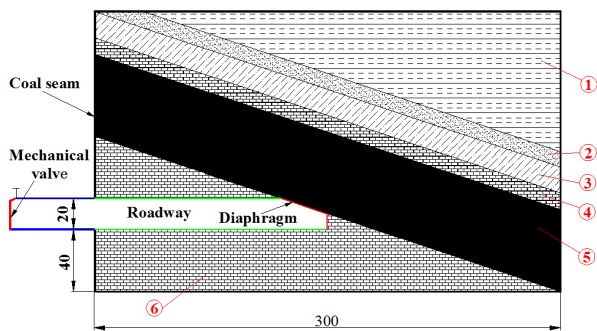


FIGURE 1 Geometry size (cm)

Numbering	Thickness / mm	Columnar section	Lithology
1	36		Siltstone
2	10		Mudstone
3	16		Limestone
4	10		Siltstone
5	50		Coal seam
6	78		Mudstone

FIGURE 2 Columnar section of coal seam

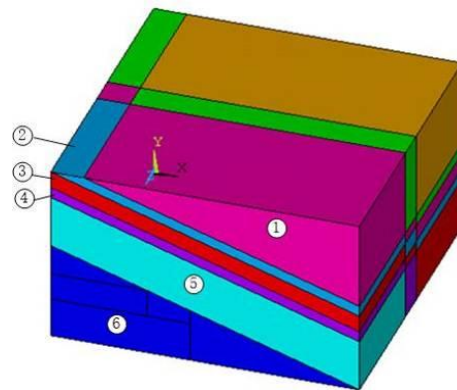


FIGURE 3 Geometric model

Strata of the model are numbered as 1~6# from top to bottom. Rock and coal are assumed to be elastic during the process of excavation. Elastic modulus and Poisson ratio of coal and rock mass in these substrata are shown in Table 1. Finite element model obtained by mapping mesh generation is shown in Figure 4.

TABLE 1 Parameters of coal and rock mass

	Elastic modulus / Pa	Poisson ratio
Coal seam	9.5×10^9	0.350
Stratum 1	1.56×10^{10}	0.290
Stratum 2	1.04×10^{10}	0.238
Stratum 3	1.68×10^{10}	0.200
Stratum 4	9.8×10^9	0.225
Stratum 5	9×10^9	0.280

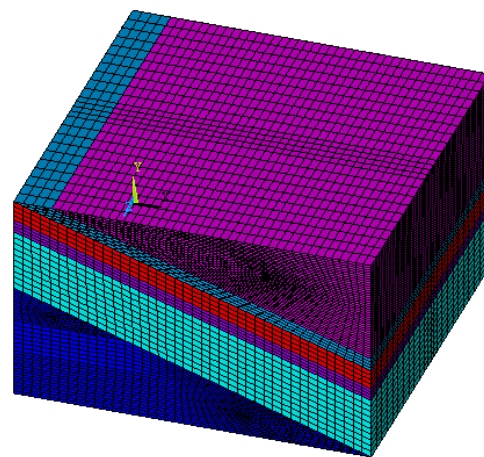


FIGURE 4 Finite element model

2.3 LOAD AND RESTRICTION

Vertical stress of 15 MPa is applied at the top boundary of the model to simulate the ground stress and gas pressure of 1 MPa is equivalent to nodal force of the coal seam which acts on upper and lower planes and components of the nodal force in the direction of X and Y are 2467220 Pa and 7399520 Pa, respectively. Planes of left and right sides, bottom, front and rear of the model are imposed restriction in the direction of X, Y and Z, respectively. Killing elements multiply stiffness of selected units by a quite small coefficient (generally 10⁻⁶), so the influence of stiffness could be neglected for

further solving process and the simulation of excavation could be realized.

3 Simulation Results

3.1 STRESS DISTRIBUTION

The major principal stress and shear stress are main factors influencing destruction of coal-body and outbursts. The modelling of the distribution of major principal stress and shear stress is useful to arrest the law of stress field of surrounding rock during the process of rock cross-cut coal uncovering.

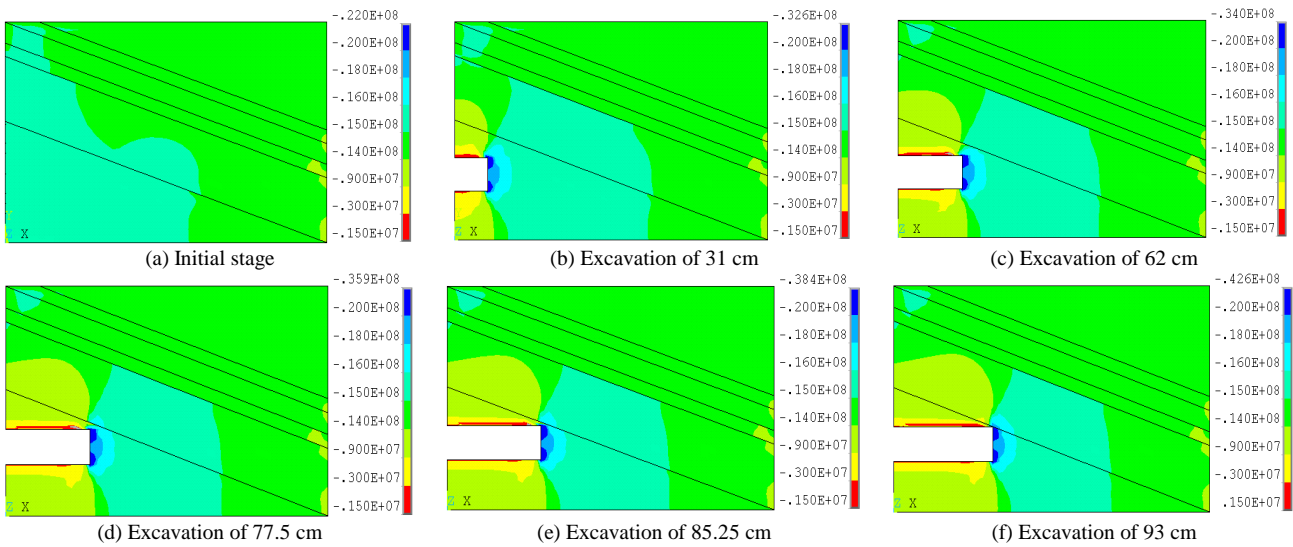


FIGURE 5 Major principal stress distribution during excavation process

During excavation process, large deformation generates in the surrounding rock of roadway. The stress of surrounding rock is redistributed and the influential area is expanded constantly from surrounding area to the far field of country rock with the advance of headway. There is a banding tension stress zone in roof and floor strata of roadway after excavation. The major principal stress in roof and floor strata decreases and it is proportional to the distance from roadway. An umbrella-shaped stress-concentrated area is formed in the coal and rock mass before the heading face and moved forward with the advance of heading face and change little during excavation process. A major principal stress

concentration region is formed in the upper and under area of the anterior heading face, which is symmetrically distributed between medial axis of roadway.

As shown in Figure 6, shear stress concentration regions in opposite direction are formed in the upper and under area of the anterior heading face and moved forward with the advance of heading face. The concentration of shear stress has significant influence on the upper area of the anterior heading face and extends from rock to coal seam. Therefore, that shear stress in coal seam increase and the influence area has an increasing tendency with the advance of heading face.

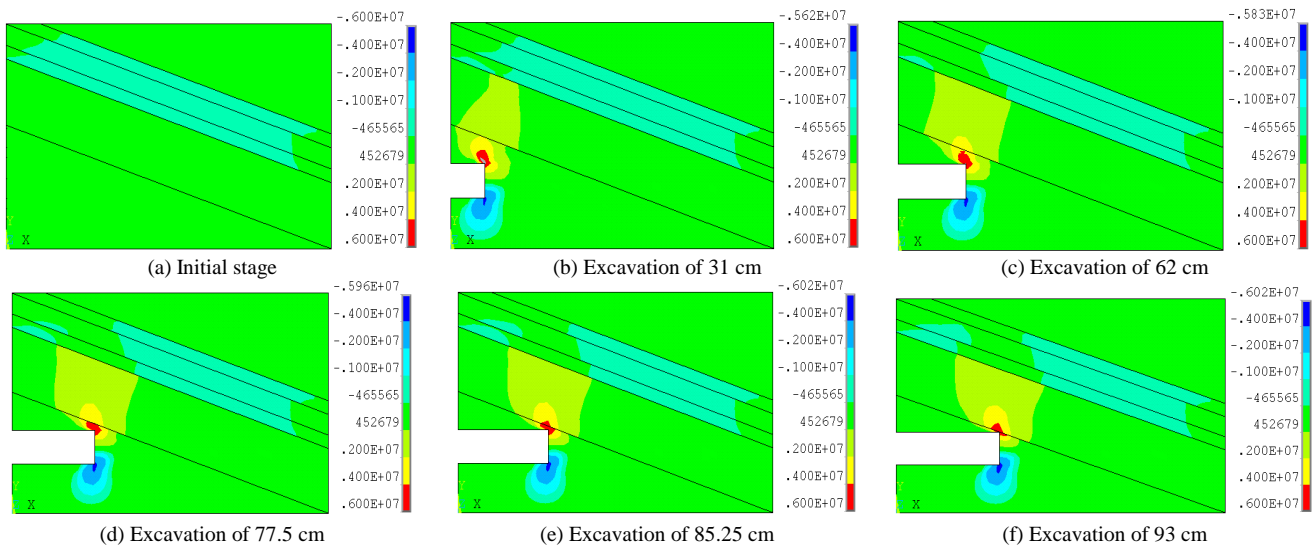


FIGURE 6 Shear stress distribution during excavation process

3.2 STRAIN DISTRIBUTION

The strain distribution of surrounding rock during the process of rock cross-cut coal uncovering is present in Figure 7.

The strain of roof and floor strata changes with the development of roadway. It increases at first and then decreases with the distance from roadway and the strain

direction changes as well. Strain concentration area moves forward and the influential area in coal seam enlarges with the advance of heading face. There is a strain concentrated region in the upper and under area of the anterior heading face and it is symmetrically distributed between medial axis of roadway. However, overlying strata strain of coal seam almost keeps invariant.

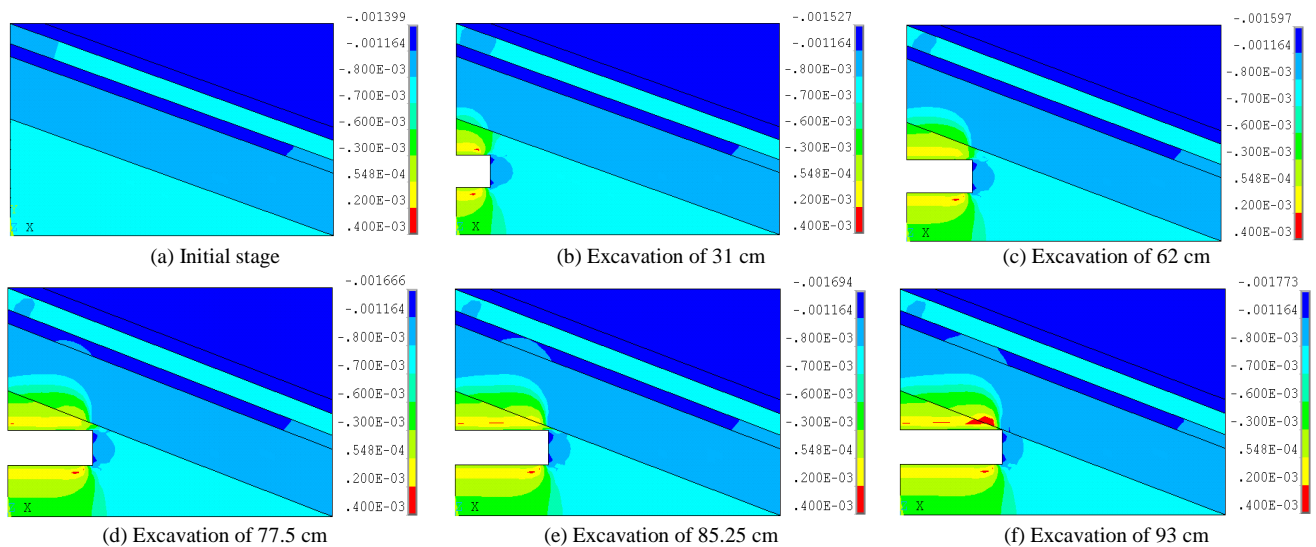


FIGURE 7 Strain distribution during excavation process

3.3 ENERGY DISTRIBUTION

The distributions of stain energy of initial state and excavation process are obtained by the function of Element Table and shown in Figure 8.

With the excavation of roadway, the strain energy of roof and floor strata changes and changed areas enlarge and move forward. Strain energy of rock strata is proportional to the distance from roadway and strain

energy of coffin corner in anterior heading face shows an uneven distribution because of the boundary effect. There is a concentrated area of strain energy in the upper and under area of the anterior heading face and the strain energy in upper area of heading face is larger than that of other areas. While, strain energy of overlying strata of coal seam almost keeps steady.

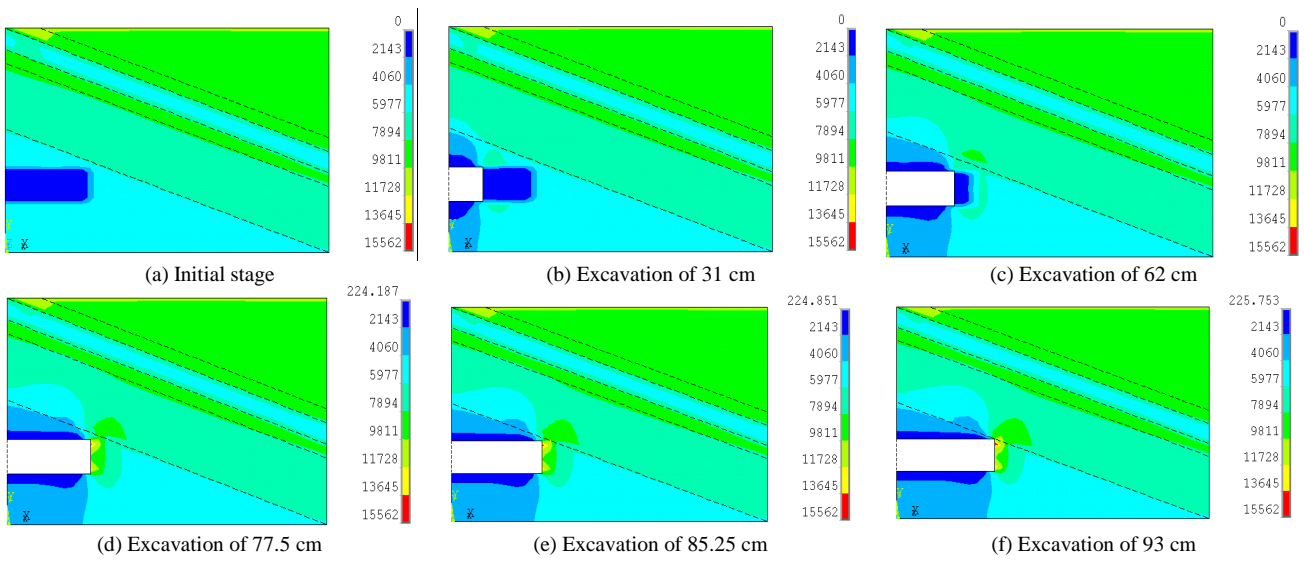


FIGURE 8 Strain energy distribution during excavation process

3.4 STRESS CHANGE IN COAL SEAM

The 5 paths parallel to the horizontal boundaries of coal seam are numbered as path 1~5 from the bottom up and downward tilt direction of coal seam is the positive direction, as is shown in Figure 9. The stress change of coal seam could be caused by the stress change of path point during excavation process. Path 1 is arranged along the nether boundary line of rock and coal and other paths are parallel to path 1 and vertical distance of each path is 10 cm.

As is shown in Figure 10, stress of the coal seam changes during the excavation process and stress change of coal seam area nearer to the roadway is more obvious. Stress of coal seam over the excavated area decreases and the minimum stress is in path 1 and followed by path 2 and stress decreasing zones move forward with the with the advance of heading face. There is a stress-increasing

area of a certain distance from the starting point. When the excavation distance is 77.5 and 93 cm, the stress-increasing areas are at the distance of 80-120 and 100-140 cm respective.

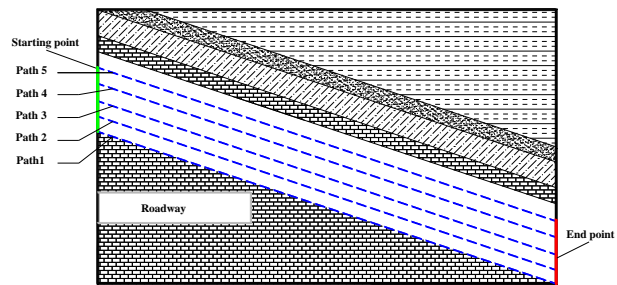
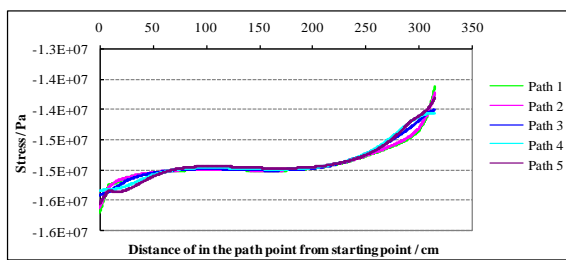
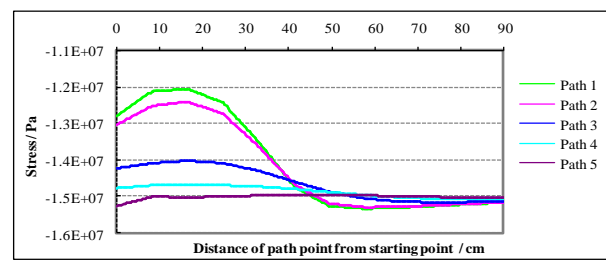


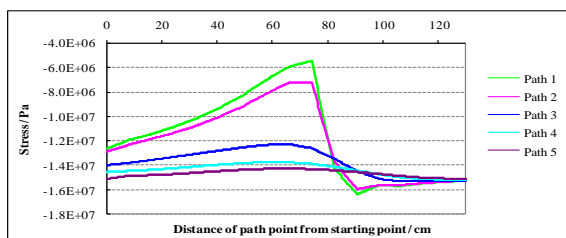
FIGURE 9 Arrangement of path in the coal seam



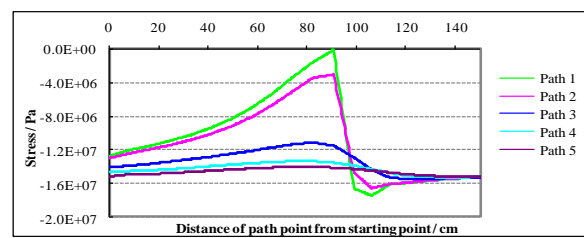
(a) Initial stage



(b) Excavation of 31 cm



(c) Excavation of 77.5 cm



(d) Excavation of 93 cm

FIGURE 10 Stress change of path in the coal seam

4 Discussion of results

During the excavation process, stress of surrounding rock is redistributed and stress concentration region is formed. When stress exceeds the strength limit of rock, surrounding rock of cross-cut would yield large deformation which extends from surrounding rock to deep rock mass. As the strength of rock is much higher than that of coal, when the heading face is far from the coal seam, stress of rock and coal and gas pressure accumulated in the coal could not crack the rock between heading face and coal seam, which bears the stress mainly. However, with the advance of the heading face, the strata thickness turns thin so that the stress affordability decreases and stress of rock and coal and gas pressure accumulated in the coal could crack the rock and this may induce the occurrence of outburst.

As rock has high stress affordability and its permeability is much lower than coal, so before uncovering the coal, motion quantity between coal and rock is small and gas accumulated in the coal releases little as well as stress of rock and coal, as a result, there will be large stress and gas pressure gradient. This can be the important factor of inducing outburst of coal and gas. At the moment of uncovering the coal, lateral stress of the uncovering area releases suddenly and the state of three dimensions of applied force turns to two dimensions. Unbalanced forced state causes rapid failure of coal and rock and this is an important factor inducing outburst.

The existence of the principal and shear stress concentration region formed in the upper and under area of the anterior heading face accumulates high strain energy and contributes to the instability and failure of coal and rock mass. The upper area with high stress is the nearest to coal seam and its destruction-resisting ability is weak. If there is, a large amount of gas accumulated in the coal, outburst is potential to occur under the disturbance of mining and the upper area is the most easily to be the releasing port of outburst.

5 Conclusions

The stress, strain and energy of roadway surrounding rock during the process of rock cross-cut coal uncovering

References

- [1] Wang G 2012 *Study on prediction of coal and gas outburst by gas content, doctoral dissertation* Qingdao: Shandong University of Science and Technology (in Chinese)
- [2] Liu Y 2009 *Technology of Improving coal-seam permeability with high pressure pulsed water jet and its application on rock cross-cut coal uncovering, master degree thesis* Chongqing: Chongqing University (in Chinese)
- [3] Paterson L 1986 A model for outbursts in coal *International Journal of Rock Mechanics & Mining Sciences & Geomechanics Abstract* 23(4) 327-32

could be numerically simulated by ANSYS and the simulation results could be regarded as the analysis basis of physical and mechanical property of surrounding rock during the process of rock cross-cut coal uncovering.

Stress field distribution of coal and rock mass is relatively uniform before excavation and a banding tension stress zone is formed in roof and floor strata after excavation. A principal and shear stress concentration area is formed in the upper and lower area of the anterior heading face and it is symmetrically distributed between medial axis of roadway and stress change of coal seam area nearer to the roadway is more obvious, while the changes of stress, strain and strain energy of overlying strata are unapparent.

Gas accumulated in the coal is difficult to release before uncovering the coal and lateral stress of the uncovering area releases suddenly after uncovering the coal. There would be large stress and gas pressure gradient at the moment of uncovering the coal. Hence, rock cross-cut coal uncovering could easily induce outburst of coal and gas.

The concentrations of the principal and shear stress formed in the upper area of the anterior heading face could contribute to the instability and failure of coal and rock mass and the upper area with high stress is the nearest to coal seam and its destruction-resisting ability is weak. The upper area is the most easily to be the releasing port of outburst and should be regarded as the key region of outburst prevention.

Acknowledgements

The authors would like to acknowledge the support of National Natural Science Foundation Project (51304128), Scientific Research Foundation of Shandong University of Science and Technology for Recruited Talents (2013RCJJ049), China Postdoctoral Science Foundation (2013M541942), Shandong Provincial Natural Science Foundation, China (ZR2013EEQ015) and Specialized Research Fund for the Doctoral Program of Higher Education (20133718120013).

- [4] Otuonye F, Sheng J 1994 A numerical simulation of gas flow during coal / gas outbursts *Geotechnical and Geological Engineering* 12(1) 15-34
- [5] Odintsev V N 1997 Sudden outburst of coal and gas-failure of natural coal as a solution of methane in a solid substance *Journal of Mining Science* 33(6) 508-16.
- [6] Xu T, Tang C A, Yang T H, Zhu W C, Liu J 1996 Numerical investigation of coal and gas outbursts in underground collieries *International Journal of Rock Mechanics & Mining Sciences* 43(6) 905-19
- [7] Wold M B, Connell L D, Choi S K 2008 The role of spatial variability in coal seam parameters on gas outburst behaviour during coal mining *International Journal of Coal Geology* 75(1) 1-14

[8] Xue S, Wang Y C, Xie J, Wang G 2011 A coupled approach to simulate initiation of outbursts of coal and gas – Model development *International Journal of Coal Geology* **86**(2) 222-30

[9] Tang C A, Liu H Y 2002 Numerical approach on outburst in crosscutting in coal seam containing gas *Chinese Journal of Rock Mechanics and Engineering* **21**(10) 1467-72 (in Chinese)

[10] Zhang C H 2007 *Numerical simulation on mechanical characteristic of surrounding rock in rock cross-cut*, master degree thesis Huainan: Anhui University of Science and Technology (in Chinese)

[11] He F 2009 *Coal-rock mechanical characteristic analysis of uncovering coal in crosscut and numerical simulation*, master degree thesis Huainan: Anhui University of Science and Technology (in Chinese)

[12] Sun D S, Zhao W H, Wang H C, Zhang J H 2009 Numerical approach on outburst in crosscutting in coal seam containing gas at Hongling coal mine *China Mining Magazine* **18**(5) 91-4 (in Chinese)

[13] Lu Y Y, Ge Z L, Li X H, Chen J F, Liu Y 2010 Investigation of a self-excited pulsed water jet for rock cross-cutting to uncover coal *Journal of China University of Mining & Technology* **39**(1) 55-8

[14] Su C Q 2011 Research on mechanical characteristics and experiment of outburst in uncovering coal seam in cross-cut, doctoral dissertation, Shandong: Shandong University of Science and Technology (in Chinese)

Authors	
	<p>G Wang, born in September, 1984, Linyi, Shangdong, China</p> <p>Current position, grades: the Associate Professor of College of Mining and Safety Engineering, Shandong University of Science and Technology, China.</p> <p>University studies: received his B.E. in Mining Engineering from Shandong University of Science and Technology in China. He received his M.E. from Shandong University of Science and Technology University in China. And He received his D.E. from Shandong University of Science and Technology University in China.</p> <p>Scientific interest: His research interest fields include coal and gas outburst, mine ventilation.</p> <p>Publications: more than 20 papers published in various journals.</p> <p>Experience: He has teaching experience of 3 years, has presided five scientific research projects.</p>
	<p>C Q Su, born in August, 1969, Liping, Guizhou, China</p> <p>Current position, grades: the Engineer of Guizhou Forestry Mining Group, China.</p> <p>University studies: received his B.E. in Mining Engineering from Hunan University of Science and Technology in China. He received his D.E. from Shandong University of Science and Technology University in China.</p> <p>Scientific interest: His research interest fields include mine ventilation, mine fire prevention.</p> <p>Publications: more than 8 papers published in various journals.</p>
	<p>H Y Wang, born in August, 1988, Zibo, Shangdong, China</p> <p>Current position, grades: postgraduate student of College of Mining and Safety Engineering, Shandong University of Science and Technology, China.</p> <p>University studies: received his B.Sc. in Safety Engineering from Shandong University of Science and Technology in China.</p> <p>Scientific interest: His research interest fields include gas disaster prevention and coal spontaneous combustion.</p> <p>Publications: 7 papers published in various journals.</p> <p>Experience: He has participated in 10 more scientific research projects.</p>
	<p>Q M Huang, born in July, 1990, Jining, Shangdong, China</p> <p>Current position, grades: postgraduate student of College of Mining and Safety Engineering, Shandong University of Science and Technology, China.</p> <p>University studies: received his B.E. in Mining Engineering from Shandong University of Science and Technology in China.</p> <p>Scientific interest: His research interest fields include coal and gas outburst, mine disaster prevention.</p>

UKF-based underground intrusion localization algorithm for optical-fibre sensing perimeter protection

Hua Zhang^{1*}, Xiaoping Jiang¹, Chenghua Li¹

¹ College of Electronics and Information Engineering, Hubei Key Laboratory of Intelligent Wireless Communications, South-Central University for Nationalities, 182 Minzu Road, Wuhan, Hubei, China

Received 1 March 2014, www.tsi.lv

Abstract

To improve the precision of the underground intrusion localization in the optical-fibre sensing perimeter protection application, an intrusion localization algorithm based on the Unscented Kalman Filter (UKF) is presented. The geometrical relationships of the sensors and the intruder are analysed and the state equation and the measurement model are deduced. Then the UKF algorithm is used to estimate and track the location of the intruder. The simulations demonstrate that the algorithm improves the intrusion localization precision and the intruder can be tracked even if no enough sensors detect the intrusion signal.

Keywords: Optical-Fibre Sensor, Underground Intrusion Detection, State Estimation, Unscented Kalman Filter

1 Introduction

The optical fibre sensing-based intrusion detection technologies have been widely used in perimeter security protection systems, with the characters of high vibrational sensitivity together with electromagnetic interference immunity. The optical fibre sensing technologies used for intrusion detection include the interferometer-based optical fibre sensors and the optical time domain reflectometry (OTDR)-based optical fibre sensors [1-9]. The OTDR-based optical sensors are sensitive to very low vibrations and can be used in intrusion detection. However, it is subject to the quality of the laser and costly [8]. The Sagnac interferometer-based optical fibre sensing system is of high sensitivity to vibrational disturbances and low cost [2, 3]. The Mach-Zehnder interferometer based optical fibre sensing technologies have the same property of high phase-sensitivity as the OTDR-based technologies and have been studied widely [4]. To improve the performance of the perimeter security system, the distributed optical fibre sensing system has been used in intrusion detecting systems [1].

In a perimeter protection system, it is important to localize the intruder when an intrusion signal is detected. The need for intrusion localization is more necessary for an underground perimeter protection system to reduce the rate of false alarm. Generally, the underground intrusion signals to be detected are acoustic (or vibrational) signals generated by the intruder. When an intrusion occurs, the waveforms sampled in the sensors are processed and analysed in amplitudes, phases and frequencies to judge the intrusion, and the time of arrival (TOA) of the intrusion signal is used to locate the position of the intruder approximately [9]. Actually, the properties of the

received intrusion signals are studied to localize the intruder by many researchers [10]. As in the interferometer-based optical-fibre sensor system, the time interval between the moment the laser was sent out and the moment the intrusion signal arrives at the receiver can not be got where the consecutive laser pulses are used. To get the precise TOAs of the intrusion signals, many signal-processing algorithms were employed [11]. However, the approaches suffer from the measurement errors for the fast speed of the laser propagating in the optical fibre, the errors of the time limit the precision of the intrusion localization to tens of meters [12].

In this paper, a state estimation based intrusion localization algorithm is proposed to get high precise underground intrusion localization estimation. The geometrical relationships of the distributed sensors and the intruder are analysed and the state equation and the measurement model are deduced. Then the Unscented Kalman Filter (UKF) is used to estimate and track the location of the intruder. The simulation demonstrates that the algorithm improves the intrusion localization precision and the intruder can be tracked even if no enough sensors detect the intrusion signal.

2 Intrusion Localization Algorithm based on the Geometric Relationship of the Sensors and the Intruder

As mentioned above, the optical-fibre sensor-based intrusion detecting technologies include the interferometer-based methods and the OTDR-based methods. Although the principles of the two methods are distinct, to detect the underground intrusion signals, both methods use the optical fibre sensors to detect the

* Corresponding author e-mail: zhanghua@mail.scuec.edu.cn

acoustic signals generated by the intruder. The acoustic signals may be generated by the excavating or ambulating actions of the intruders. As the acoustic signals are detected by the optical fibre sensors, the signals are processed and the TOAs are got. Then the geometrical equations can be deduced from the locations of the distributed sensors and the differences of the TOAs. For convenience, a Sagnac interferometer-based optical fibre sensing system is used to detect the underground intrusion in our work.

When the underground intrusion signal occurs, the resultant vibrations or acoustic signals can be detected by the optical-fibre sensors. The detected signals may be of the different phases or be of various light intensities. As the sensing light signals are converted to the electric signals and then converted into digital signals, the digital signals are processed and analysed in amplitudes and frequencies. And the signals with certain amplitudes and waveforms are judged as the markers of the intrusion. Then the time intervals between the moment the laser pulse was sent out and the moment the 0-phase of the intrusion signal waveform can be got. The time interval is the time the acoustic signal costs in propagating from the intruder to the sensor. So it is called time of arrival (TOA) of the intrusion signal.

When an intrusion signal is detected, the moment t which is called TOA can be gotten by signal processing. If the moment the intruder generated the vibrational signal is t_0 , the time interval $(t-t_0)$ includes the time for the acoustic signal arriving at the sensor and the time for the laser propagating in the optical fibre. As shown in Figure 1, the distance from the intruder to the sensor is equal to the distance the acoustic signal transports in TOA of the intrusion signal. As the locations of the sensors are fixed, the time for the laser is almost constant and can be calibrated previously. Then the geometrical relationship between the sensor i and the intruder is:

$$\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2} = v_l(t-t_0-T_i), \quad (1)$$

where, (x_i, y_i, z_i) is the location of the i^{th} sensor, (x, y, z) is the location of the intruder, and v_l is the transporting speed of the vibrational signal generated by the intruder, t_0 is the moment the intruder generated the vibrational signal, T_i is the time for the laser propagating in the optical fibre of the i^{th} sensor, and t_i is the moment at which the intrusion signal in the i^{th} sensor is detected in the receiver.

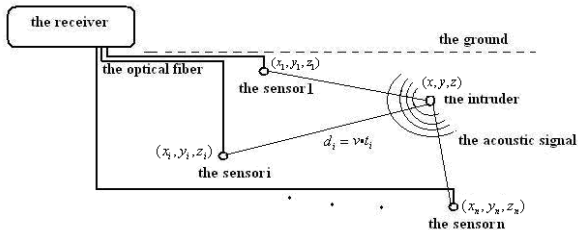


FIGURE 1 In an underground perimeter protection system, the distances from the intruder to the sensors are equal to the distance which the acoustic signals transport from the intruder to the sensors.

As in the Equation (1), the parameters t_0 and v_l are unknown, the parameters can be ignored by using the different distances of the sensors,

$$\begin{aligned} & \sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2} - \sqrt{(x-x_j)^2+(y-y_j)^2+(z-z_j)^2} \\ & = v_l \cdot \left\| (t_i-t_j-(T_i-T_j)) \right\| \end{aligned}, \quad (2)$$

where $\|(\cdot)\|$ is the absolute value sign.

As long as the intrusion signal is detected by enough sensors, the parameters (x, y, z) and v_l can be computed with optimal estimation methods such as the Least-Square methods. The Least-Square based intrusion localization algorithm is as follows:

Firstly, the equation (2) can be written as,

$$\begin{aligned} & \sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2} \\ & - \sqrt{(x-x_j)^2+(y-y_j)^2+(z-z_j)^2} \\ & - v_l \left\| (t_i-t_j-(T_i-T_j)) \right\| = 0 \end{aligned} \quad (3)$$

Then the left side of the equation (3) can be denoted by a function,

$$\begin{aligned} f_{ij}(x, y, z, v) &= \sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2} \\ & - \sqrt{(x-x_j)^2+(y-y_j)^2+(z-z_j)^2} \\ & - v_l \left\| (t_i-t_j-(T_i-T_j)) \right\| \end{aligned} \quad (4)$$

To solve the equation (3), use the linearization about the nominal value $(\hat{x}, \hat{y}, \hat{z}, \hat{v})$,

$$x = \hat{x} + \Delta x, y = \hat{y} + \Delta y, z = \hat{z} + \Delta z, v = \hat{v} + \Delta v. \quad (5)$$

Substitute equation (5) to equation (4), we get

$$\begin{aligned} f_{ij}(x, y, z, v) &= f_{ij}(\hat{x}, \hat{y}, \hat{z}, \hat{v}) \\ & + \left[\frac{\partial f_{ij}}{\partial x} \quad \frac{\partial f_{ij}}{\partial y} \quad \frac{\partial f_{ij}}{\partial z} \quad \frac{\partial f_{ij}}{\partial v} \right] \begin{Bmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta v \end{Bmatrix} \end{aligned} \quad (6)$$

From equation (3),

$$R_{ij} = 0 - f_{ij}(\hat{x}, \hat{y}, \hat{z}, \hat{v}) = \left[\frac{\partial f_{ij}}{\partial x} \quad \frac{\partial f_{ij}}{\partial y} \quad \frac{\partial f_{ij}}{\partial z} \quad \frac{\partial f_{ij}}{\partial v} \right] \begin{Bmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta v \end{Bmatrix} \quad (7)$$

If there are n sensors, which detected the intrusion, for $(i=1,2,\dots,n-1)$ and $(j=2,3,\dots,n)$ the equation (7) can be written as,

$$\{R\} = \{A\} \cdot \{\Delta E\}, \tag{8}$$

where,

$$R = \begin{Bmatrix} R_{12} \\ R_{13} \\ \cdot \\ \cdot \\ R_{n-1,n} \end{Bmatrix}, A = \begin{Bmatrix} \frac{\partial f_{12}}{\partial x} & \frac{\partial f_{12}}{\partial y} & \frac{\partial f_{12}}{\partial z} & \frac{\partial f_{12}}{\partial v} \\ \frac{\partial f_{13}}{\partial x} & \frac{\partial f_{13}}{\partial y} & \frac{\partial f_{13}}{\partial z} & \frac{\partial f_{13}}{\partial v} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial f_{n-1,n}}{\partial x} & \frac{\partial f_{n-1,n}}{\partial y} & \frac{\partial f_{n-1,n}}{\partial z} & \frac{\partial f_{n-1,n}}{\partial v} \end{Bmatrix} \tag{9}$$

and

$$\Delta E = (\Delta x \ \Delta y \ \Delta z \ \Delta v)^T. \tag{10}$$

Then the method of least squares can be used to get the optimal estimation of the nominal value $(\hat{x}, \hat{y}, \hat{v})$. Here we use the minimal residual method of least squares by solving ΔZ ,

$$\{\Delta E\}_k = [A_k^T A_k]^{-1} \cdot A_k^T \cdot \{R\}_k, \tag{11}$$

where $k=1, 2, \dots$ is the number of iterations, and the estimation in k -step iteration is

$$\begin{Bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \\ \hat{v} \end{Bmatrix}_k = \begin{Bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \\ \hat{v} \end{Bmatrix}_{k-1} + \begin{Bmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta v \end{Bmatrix}_{k-1}. \tag{12}$$

The partial derivative terms in equation (7), (9) are given by

$$\begin{cases} \frac{\partial f_{ij}}{\partial x} = \frac{(x-x_i)}{\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2}} - \frac{(x-x_j)}{\sqrt{(x-x_j)^2+(y-y_j)^2+(z-z_j)^2}} \\ \frac{\partial f_{ij}}{\partial y} = \frac{(y-y_i)}{\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2}} - \frac{(y-y_j)}{\sqrt{(x-x_j)^2+(y-y_j)^2+(z-z_j)^2}} \\ \frac{\partial f_{ij}}{\partial z} = \frac{(z-z_i)}{\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2}} - \frac{(z-z_j)}{\sqrt{(x-x_j)^2+(y-y_j)^2+(z-z_j)^2}} \\ \frac{\partial f_{12}}{\partial v} = -(t_i-t_j-(T_i-T_j)) \end{cases} \tag{13}$$

The algorithms from equation (4) to (13) above are repeated recursively, the iteration is going on until ΔE is less than a set tolerance. However, as the geometrical

relationship in Equation (2) does not consider the noises in the parameters, the number of the iterations of the algorithm may be too large and the algorithm results in bad precision. Especially, when the number of the distributed sensors, which detected the intrusion signal is less than 4, the errors of the location estimation increase remarkably.

3 UKF-based Intrusion Localization Algorithm

To improve the accuracy of the location estimation of the intruder, the state estimation methods can be used to track the location of the intruder when the measurement noises and the system noises are considered. The state equations and the measurement model are deduced from the geometric relationship in equation (2) and the UKF algorithms are used for state estimation.

3.1 THE SYSTEM EQUATION AND THE MEASUREMENT MODEL FOR INTRUSION LOCALIZATION

As in equation (2), the speed of the vibrational signals propagating underground is unknown. To improve the precision of the location estimated, the unknown speed of the vibrational signals and the moment the intruder generated the vibrational signals as well as the location of the intruder and the moving speed of the intruder, are considered as the state parameters, i.e., $X = [x \ y \ z \ v_x \ v_y \ v_z \ v_t \ t_0]^T$. For simplification, the moving speed of the intruder is considered almost constant, i.e. the variation of the moving speed is zero, and a zero-mean Gaussian noise is added. Moreover, the speed of the vibrational signal is considered constant and the zero-mean Gaussian noise is added. Then we get the state equations as:

$$\dot{X} = AX + W, \tag{14}$$

where,

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{15}$$

And W is the noise vector of the state parameters, the means of which are considered zero and the covariance matrix $R = diag(\sigma_1 \sigma_2 \dots \sigma_8)$.

The measurement parameters are the moments when the intrusion signal arrives at the sensors, t_i , i.e., $Y = [t_1 t_2 \dots t_n]^T$. From Equation (1), we get

$$t_i = \frac{1}{v_i} \cdot \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2} + t_0 + T_i. \quad (16)$$

Then the measurement model can be written as:

$$Y = G(X) + V, \quad (17)$$

where, $G(\cdot)$ is the measurement function vectors as noted in Equation (5), and V is the measurement noise vector, with the mean m and the covariance matrix $Q = \text{diag}(\sigma_{t_1} \sigma_{t_2} \dots \sigma_{t_n})$.

3.2 THE UKF ALGORITHM

As the measurement equation in (17) is nonlinear, the original Kalman Filter cannot be used for state estimation directly. If the equation is linearized, the extended Kalman Filter (EKF) can be used to estimate the location of the intruder. However, the errors in the linearization may result in state estimation errors [13, 14]. As the unscented Kalman Filter (UKF) algorithm is excellent in nonlinear system, it is adopted in our work [15, 16]. The algorithm can be described as follows:

3.2.1 Initiation

At first, the initial mean and covariance of the 8-dimensional state sector can be computed as:

$$\hat{X}_0 = E(X_0), \quad (18)$$

$$P_0 = E[(X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^T]. \quad (19)$$

3.2.2 Sigma Point Sample and the Weight

The symmetric sampling method is used and $(2n+1)$ points $\{x_i(k|k) | i=0, 1, \dots, 2n, k \geq 1\}$ are sampled. The points and the weight are selected as follows:

$$\chi_0(k-1) = \hat{X}(k-1), \quad (20)$$

$$\chi_i(k-1) = \hat{X}(k-1) + \left(\sqrt{(n+\lambda)P_{xx}(k-1)} \right)_i, \quad i=1, \dots, n, \quad (21)$$

$$\chi_i(k-1) = \hat{X}(k-1) - \left(\sqrt{(n+\lambda)P_{xx}(k-1)} \right)_i, \quad i=n+1, \dots, 2n, \quad (22)$$

where, n is the dimension of feature state, $\lambda = \alpha^2(n+\kappa) - n$ is a scale parameter. The α is constant which determines the spread of the sigma points around $\hat{X}(k-1)$ and is usually set to a small positive value. In addition, the constant κ is another scale parameter, which is set to $(3-n)$. β is used to incorporate prior knowledge of the distribution of the system states.

In addition, two weights ω_i^m and ω_i^c are used to compute the mean and covariance of the state estimation:

$$\omega_0^m = \frac{\lambda}{n + \lambda}, \quad (23)$$

$$\omega_0^c = \frac{\lambda}{n + \lambda} + (1 - \alpha^2 + \beta), \quad (24)$$

$$\omega_i^m = \omega_i^c = \frac{1}{2(n + \lambda)}, \quad i=1, \dots, 2n. \quad (25)$$

3.2.3 Time Update

The predicted mean and covariance are computed as follows:

$$\chi_i(k|k-1) = tA\chi_i(k-1), \quad (26)$$

$$\hat{X}(k|k-1) = \sum_{i=0}^{2n} \omega_i^m \chi_i(k|k-1), \quad (27)$$

$$P_{xx}(k|k-1) = \sum_{i=0}^{2n} \omega_i^m [\chi_i(k|k-1) - \hat{X}(k|k-1)][\chi_i(k|k-1) - \hat{X}(k|k-1)]^T, \quad (28)$$

$$\hat{X}(k|k-1) = \sum_{i=0}^{2n} \omega_i^c \chi_i(k|k-1), \quad (29)$$

$$Y_i(k|k-1) = G(\chi_i(k|k-1)), \quad (29)$$

$$\hat{Y}(k|k-1) = \sum_{i=0}^{2n} \omega_i^c Y_i(k|k-1). \quad (30)$$

3.2.4 Measurement Update

Moreover, the predicted observation mean, innovation covariance and the cross relation matrix are computed as follows:

$$P_{yy}(k) = \sum_{i=0}^{2n} \omega_i^c [Y_i(k|k-1) - \hat{Y}(k|k-1)][Y_i(k|k-1) - \hat{Y}(k|k-1)]^T, \quad (31)$$

$$\hat{Y}(k|k-1) = \sum_{i=0}^{2n} \omega_i^c Y_i(k|k-1)$$

$$P_{xy}(k) = \sum_{i=0}^{2n} \omega_i^c [\chi_i(k|k-1) - \hat{X}(k|k-1)][Y_i(k|k-1) - \hat{Y}(k|k-1)]^T, \quad (32)$$

$$K(k) = P_{xy}(k)P_{yy}^{-1}(k), \quad (33)$$

$$\hat{X}(k) = \hat{X}(k|k-1) + K(k)[Y(k) - \hat{Y}(k|k-1)], \quad (34)$$

$$P_{xx}(k) = P_{xx}(k|k-1) - K(k)P_{xy}(k)K^T(k). \quad (35)$$

The algorithms above are repeated from equation (20) to (35), the iteration is going on to estimate and track the location of the intruder.

4 Simulations and Experiments

To test the precision of the intrusion localization algorithm proposed, simulations are performed in given data. In the simulations, the distributed Sagnac-based optical-fibre sensors are used and the sensors assumed to be located in lines and rows as shown in Figure 2, and the distance between each pair of the neighbouring sensors is 50 meters, and all the sensors are assumed to be buried 1.5 meters below the ground. The propagating speed of the vibrational signal the intruder generated underground is assumed to be constant, i.e. 1000m/S. Generally, the sampling rate of the receiver is above 10k times per second. So the errors of the TOAs are considered below 0.1mS and the measurement noise is considered zero mean and the covariance 0.1. The intruder is considered moving in a speed 1m/S, and the initial $R = \text{diag}(1,1,1,1,1,1,1)$. The number of signal points is set to 21 in simulation.

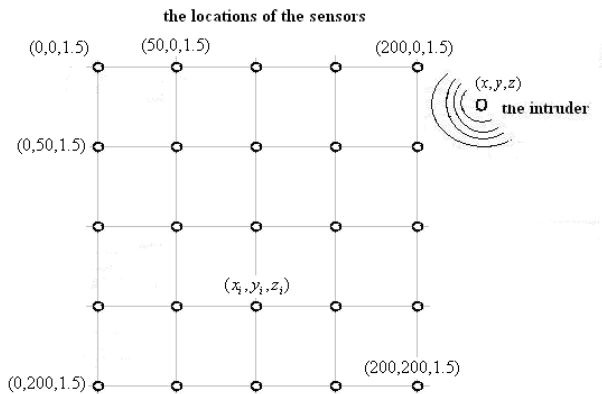


FIGURE 2 The locations of the sensors for simulation

Figure 3 depicts the result of the simulations. The error of the locations is below 0.5 meters and the locations of the intruder can be tracked precisely. Moreover, the errors of the estimation under various numbers of the sensors, which detected the intrusion are listed in table 1. It demonstrates that even if only one

sensor detects the intruder, the algorithm can track the intruder precisely.

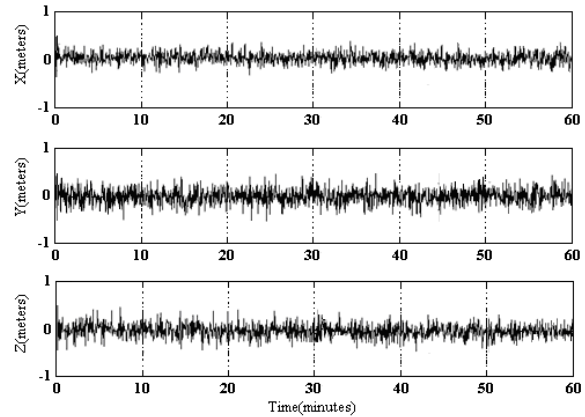


FIGURE 3 The Simulation of the Intruder Localization Estimation

TABLE 1 The statistical errors with various number of sensors, which detected the intrusion

Number of the sensors	The parameters	Errors of estimation (meters)
1	x	1.46
1	y	1.722
1	z	2.21
3	x	0.73
3	y	0.92
3	z	0.88
5	x	0.18
5	y	0.22
5	z	0.3

5 Conclusions

To improve the precision of the underground intrusion localization in the optical-fibre sensing perimeter protection application, an UKF-based intrusion localization algorithm is proposed in the paper. The state equation and the measurement model are deduced from the geometrical relationship of the sensors and the intruder, and the UKF algorithm is used to estimate and track the location of the intruder. The simulation demonstrates that the algorithm improves the intrusion localization precision and the intruder can be tracked even if no enough sensors detect the intrusion signal.

Acknowledgments

This work was supported by the open fund of Hubei Key Laboratory of Intelligent Wireless Communications under Grant No. IWC2012014, the General Program of the Natural Science Fund of Hubei Province, China under Grant No. 2012FFC13301, the Key Technologies R&D Program of Wuhan City, China under Grant No. 201212521825 and the General Program of National Natural Science Foundation of China under Grant No. 61201448.

References

- [1] Juarez J C, Maier E W, et al 2005 Distributed Fiber-Optic Intrusion Sensor System *Journal of Lightwave Technology* **23**(6) 2081-87
- [2] Szustakowski M, Yczkowski M, et al 2004 Sensitivity of perimeter sensor based on Sagnac interferometer *Light guides and Their Applications* **5576** 319-23
- [3] Culshaw B 2006 The optical fibre Sagnac interferometer: An overview of its principles and applications *Measurement Science and Technology* **17** 1-16
- [4] Jahed N M S, Nurmohammadi T, et al 2009 Enhanced resolution fiber optic strain sensor based on Mach-Zehnder interferometer and displacement sensing principles *Proc. ELECO International Conference on Electrical and Electronics Engineering Turkish/American IEEE: Bursa* p II-302
- [5] Linze N, Mégret P, et al 2012 Development of an Intrusion Sensor Based on a Polarization-OTDR System *IEEE Sensors Journal* **12**(10) 3005-09
- [6] Li X L, Sun Q Z, et al 2012 Hybrid TDM/WDM-Based Fiber-Optic Sensor Network for Perimeter Intrusion Detection *Journal of Lightwave Technology* **30**(8) 1113-20
- [7] Zyczkowski M, Szustakowski M, et al 2004 Fiber optic perimeter protection sensor with intruder localization *Proc. SPIE on Unmanned/Unattended Sensors and Sensor Networks* **5611** UK/American SPIE: London p 71
- [8] Park J 2003 Fiber optic intrusion sensor using coherent optical time domain reflectometer *Japanese Journal of Applied Physics* **42**(6a) 3481-82
- [9] Mcaulay A D, Wang Jian 2004 A Sagnac interferometer sensor system for intrusion detection and localization *Proc. SPIE on Enabling Photonic Technologies for Aerospace Applications* **5435** American SPIE: Orlando FL p 114
- [10] Kondrat M, Szustakowski M, et al 2006 Two-interferometer fiber optic sensor with disturbance localization *Proc. SPIE on Unattended Sensors and Sensor Networks* **6394** Sweden/ American SPIE: Stockholm p 0T-1
- [11] Zhang Y, Chen J M 2012 Location method of Distributed Fiber-optic Perimeter Security System Based on Mach-Zehnder Interferometer *Chinese Journal of Lasers* **39**(6) 0605005-1-4
- [12] Zyczkowski M, Ciurapinski W 2007 Fiber optic sensor with disturbance localization in one optical fiber *Proc. SPIE, on Optical Sensing Technology and Applications* **6585** Czech Republic/ American SPIE: Prague p 1K-1
- [13] Belkacem S, et al 2010 Robust Non-Linear Direct Torque and Flux Control of Adjustable Speed Sensorless PMSM Drive Based on SVM Using a PI Predictive Controller *Journal of Engineering Science and Technology Review* **3**(1) 168-75
- [14] Su S C, Zhang W 2013 Fault Prediction of Nonlinear System Using Time Series Novelty Estimation *Journal of Digital Information Management* **11**(3) 207-12
- [15] Juliet S J, Uhlmann J K, et al 2000 A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators *IEEE Transactions on Automatic Control* **AC-45**(3) 477-82
- [16] Crassidis J L, Markley F L 2003 Unscented Filtering for Spacecraft Attitude Estimation *Journal of Guidance and Dynamics* **26**(4) 536-42

Authors



Hua Zhang, born in February, 1972, Wuhan City, Hubei Province, P.R. China

Current position, grades: Lecturer of College of Electronics and Information Engineering, South-Central University for Nationalities, P.R. China.

University studies: Graduated from Huazhong University of Science and Technology, China in 2011, received a doctor's degree in Control Science and Engineering.

Scientific interest: Autonomous Navigation, Image Analysis and Intelligent Control.

Publications: Presided over 3 scientific research projects the completion of provincial; more than 10 papers published in various journals.

Experience: Graduated from Huazhong University of Science and Technology, China in 2011, received a doctor's degree in Control Science and Engineering, was approved as a tutor of graduate students in 2013; has completed 3 scientific research projects; more than 10 papers published in various journals.



Xiaoping Jiang, born in February, 1974, Wuhan City, Hubei Province, P.R. China

Current position, grades: Associate professor of College of Electronics and Information Engineering, South-Central University for Nationalities.

University studies: Graduated from Huazhong University of Science and Technology in 2007, received a doctor's degree in Communication Engineering.

Scientific interest: Communication signal Processing, Intelligent signal analysis and multimedia processing.

Publications: Presided over 5 scientific research projects the completion of national and provincial; more than 10 papers published in various journals.

Experience: Graduated from Huazhong University of Science and Technology in 2007, received a doctor's degree in Communication Engineering, was approved as a tutor of graduate students in 2011; has completed 4 scientific research projects the completion of national and provincial; more than 10 papers published in various journals.



Chenghua Li, born in February, 1972, Wuhan City, Hubei Province, P.R. China

Current position, grades: Associate professor of College of Electronics and Information Engineering, South-Central University for Nationalities.

University studies: Graduated from Huazhong University of Science and Technology in 2009, received a doctor's degree in Cloud Computing.

Scientific interest: Streaming services, data mining and large-scale computing services, etc.

Publications: Presided over 5 scientific research projects the completion of national and provincial; more than 10 papers published in various journals.

Experience: Graduated from Huazhong University of Science and Technology in 2009, received a doctor's degree in Communication Engineering, was approved as a tutor of graduate students in 2011; has completed 5 scientific research projects the completion of national and provincial; more than 10 papers published in various journals.

A novel compressed air solar energy photo-thermal generating electricity system

Junxia Zhang*, Hongxing Zhao

School of Energy and Engineering, Yulin University, China

Received 1 March 2014, www.tsi.lv

Abstract

On the basis of comparing the solar photovoltaic and photo-thermal generating electricity advantages and disadvantages, to overcome the phase-change losses caused by water evaporated into vapour, a compressed air solar energy photo-thermal generating electricity system was proposed in the present work. Air was compressed with compressor and heated with solar heater so as to get high temperature and pressure. High-temperature and high pressure air passes into turbine to generate electricity. The entire design is simple and compact, safe and reliability, energy saving and environmental protection. Thermodynamic cycle analysis was carried out. It comes to conclusion that practical efficiency depends on pressure ratio, the compressor and turbine efficiencies and solar photo-thermal conversion efficiency.

Keywords: Compressed air, solar energy, photo-thermal, generating electricity

1 Introduction

Solar energy is one of the most abundant renewable energy, which can be directly or indirectly used. Solar radiation features strong dispersion and low energy flux density. It is suitable to obtain low temperature thermal resources. At present, it has been widely used for various fields, including heating, solar drying, solar heat, solar air conditioning, solar thermal power generation in area [1].

Solar energy power generation mainly have two categories: photovoltaic and photo-thermal power generation. The three phase alternating current produced with the solar photo-thermal generating electricity can be directly connected to the power grid, and uses regenerative characteristics to achieve continuous power generation, which may be to replace the thermal power and nuclear power in the future. Compared to photovoltaic and wind power, the solar photo-thermal generating electricity belongs to the power source that is more easily accepted by power grid. The cumulative installed capacity of the global solar photo-thermal power generation in 2011 reached 1760 MW, representing an increase of 35% over last year, and new installed capacity of 450 MW [2-6]. Compared to 2010, although the new installed capacity of the solar photo-thermal power generation declined, but several large-scale projects are being under the construction. The global solar thermal power generation in 2012 continues to maintain the momentum of steady growth.

The solar photo-thermal generating electricity refers to the use of large-scale arrays of parabolic or dish-shaped mirror to collect the solar thermal energy to provide steam through a heat exchanger device,

combined with the traditional process of turbine generator, so as to achieve the purpose of power generation. The use of the solar photo-thermal generating electricity technology avoids expensive silicon photoelectric conversion process, and can greatly reduce the cost of the solar power. Moreover, this form of solar energy using solar photo-thermal has great advantages, solar hot water can be stored in huge containers, a few hours after the sun goes down, will still be able to drive a steam turbine power generation, such as solar slotted disc and tower power generation devices. Zhang et al. [7] proposed a technical solution on the solar thermal power-heating. Optical plant has a scale of 10 condenser loops. Turbine generator has the installed capacity of 3 MW, which generates power in the non-heating period and uses the steam turbine cycle cooling water as a heating source in the heating period. It is equipped with the thermal storage system, which meets the need of 24 h heating. Zhang et al. [8] proposed another technical solution to couple the solar thermal power generation system with conventional thermal power plants by means of different ways so as to use solar optical thermal systems in place of some coal consumption, reducing the fired coal consumption. They adopted a trough solar thermal power system, geometric concentration ratio is in the range of 10 to 100, and the heat transfer medium temperature is up to 400°C. A four rank heat exchanger is used to collect heat from the heat collector field, and then heats the water by means of thermal oil carrier. The former two-stage heater heats some feed water to produce the superheated steam, which has the same parameters as the first and second extraction vapour. The superheated vapour enters the first and second extraction point to supply vapour to

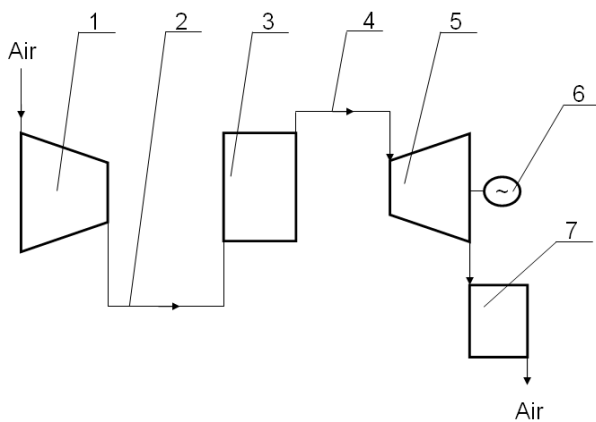
* *Corresponding author* e-mail: wyb700411@163.com

turbine. The latter two-stage heater is an alternative heater of the high-pressure heaters to feed water and low pressure heater to condensate. Based on their computational analysis, they concluded that it has a minimum coal consumption to adopt a complementary manner of the four-stage solar thermal with thermal power heat transfer. In order to overcome the lack of sunlight and night without illumination, the current solar thermal power systems use the thermal storage system, Wang [9] reviewed the prospects of the thermal storage system used for solar thermal power, including molten salt (sodium and potassium), concrete with high temperature thermal storage and phase-change alloy.

However, the existing solar photo-thermal power causes phase change loss because water transforms into vapour. To avoid phase change loss for improving the system efficiency of the solar photo-thermal power generation, this paper proposes a compressed air solar photo-thermal power generation system. The system uses the air as the working fluid, and the compressor and the solar heater were used to raise the temperature and pressure of the air, and then high-temperature and high-pressure air enters into the turbine to generate electricity. The entire design is simple and compact, safe and reliable, energy saving and environmental protection.

2 System structure and work principle

Compressed air solar energy photo-thermal generating electricity system consists of the compressor, the turbine, and the generator, the cooler and the pipes, as shown in Figure 1. The outlet of the compressor is connected to the high-pressure air pipe, the other end of the high-pressure air pipe is connected to the solar heater, the other end of the solar heater is connected to high-temperature air pipe, and the other end of the high-temperature air pipe is connected to the turbine. The rotation shaft of the turbine is connected to the generator, and the outlet of the turbine is connected with the cooler.



1-compressor; 2-high-pressure air pipe; 3-solar heater; 4-high-temperature air pipe; 5-turbine; 6-generator; 7-cooler

FIGURE 1 Schematic diagram of compressed air solar photo-thermal generating electricity

When the air enters into the compressor, the pressure is lifted, and transforms into high-pressure air. High-pressure air flows through into the solar heater, heating by sunlight, transforming into high-temperature air. High-temperature and high-pressure air passes through into the turbine, promotes the expansion work and the impeller and drives generators. The exhaust gas discharged by the turbine is passed into the cooler; the waste heat of the exhaust gas is utilized.

3 Thermodynamic cycle analysis

3.1 IDEAL THERMODYNAMIC CYCLE ANALYSIS

The air is treated as an ideal gas. Ignoring friction and the area loss due to the flow in piping, valves, and the elbow, the ideal thermodynamic cycle is given in Figure 2.

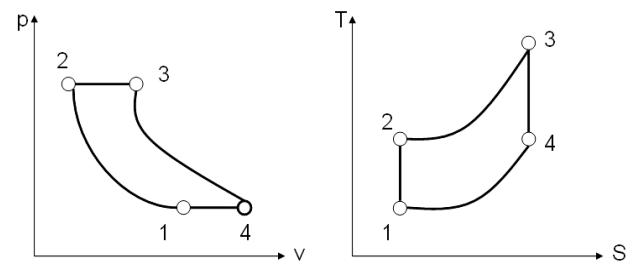


FIGURE 2 Ideal thermodynamics cycle schematics

The 1-2 is a process of the adiabatic compression of air. The compressor adiabatically compresses air from the ambient pressure to P_1 . According to the first law of thermodynamics, the compression work per kilogram air can be represented by the enthalpy difference of the initial and the end states.

$$W_{1-2} = h_2 - h_1 \tag{1}$$

The 2-3 is an air isobaric heating process. Under the solar radiation, the high-pressure air is heated to T_1 . In the process, the done work is zero; the heat absorbed by per kilogram air can be expressed as

$$Q_{2-3} = h_3 - h_2 \tag{2}$$

The 3-4 is a process of the adiabatic expansion of air. The high-pressure air expands in the turbine, and drives the generator to generate electricity. The done work per kilogram air in this process is given by the following formula:

$$W_{3-4} = h_4 - h_3 \tag{3}$$

The 4-1 is a process of the exhaust gas cooling. The exhaust gas discharged from the turbine is cooled to ambient temperature. The utilized waste heat can be expressed as

$$Q_{4-1} = h_4 - h_1 \tag{4}$$

According to the Carnot cycle, ideal cycle efficiency of compressed air solar photo-thermal generating electricity can be expressed as follows:

$$\eta = 1 - \frac{T_0}{T_1} \tag{5}$$

Given ambient temperature $T_0 = 273K$ and heating temperature $T_1 = 273-1000K$, figure 3 shows ideal cycle efficiency diagram. When T_1 ranges at 273-500K, ideal cycle efficiency obviously increases; when T_1 is larger than 500K, ideal cycle efficiency slightly increases. However, expensive exchanger materials and larger heat area need to use because of higher heating temperature, which causes an increase in the investment cost. Therefore, the heating temperature 500K is the optimal selection.

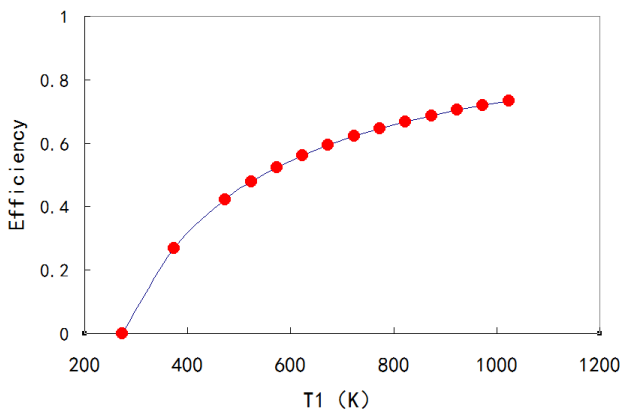


FIGURE 3 Ideal cycle efficiency diagram

3.2 PRACTICAL THERMODYNAMIC CYCLE ANALYSIS

The obtained electricity per kilogram air in unit time for compressed air solar photo-thermal generating electricity can be expressed as follows

$$P_p = Q_{2-3} + W_{1-2} - Q_{4-1} \tag{6}$$

Assumed that the solar photo-thermal conversion efficiency is expressed as η_h , efficiency of compressor is η_c , the ideal electricity per kilogram air in unit time for compressed air solar photo-thermal generating electricity system is

$$P_i = \frac{1}{\eta_T} \left(\frac{Q_{2-3}}{\eta_h} + \frac{W_{1-2}}{\eta_c} \right) - Q_{4-1} \tag{7}$$

Practical cycle efficiency of compressed air solar photo-thermal generating electricity, η_{TP} , is expressed as:

$$\eta_{TP} = \frac{P_p}{P_i} \tag{8}$$

According to thermodynamics, air adiabatic index is k , ambient temperature is T_0 , ambient pressure is P_0 , the net work of the compressor depends on pressure ratio,

$$\varepsilon = \frac{P_1}{P_0}$$

$$W_{1-2} = \frac{k}{k-1} RT_0 \left(\varepsilon^{\frac{k-1}{k}} - 1 \right) \tag{9}$$

Therefore, η_{TP} depends on η_h , η_c , η_T and ε .

Given ambient temperature =273K, the outlet air temperature of the solar heater $T_1 = 500K$, the compressor efficiency $\eta_c = 50\%$, $\eta_T = 80\%$ and pressure ratio $\varepsilon = 3$, Figure 4 shows effect of η_h on η_{TP} . As η_h increases, η_{TP} ascends. Given the solar photo-thermal conversion efficiency $\eta_h = 30\%$, other parameters is the same as the Figure 4.

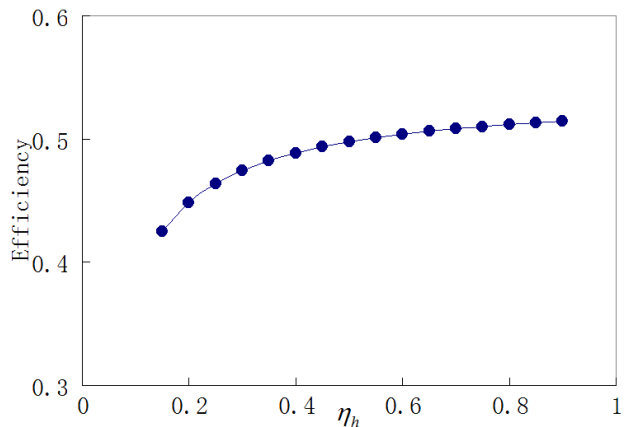


FIGURE 4 Effect of η_h on η_{TP}

Figure 5 shows effect of η_c on η_{TP} . η_{TP} Obviously increases with η_c .

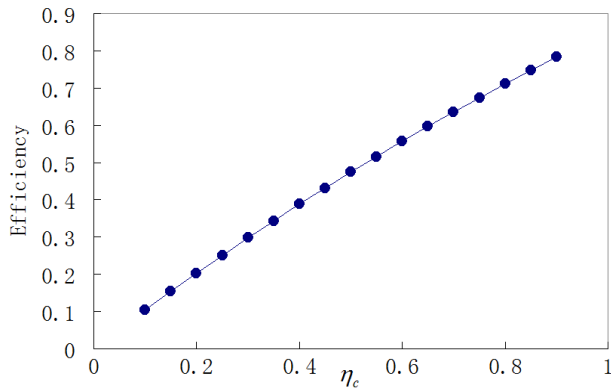


FIGURE 5 Effect of η_c on η_{TP}

Figure 6 shows effect of ε on η_{TP} . When $\varepsilon < 3$, η_{TP} obviously increases; when $\varepsilon > 3$, η_{TP} slightly increases. Therefore, $\varepsilon = 3$ can obtain the optimum η_{TP} .

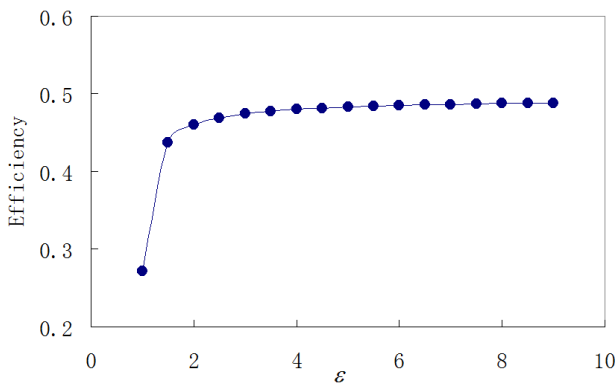


FIGURE 6 Effect of ε on η_{TP}

Figure 7 shows effect of turbine efficiency (η_T) on η_{TP} . η_{TP} Obviously increases with η_c .

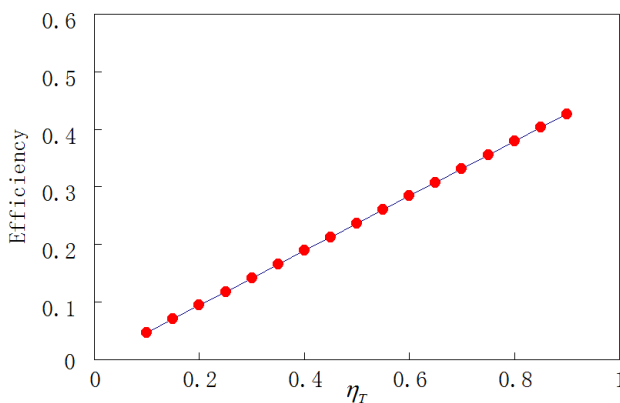


FIGURE 7 Effect of η_T on η_{TP}

3.3 THERMODYNAMIC CYCLE ANALYSIS

As for efficiency, the use period, the investment cost, as well as the environmental impact, the compressed air solar photo-thermal generating system may achieve better performance compared with existing solar photovoltaic and solar photo-thermal power generation system, and is a competitive technology. Solar photovoltaic battery will cause environmental pollution, solar pond and solar photo-thermal power generation system is constrained by the geographical conditions, the working fluid also pollutes the environment. Pollution-free compressed air solar power generation system always uses the air in the power generation process, which will not cause environmental problems. Compressed air solar photo-thermal power generation system takes advantage of existing the compressors, solar heaters and turbines, and not too many restrictions on the geographical conditions.

4 Conclusions

In this paper, the compressed air solar photo-thermal generating electricity was proposed. The system structure, working principle, efficiency and feasibility analysis were discussed. The compressed air solar photo-thermal generating electricity device utilizes existing compressor, solar heaters and turbine technology to produce electricity. This way to generating electricity is an environmentally friendly, without geographical restrictions, easy to implement device.

Thermodynamic knowledge is used to calculate efficiency of the compressed air solar photo-thermal generating electricity, respectively, discussing the effect of compressor and turbine efficiencies, the solar heater photo-thermal conversion efficiency and pressure ratio. Actual cycle efficiency will be increased as the efficiencies of the compressor and turbine, the heat conversion efficiency of the solar heater and the pressure ratio increases, the compressor efficiency has a more significant impact on the efficiency. Optimum air temperature of the solar heater outlet is 500 K, and optimum pressure ratio is 3.0.

5 Notation

- h - the enthalpy;
- k - the adiabatic index;
- P - the pressure;
- Q - the heat;
- R - the gas constant;
- T - the temperature;
- W - the work;
- η - the efficiency;
- ε - the pressure ratio.

References

- [1] Ji J 2013 *Adv. New Renew. Energy*. **1**(1) 8-31
 [2] Shun J P, Li Q H 2012 *Sol. Energy* **20** 54-5
 [3] Yu J, Che J, Zhang J 2008 *World SCI-TECH R&D* **30**(1) 56-9
 [4] Chen X Y, Tian P F 2013 *Coal Technol.* **32**(10) 206
 [5] Fang Z J, Chen G T 2009 *Chinese J. Laser.* **36**(1) 5-14
 [6] Ming T Z, Lin W, Cheng S J 2008 *Adv. Pow. Syst. Hydroelectric Eng.* **24**(6) 1-6
 [7] Zhang G Q, Hu H L, Liu Y Z 2013a *Sol. Energy* **21** 35-9
 [8] Zhang G Q, Hu H L, Liu Y Z 2013b *Sol. Energy* **17** 24-8
 [9] Wang H 2013 *Power Energy* **3** 399-400

Authors



Junxia Zhang, born on July 7, 1974, Fengzhen, Inner Mongolia, China

Current position, grades: Yulin University, Lecturer

University studies: Inner Mongolia University of Science and Technology, Inner Mongolia, China, 2002; Ph.D. degree in University of Science and Technology Beijing, Beijing, China, 2007

Scientific interest: Solar energy, energy conversion and application

Publications: More than 20 papers; 13 papers were in indexed list of SCI, EI and ISTP; 10 China patents were granted.

Experience: Zhang worked in companies for five years and got an award for the design of cryogenic liquid tank trailer. She has been worked in college for four years, taught twelve professional and non-professional courses, and guided undergraduate graduation thesis of more than 10 students, in which three students got the outstanding paper award. Moreover, she hosted a teaching reform project and involved in two teaching reform project. She hosted the project of Yulin College high-level talent, and participated in national, provincial and municipal research projects.



Hongxing Zhao, born on May 27, 1961, Yulin, Shaanxi, China

Current position, grades: Yulin University, Professor

University studies: Northwestern Polytechnical University, Shaanxi Province, China, 1981; Northwest University, Shaanxi Province, China, 1987; Ph.D. degree in Northwestern Polytechnical University, Shaanxi Province, China, 2002

Scientific interest: Mathematical and Computer application

Publications: More than 30 papers

Experience: Zhao worked in Yulin University for many years. He teaches many courses, including higher mathematics, linear algebra, probability and statistics, algebra and computer applications. He won the title of outstanding teachers for several times. Many research findings published in the indexed list of the SCI, EI and ISTP academic journals. Some research findings were considered as the leading domestic level. Meanwhile, he presided over and participated in a number of national, provincial, and municipal research projects.

An experimental study on coal gas desorption laws with different particle size

Z H Wen*, J P Wei, H T Zhang, S H Dai

School of Safety Science & Engineering, Henan Polytechnic University, Jiaozuo, China;

State Key Laboratory Cultivation Base for Gas Geology and Gas Control, Jiaozuo, 454150, China

Received 27 March 2014, www.tsi.lv

Abstract

Based on the self-designed coal containing methane gas desorption law experimental system and coal gas desorption kinetics law, the gas desorption law for different particle size coal under isothermal-isobaric condition are measured by combining combined simulation measurement with the theoretical analysis. The effect of particle size on coal gas desorption laws are obtained through fitting analysis on experimental data: 1. Under the same sorption and desorption conditions, coal particle size was inversely related to total desorption gas amount in the same period. 2. For coal with any particle size, the gas desorption amount firstly increased as time, and finally reached the maximum value, which was equal to the gas sorption amount unit per mass. 3. For coal with higher gas and coal outburst risk trend, the effect of particle size on initial desorption velocity and total desorption amount with time was less. 4. Within the limit particle size, the gas initial desorption velocity had a negative relationship with particle size. Finally we theoretically analysed the effect of particle size on coal gas desorption laws.

Keywords: Particle size, Gas desorption, Gas adsorption, Pore

1 Introduction

Gas desorption law in coal can reflect the gas and coal outburst risk and be applied to determine the loss during the gas content measurement. Many empirical formulas based on the relationship between desorption amount and time has been presented by worldwide researchers through many experiments. These formulas can be classified into two types, one was power functions, represented by R.M. Barrer formula [1]. The other was exponential functions, with E.M. Eirey formula as the representative one.

However, to some degree, these results have their limitation and range ability because they were obtained based on some certain conditions and coal samples. In fact, there are many factors which can affect gas desorption law, for example, gas absorption ability of coal, gas absorption pressure, coal failure type, particle size, moisture in coal sample, temperature and so on [5]. In particle size study, Yang Qiluan, Cao Yaolin, Wang Zhao Feng [1-6] pointed out that there existed a limit particle size. When sample particle size was less than limit particle size, gas desorption intensity and attenuation quotient would decrease with sample particle size increasing. While when the sample particle size is bigger than limit particle size, attenuate trend became little. However, predecessors research results are widely accepted and applied to engineering nowadays, the validity and systematises of these results also need to be verified and adjusted through experiments and theories.

At the same time, reports about mechanism of particle size affecting gas desorption are rare. Therefore, it is necessary to carry out gas desorption law study of different particle size under isothermal-isobaric condition and analyse the mechanism.

2 Gas desorption kinetics law

In gas desorption kinetics law study, researchers presented many empirical formulas, diffusion models, permeability models. For less uncertainty constants in these formulas or models, they are taken as predictive index for coal and gas outburst. However, they have the same deficiency, when $t=0$, the testing data fitting results are bad. In some cases, when $t < 1h$, the fitting results agree with testing data. While in some cases, an available result can be obtained when $t \rightarrow \infty$. So far, the only verified model is three-constant diffusion control model presented by Chen Chang Guo [7].

In this model, gas desorption includes two parts, one is the desorption amount of coal seam/particle and open macro pore surface, $Q_0(t)$. For the surface is directly contacted with surrounding environment, gas can release freely. The other part is the desorption amount of coal seam/particle of internal pore surface, $Q_d(t)$. A diffusion process is necessary for gas to exchange with surrounding environment. So the total gas desorption amount $Q(t)$ is equal to $Q_0(t)$ surplus $Q_d(t)$, i.e. $Q(t) = Q_0(t) + Q_d(t)$.

The gas desorption in the first part belongs to a physical process, which can finish instantaneously. So

* *Corresponding author* e-mail: wenzhihui@hpu.edu.cn

$Q_0(t)$ is a constant Q_0 . In the second part, gas diffusion in coal/particle is the main factors affecting coal desorption/sorption velocity. In three-constant diffusion control model, gas desorption kinetics formula can be described as following:

$$Q(t) = Q_0 + Q_d(\infty) \sqrt{1 - \exp\left(\frac{4\pi^2 D}{d^2}\right)}, \quad (1)$$

where, $Q(t)$ is gas desorption at t time, Q_0 is a constant reference to gas surface desorption, $Q_d(\infty)$ is gas desorption when time tend to infinity, D is diffusion coefficient, d is particle equivalent diameter.

3 Different particle size coal samples gas desorption experiment

3.1 EXPERIMENT SYSTEM

A gas-containing desorption experiment system with desorption apparatus and volume adjusting apparatus were developed. The system was established based on the measurement method of coal methane adsorption (MT/T752-1997). The whole system includes vacuum degassing element, gas adsorption-desorption element, desorption apparatus element and high-pressure gas element.

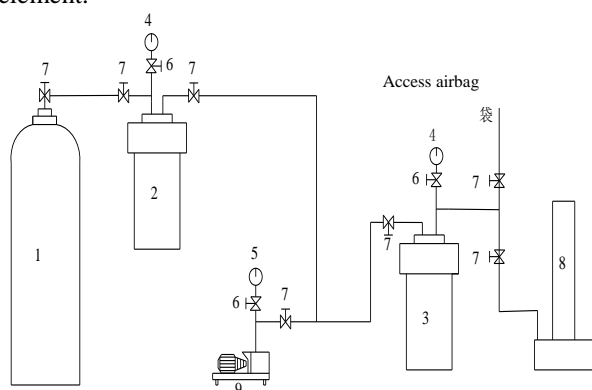


FIGURE 1 Schematic drawing of experiment system

1. high-pressure gas bottle; 2. gasing tank; 3. coal sample tank;
4. precision pressure gauge; 5. vacuum gauge; 6. pressure gauge switch;
7. high-pressure disconnecting valve; 8. desorption apparatus;
9. vacuum pump

3.2 EXPERIMENT PRINCIPLE

Making constant mass dry vacuum degassed coal sample in coal sample tank is connected with high-pressure gas bottle under a selected gas sorption balance pressure. When it reaches to the balance point, testing the gas desorption process of coal sample disposed in air and releasing the pressure at a constant desorption temperature.

3.3 EXPERIMENT METHOD

Two kinds of coal samples were collected from Yian coal mine of Xin'an coalfield and Zhaogu No.1 coal mine of Jiaozuo coalfield respectively. The former were soft coal seam, which had been heavily destroyed with high gas and coal outburst risk. The latter was protogenesis structure coal without gas and coal outburst risk. These coal samples were powdered and classified into small particle sizes: 0.2~0.5 mm, 0.5~1 mm, 1~3 mm, respectively, and then enclosed into six ground glass containers labelled with different label. The testing procedure includes five steps: coal sample pre-treatment, vacuum degassing, gas adsorption equilibrium, determination of the coal sample desorption process, and determination on data processing.

1) Coal sample pre-treatment. Prepared coal is dried with temperature (105 ± 1) °C for about 1.5 h, and keep these coal samples in air-dried state; then we take amount of coal into coal tank. Then the tank is fully compacted as much as possible to reduce the dead space volume in the tank. Cover the sample with cotton and 80 mesh copper net, and seal coal tank.

2) Coal vacuum degassing. Open water bath and vacuum pump. Set the water bath temperature be (60 ± 1) °C. Open the coal tank valve, vacuum degas the coal until the vacuum table to 2 h-0.1 MPa.

3) Gas adsorption equilibrium. After degassing, adjust the water bath temperature be (30 ± 1) °C; unscrew the high pressure gas cylinder valve and inflatable tank valve, and connect high pressure gas cylinders with the inflatable tank. Buffer the pressure of the coal gas into tank. When the pressure of the Inflatable tank comes to 1.2 times of the ideal equilibrium pressure, we close the high pressure gas cylinder valve, and wait to inflate the samples tank, then open the valve connected the samples tank and the inflatable tank, the inflatable tank inflate the samples tank. Depending on the size of the coal, samples tank will get equilibrium after adsorb gas for more than 12 h.

4) Determination of the coal sample desorption process. Firstly read stopwatch, air bags, desorption apparatus, measure, and record the temperature and air pressure. then connect air bags with desorption apparatus, and close the valve between the coal sample tank and inflated tank; open the valve connecting vacuum airbag, so that the free gas in the coal sample tank can go into the vacuum airbag. When the pressure of the tank go to 0, open the valve connected the vacuum airbag, and the valve connected the desorption apparatus, and press the stopwatch and record the start time, read and record the amount of desorption gas.

5) Determination on data processing. For comparative analysis of different coal gas desorption, gas desorption volume will be measured under standard conditions, conversion formula is as follows:

$$Q_t = \frac{273.2}{101325(273.2 + t_w)} (P_{atm} - 9.81h_w - P_s) Q_t^*, \quad (2)$$

where Q_t^* - Standard state of the total gas desorption, cm^3 ;
 Q_t - Experimental conditions the total measured gas desorption, cm^3 ;
 T_w - The amount of water pipe, $^{\circ}C$;
 P_{atm} - Atmospheric pressure, Pa;
 h_w - Read the amount of data within the water column tube, mm;
 P_s - saturated water vapour pressure, Pa.

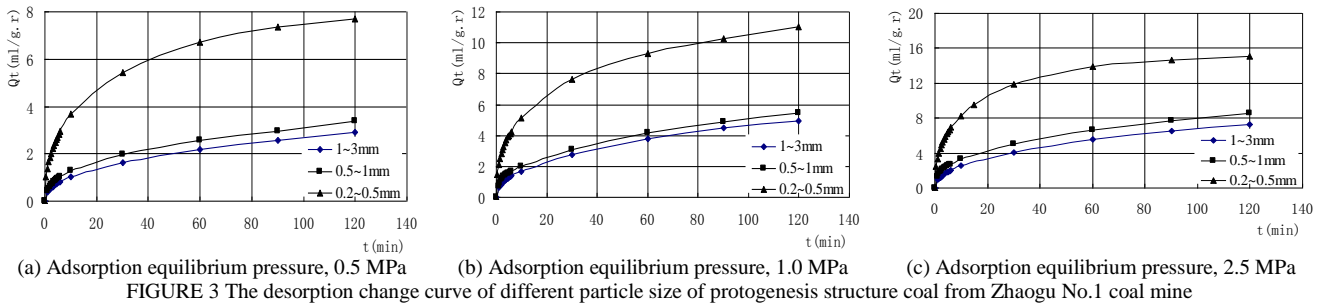
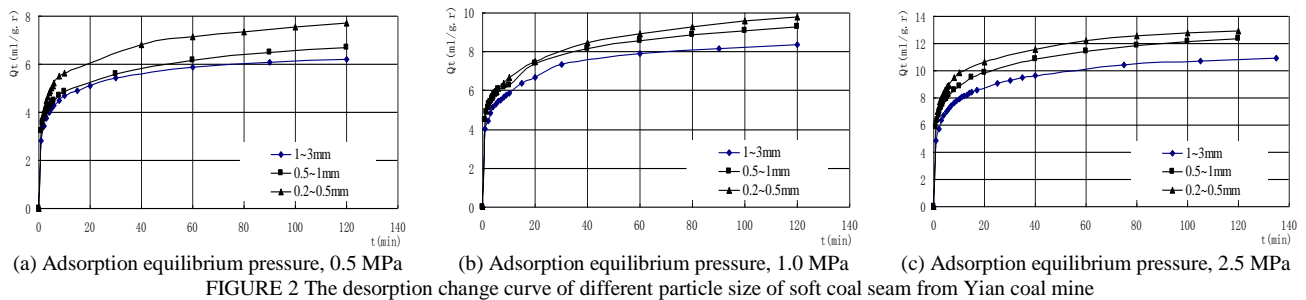
During the test, as desorption environment temperature was maintained at $30 \pm 1^{\circ}C$, gas outlet pressure is approximate 0.1 MPa (ignore the water column influence in the desorption apparatus). Therefore,

we took this test as an isothermal-isobaric desorption procedure condition of all these particle coal samples.

4 Results and analysis

4.1 PARTICLE SIZE EFFECTS ON GAS DESORPTION LAW

The particle size of coal sample was less than limit particle size, which was accordance with the sampling in the field. In order to study effects of particle size on gas desorption law, the above 6 samples were tested under the same sorption and desorption conditions. The fitting results were showed as following figures.



From above Fig. 2 and Fig. 3, we learnt that when the particle size was less than limit particle size, the heavily destroyed coal was analogous to the effects of protogenesis structure coal on gas desorption process. Under the same gas pressure, gas content was not affected by particle size. When the particle size was big, coal desorption velocity was low. When the particle size was small, velocity was fast. The reason was that the gas flow resistance increases with particle size increasing, thus the gas desorption velocity decreases. So particle size can affect the gas desorption of gas-containing coal, this effect was mainly reflected in the following aspects:

1) Under the same sampling place and absorption and desorption conditions, the desorption total amount had a negative relationship with particle size in the same period.

2) Particle size would not affect the total desorption amount, and the total desorption amount change curve was always a monotone increasing with top limit, which was the gas sorption amount unit per mass.

3) We learnt that the higher gas and coal outburst risk trend is, the less effect of particle size on initial desorption velocity and total desorption amount with time changing will be.

4) The displacements of desorption gas and hindrance force acting on the gas were influenced by coal particle size. Smaller particle size leads to shorter gas displacement and higher gas desorption rate [9]. The initial rate of gas desorption decreases with an increase in particle size. Extreme particle size was proposed by Yang [10]. Extreme particle size is related to physical-chemical properties of coal, but is irrelevant to damage extent. When coal particle size is larger than extreme particle size, the coal is divided into lots of fine particles by the cracks and large pores. It is regarded that the hindrance force of pores within the coal having extreme particle size are constant. Moreover, the hindrance of the cracks and large pores is far lower than that of all of the coals with extreme particle size. Hence, when the coal size was up to a certain value (extreme particle size), the initial rate of gas desorption did not decrease anymore.

5) The gas initial desorption velocity decreased with particle size increasing. According to the limit particle size theory, coal sample initial desorption velocity would reach to be constant, which was verified by the testing data with particle size 0.5~1 mm, 1~3 mm. Furthermore, it also showed the validity of this conclusion.

4.2 THEORY ANALYSIS ON THE PARTICLE SIZE EFFECTS ON THE GAS DESORPTION LAW

According to the gas desorption kinetics law, coal seam or particle desorption is affected by coal seam/particle surface, open macro pore surface and coal internal pore space. The particle size firstly affected the coal total surface area, secondly the pore space gas molecule in or out the coal particle.

Total specific surface area: Coal is a complex porous medium and nature sorbent. The surface area proportion with microspore diameter less than 10⁻⁴ mm, up to 97.3%, specific area to 200 m²/g determined the adsorption volume to some degree. Through analysing the coal sample with particle size (0.5~1, 5~10 mm) pore structure testing data, Zhang Xiaodong [8] pointed out that particle size had little influence on the pore specific area, especially for microspore. In addition, particle size had little effect on the dry coal desorption amount. So, from the specific area point, particle size hardly affected the gas desorption.

Particle coal internal pore space: The particle size can reflect the gas desorption transportation distance and resistance. For the same coal sample, the bigger the particle size, the longer desorption transportation distance and the bigger the resistance. So in the per unit time, the desorption amount was small. At the same time, during the sorption process, the coal sample body would swell, leading the coal microstructure to irreversible deformation. Under the same condition, for the coal sample with small particle size, this deformation process might enhance the micro connectivity and enlarge the surface area, causing the adsorption amount increasing. Of course, the gas desorption intensity in per unit time and given time gas desorption would increase meanwhile.

5 Conclusions

The gas desorption of coal is influenced by many factors. These factors can be categorized into two types. One is

References

- [1] Barrer R M 1951 Diffusion in and through solid, Ph.D. Thesis *Cambridge Univ.* England
- [2] Airey E M 1968 Gas emission from broken coal *Int. J. Rock Mech. Min.* **5** 2074-83
- [3] Wang Y A 1996 Gas control in coal mine, in *Coal Mine Safety Manual: Part 2* China Coal Industry Publishing House: Beijing
- [4] Yang Q L 1986 Study on coal gas diffusion law, *Safety in Coal Mines* **18** 9-17
- [5] Cao Y L, Qiu H S 2007 Gas desorption law of classic coal core study *China Min. Mag.* **16** 119-223
- [6] Wang Z F 2001 Gas desorption law in air, water and mud medium study and application, Ph.D. Thesis *China Univ. Min. Technol.* Beijing, China

the physical-chemical properties of deformed coal, such as metamorphic grade, adsorption capacity, damage type, internal moisture, and particle size, etc. The other is external factors, which include gas adsorption equilibrium pressure and environmental temperature, etc. [11].

Through studying, the effects of particle size on gas desorption law, we can obtain the following conclusions:

1) Under the same absorption and desorption conditions, coal particle size is inversely related to total desorption gas amount in the same period.

2) For coal with any particle size, the gas desorption amount firstly increased with time, and finally reached the maximum value, which was equal to the gas sorption amount unit per mass.

3) For coal with higher gas and coal outburst risk trend, particle size on initial desorption velocity and total desorption amount show less effect with time.





4) Within the limit particle size, the gas initial desorption velocity had a negative relationship with particle size.

The sample particle size can affect the total specific surface area. However, the results showed that it had little influences on the gas absorption or desorption. Particle size can affect the pore space that gas molecular transport, which can directly influence the desorption pathway and resistance. Meanwhile, the swelling deformation induced by gas sorption can affect gas desorption of different particle size. This result can explain the reason why gas desorption law of different particle size were obviously different under the same conditions.

Acknowledgements

The research presented in this paper was funded by the National Natural Science Foundations of China (No. 51104059, 51004041), the National Key Basic Research and Development Program of China (973 Program) (No. 2012CB723103), the Chinese Ministry of Education Innovation Team Development Plan (No. IRT1235) and the Open Project of State Key Laboratory Cultivation Base for Gas Geology and Gas control (No. WS2013B04, WS2012B07).

- [7] Chen C G 1995 Study of coal physical-chemistry structure and gas adsorption/desorption mechanism, PhD Thesis *ChongQing University* Chongqing, China
- [8] Zhang X D, Sang S J 2005 Isotherm adsorption of coal samples with different grain size *J. China Univ. Min. Tech.* **34** 427-32
- [9] Li X Y 2003 Characteristics and effecting factors of adsorption time of coalbed gas reservoir *Nat. Gas Geosci.* **6** 502-5
- [10] Yang Q L 1987 Experimental study of gas desorption law of coal *Safety Coal Mines* **19** 9-17
- [11] Wen Z H 2008 Experimental study on gas desorption laws of tectonically coal, Master degree dissertation *Henan Polytech. Univ.* Jiaozuo, China

Authors	
	<p>Zhihui Wen, born on March 24, 1982, Xingtai city, Hebei province of P.R.C</p> <p>Current position, grades: member of State Key Laboratory Cultivation Base for Gas Geology and Gas Control, Lecturer University studies: Henan Polytechnic University Scientific interest: coal mine gas disasters control, numerical simulation Publications: have published more than 20 papers, including more than 5 papers indexed by EI or SCI.</p>
	<p>Jianping Wei, born on May 25, 1971, Suiping county, Henan province of P.R.C</p> <p>Current position, grades: Vice-minister of State Key Laboratory Cultivation Base for Gas Geology and Gas Control, Professor University studies: China University of Mining and Technology Scientific interest: coal mine gas disasters control Publications: have published more than 60 papers, including more than 20 papers indexed by EI or SCI.</p>
	<p>Hongtu Zhang, born in 1988, Suiping county, Henan province of P.R.C</p> <p>Current position, grades: student of State Key Laboratory Cultivation Base for Gas Geology and Gas Control, PhD student University studies: Henan Polytechnic University Scientific interest: coal mine gas disasters control Publications: have published more than 5 papers, including more than 2 papers indexed by EI or SCI.</p>
	<p>Shaohua Dai, born in 1990, Nanyang city, Henan province of P.R.C</p> <p>Current position, grades: student of State Key Laboratory Cultivation Base for Gas Geology and Gas Control, Master student University studies: Henan Polytechnic University Scientific interest: coal mine gas disasters control Publications: have published more than 2 papers.</p>

Application of fractal theory to dam deformation forecast

Zhigang Yin*, Guohui Gao

Changchun Institute of Technology, Changchun, China

Jilin Province Water Project Security and Disaster Prevention and Control Project Laboratory, Changchun, China

Received 1 March 2014, www.tsi.lv

Abstract

Based on the safety observation data of dam, the establishment of the prediction model of dam deformation is very important for safe operation of the reservoir. The early deformation curve of the dam has self-similar fractal feature. The fractal interpolation function can be applied to not only processing the dam prototype observation data but also forecasting the rule of early dam deformation. In this paper, the reservoir dam deformation can be analysed and predicted by the fractal interpolation function. Analysis shows that, the method for dam deformation prediction of maximum error is 8.0%. Therefore, regarding the half-year short-term forecast, the forecast result obtained from fractal interpolation function method could be reliable.

Keywords: dam, the fractal interpolation function, forecast

1 Introduction

As a water retaining structure, dam can bring social and economic benefits, but its potential safety problem also brings great threat to people's lives and properties. Although china's government has paid increasing attention on the dam safety, the security situation of the dam is still not optimistic [1-2]. Since China's vast areas are densely populated, dam-breaking in the remote region could still cause tremendous casualties and property losses. Therefore, establishing dam deformation and stress change forecast models has great significance in forecasting the dam security for some time to come, timely warning and taking effective remedial measures before the accident happens so as to minimize the disaster loss.

At present, the forecast model established by analysing the dam observation data mainly includes statistic model, grey theory prediction model, neural network method and so on. These forecast models demand to establish corresponding multi-parameter mathematical equations based on the dam operation test data. During application, the determination of the parameters could have an important effect on the forecast result. The fractal interpolation function put forward by Barnsley in 1986, the numerical value of a time-varying random variable could be forecasted as long as a parameter (scale factor) is determined. The method is especially suitable for forecasting dam deformation. The fractal theory has outstanding advantages in revealing all kinds of universal rules in dam deformation and other complicated phenomenon, so that it provides a new method for processing the dam prototype observation data and predicting dam deformation.

2 Method and forecast model

2.1 FRACTAL THEORY

For the given closed interval $I = [a, b]$, supposed that $a = x_0 < x_1 < \dots < x_N = b$ is a partition of I , wherein N is more than or equal to 2, and y_0, y_1, \dots, y_N is randomly a group of real numbers, $K = I \times R$ [3], $I_i = [x_{i-1}, x_i]$, $i = 1, 2, \dots, N$, supposed that L_i is a compression homeomorphism of $I \rightarrow I_i$, L_i as

$$L_i(x_0) = x_{i-1}, L_i(x_N) = x_i. \quad (1)$$

If $0 < l_i < 1$, and then

$$|L_i(u_1) - L_i(u_2)| \leq l_i |u_1 - u_2|, \forall u_1, u_2 \in I. \quad (2)$$

Supposed that F_i is the continuous function of $K \rightarrow R$, F_i as

$$F_i(x_0, y_0) = y_{i-1}, F_i(x_N, y_N) = y_i. \quad (3)$$

If $0 \leq q_i < 1$, and then

$$|F_i(u, v_1) - F_i(u, v_2)| \leq q_i |v_1 - v_2|, \forall u \in I, v_1, v_2 \in R. \quad (4)$$

Define the mapping $\omega_i : K \rightarrow K$:

$$\omega_i \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} L_i(x) \\ F_i(x, y) \end{pmatrix}, i = 1, 2, \dots, N \quad (5)$$

* *Corresponding author* e-mail: yzhg1972@163.com

and then $\{K; \omega_i, i=1,2,\dots,N\}$ constitute an iterated function system.

Regarding the continuous function f in I (given closed interval), G is the image of f , $G = \text{Graph}(f) = \{(x, f(x)), x \in I\}$ is the Invariant set of the iterated function system $\{K; \omega_i, i=1,2,\dots,N\}$, that is to say, $G = \bigcup_{i=1}^N \omega_i(G)$ and, $f(x_i) = y_i, i=1,2,\dots,N$, it is considered that such f is the fractal difference function corresponding to $\{K; \omega_i, i=1,2,\dots,N\}$ (FIF) [3].

Supposed that the data set $(x_i, y_i): i=1,2,\dots,N$ is given, the process for constructing the iterated function system is shown as below. When $L_i(x)$ and $F_i(x, y)$ are linear functions, formula (5) could be expressed as below:

$$\omega_i \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_i & 0 \\ c_i & d_i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e_i \\ f_i \end{pmatrix}, i=1,2,\dots,N \tag{6}$$

Then $L_i(x) = a_i x + e_i$, $F_i(x, y) = c_i x + d_i y + f_i$. According to formulae (2) and (4), $0 < |a_i| < 1$, $|d_i| < 1$.

The following equation set could be derived from formulae (1) and (3).

$$\begin{cases} a_i x_0 + e_i = x_{i-1} \\ a_i x_N + e_i = x_i \\ c_i x_0 + d_i y_0 + f_i = y_{i-1} \\ c_i x_N + d_i y_N + f_i = y_i \end{cases} \tag{7}$$

2.2 FRACTAL PREDICTION METHOD

The dam deformation curve has the self-similarity. The settlement of the observing point could be a random process. According to the principle of fractal interpolation function, supposed that the settlement is a continuous self-similar random process $Y = \{Y(t), t \in T\}$, namely $Y(t) = \mu^{-H} y(\mu t)$, $t \in T$, $\forall \mu > 0$, $0 \leq H < 1$ H is Hurst coefficient, which has great significance in indicating the degree of self-similarity. If the Hurst coefficient is larger, closer to 1, the self-similarity (fractal characteristic) of the system would be more distinct.

Supposed that the accumulated settlement value of the dam observation point is $\{(t_i, y_i), i=0,1,\dots,N\}$, $t_i \in [t_0, t_N]$, and $y_i \in [a, b]$, the following equations could be derived according to the fractal difference formula: $L_i(t) = t_{i-1} + (t_i - t_{i-1})(t - t_0) / (t_N - t_0)$, $F_i(t, y) = c_i t + H y + f_i$, $\omega_i(t, y) = (L_i(t), F_i(t, y))$, $i=1,2,\dots,N$. An iterated function system could be

constructed in this way. The attractor is the image of the desired fractal interpolation function $f(x)$.

Afterwards, the fractal interpolation function is extrapolated so as to forecast the dam deformation based on the existing dam observation data. The data could be extrapolated via the following method by using the interpolation function: the fractal interpolation function within the interval $[t_0, t_N]$ could be obtained from the fractal interpolation formula and then it is extrapolated. The iterated function within the interval $[t_N, t_{N+1}]$ could be defined, namely,

$$\begin{aligned} L_{N+1}(t) &= t_N + (t_{N+1} - t_N)(t - t_0) / (t_{N+1} - t_0), \\ F_{N+1}(t, y) &= c_{N+1} + H y + f_{N+1}, \quad \text{wherein,} \\ c_{N+1} &= (y_{N+1} - y_N - H(y_{N+1} - y_0)) / (t_{N+1} - t_0), \\ f_{N+1} &= y_N - H y_0 - c_{N+1} t_0 \tag{4}. \end{aligned}$$

3 Application of fractal prediction method

TABLE 1 Observation data about deformation at the top of the earth dam of a reservoir

Time interval (month)	Settlement value within the time interval (mm)	Accumulated settlement value (mm)
1	0	0
2	23	-23
3	15	-38
4	30	-68
5	32	-100
6	18	-118
7	1	-119
8	-5	-114
9	3.6	-117.6
10	2.1	-119.7
11	0.3	-120
12	8.3	-128.3
13	4.4	-132.7
14	2.7	-135.4
15	-2.4	-133

The observation curve obtained by time series method is shown in Fig. 1.

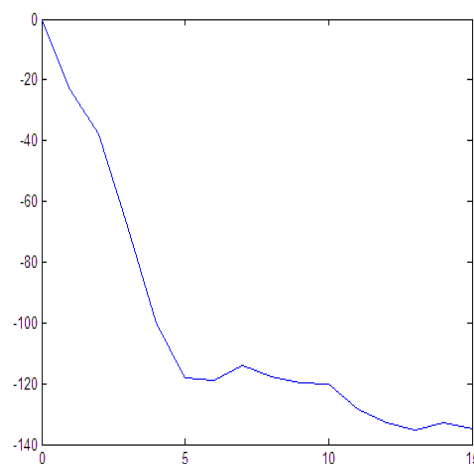


FIGURE 1 Dam top deformation curve (time series method)

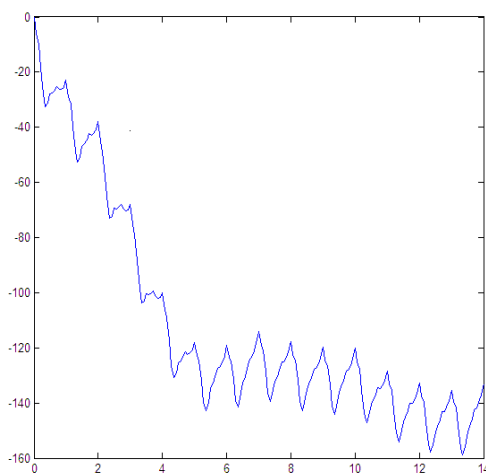


FIGURE 2 Dam top deformation curve (fractal geometry method)

Fig. 2 is the dam top deformation curve fitting by the interpolation function. Due to the limitation of the sampling frequency and the data storage capacity, the dam deformation observation data available is limited and intermittent discrete data. The fitting of original continuous data is realized via mathematic method. Obviously, the deformation curve obtained via fractal method better fits the unstable dam deformation in early period. The fractal method is mainly adopted for obtaining the dam deformation condition in the future according to the observation data.

References

[1] Xing L S 2001 Analysis on China's hydropower station dam accident and security countermeasures *Adv. Sci. Technol. Water Resour.* **21** 26-32 (In Chinese)

TABLE 2 Comparison between predicted value and measured value

Time period (month)	Measured value (mm)	Predicted value	Relative error
16	-134.7	-137.7	2.20%
17	-134.1	-140.6	4.80%
18	-134.6	-143.4	6.50%
19	-135.2	-146.1	8.00%

From Table 2, it can be observed that the dam deformation is predicted by using fractal interpolation function; the error of the forecasting result will become larger along with the increase of the time period. This is related with the selected scale factor d_i , whose optimization needs further study. The maximal error of the method used in this project is 8.0%. Therefore, regarding the half-year short-term forecast, the forecast result obtained from fractal interpolation function method could be reliable.

4 Conclusion

Considering that the early deformation curve of the dam has self-similar fractal feature, the fractal interpolation function can be applied to not only processing the dam prototype observation data but also forecasting the rule of early dam deformation, with the forecast precision well consistent with the measured value. Therefore, the fractal interpolation function is particularly suitable for forecasting the dam deformation.

[2] Wei D R 2000 Progress and development prospect of hydropower station dam security monitoring during the ninth five-year plan period *Dam Monit. Technol.* **3** 33-38 (In Chinese)
 [3] Sha Z, Ruan H J 2005 *Fractal and Fitting*, Zhejiang Univ. Press: Zhejiang, China (In Chinese)
 [4] Wang Y D, Hua X H 2008 Deformation forecast method based on fractal theory *Geospatial Inform.* **6** 114-116 (In Chinese)

Authors	
	<p>Zhigang Yin, born on July 1, 1972, Changchun city, Jilin Province, China</p> <p>Current position, grades: Changchun Institute of Technology, Assistant Professor University studies: Changchun Institute of Technology Scientific interest: Mathematics, Water Project Security and Disaster Prevention and Control Publications: The Research of the Spectral Features of Vibration Signal From Underground Railway Based on Wavelet Transform Experience: 2009.01- Changchun Institute of Technology, Assistant professor 2005.03-2009.01 Ph.D. degree in civil engineering, Tongji University, Shanghai, China, Changchun Institute of Technology, 1991.08-1995.07 Fuzhou University (Faculty of Civil Engineering), 1991</p>
	<p>Guohui Gao, born on July 17, 1989, Zhengzhou city, Henan Province, China</p> <p>Current position, grades: Changchun Institute of Technology, student University studies: Changchun Institute of Technology Scientific interest: The theory and application of hydraulic structures Experience: 2008.09-2012.06, Changchun Institute of Technology, 2012.09, Changchun Institute of Technology.</p>

Three dimensional simulation of weft knitted fabric based on surface model

Yinglin Li^{1,2}, Lianhe Yang^{3*}, Suying Chen⁴, Lei Xu^{1,2}

¹ School of Textiles, Tianjin Polytechnic University, Tianjin 300387, China

² Key Laboratory of Advanced Textile Composites, Ministry of Education, Tianjin 300387, China

³ School of Computer Science & Software Engineering, Tianjin Polytechnic University, Tianjin 300387, China

⁴ School of Textile and Clothing, Qingdao University, Qingdao 266071, China

Received 1 March 2014, www.tsi.lv

Abstract

In order to enhance the visual simulation effect of weft knitted fabric in weft knitting CAD system, a three dimensional surface model of weft knitted structures is developed and realized on the computer screen. It is assumed that the yarn centreline is continuous when it is stressed to bend, and the cross-section of yarn is circular and uniform. Three dimensional geometric models are created by using second order continuous NURBS curves. According to the trend of the yarn centreline, a surface model of the yarn is created by sweeping approach. With developing tools of Visual C++ and OpenGL graphics library, the surface model of loop is achieved by the method of mesh chips on computer screen, and the three dimensional simulation of weft knitted fabric is realized. The simulation effect based on surface model can be seen from the simulation results of basic weft knitted fabric.

Keywords: Weft knitting, Loop structure, NURBS; Simulation, Surface model

1 Introduction

By using computer aided design (CAD) technology, designers can easily design fabrics in virtual mode [1]. At the same time, the efficiency and quality of fabric design has been greatly improved, and the repeated process of knitting experimenting, proofing can be avoided. CAD systems are convenient tools for knitting fabric manufactures.

Loop modelling is the foundation of knitting CAD technology. Research on weft knitted fabric structure has continued for 70 years since the forties of the last century. Loop modelling is the foundation of knitting CAD technology. Research on weft knitted fabric structure has continued for 70 years since the forties of the last century. Many geometrical models have been created by researchers for plain knitted fabric [2], including those of Chamberlain [3], Peirce [4, 5], Leaf and Glaskin [6], Leaf [7], Munden [8, 9], Postle [10], Demiroz [11], Kurbak [12-14], Su Liu [15], Ranju Meng [16] and Yinglin Li [17,18]. The geometric model of loop basically perfects, which is the foundation for the simulation of fabric.

However, there are more problems in the fabric simulation. Kurbak carried out simulation of fabric using 3DS Max [12], which is convenient, but is not conducive to the development of CAD software. Su Liu [15] applied sphere function in OpenGL to achieve simulation of fabric, but this method is not conducive to physical attributes of model.

In this paper, loop geometrical model is created by using the NURBS spline curve [19-22], and the three dimensional simulation of the loop and fabric is realized by using the surface model.

2 The geometric modelling of plain loop

2.1 ANALYSIS OF LOOP STRUCTURE

Knitted structures are progressively built-up from row of intermeshed loops (as shown in Figure 1). The basic constitutional unit of knitted fabrics is the loop. It is a bending space curve. The plain loop is composed of needle loop, sinker loop and loop column [23, 24].

The loop model should be developed before the simulation of the knitted fabric. It is assumed that the yarn in the loop is uniform when the fabric is fully relaxed, the cross-section is circular and the loop is evenly symmetrical, which can be used to better describe the actual shape of the loop and the buckling state of the yarn in the loop, the centreline of the yarn should be a second order continuous spatial curve.

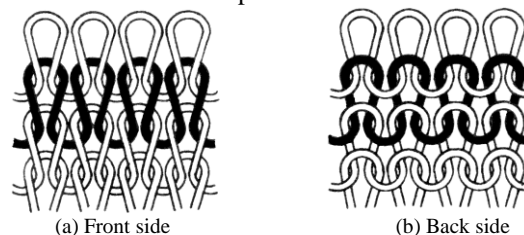
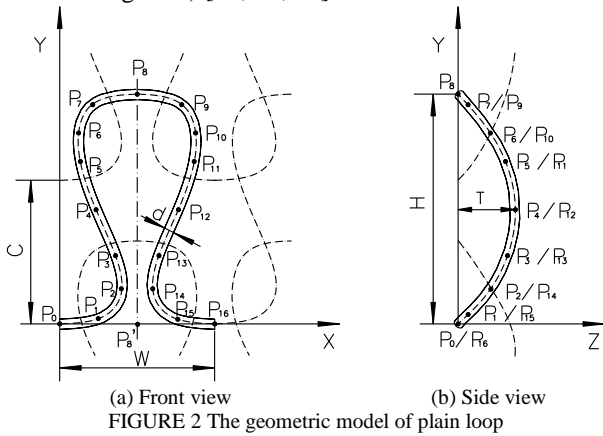


FIGURE 1 Plain knitted fabric

* Corresponding author e-mail: yanglh@tjpu.edu.cn

2.2 THE GEOMETRIC MODEL OF PLAIN LOOP

Plain loop is the basic unit of weft knitted fabric. The basic structure parameters of plain loop model are: the loop width W , the loop height H , the yarn diameter d , the thickness T of the fabric, and the course spacing C (as shown in Figure 2) [17, 23, 24].



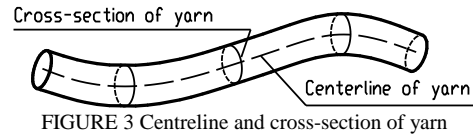
This paper uses the three dimensional loop geometric model which is developed with the data points (that loop passed through, determined according to the geometric characteristics of the loop) of the 17 type based on spline curves in the literature [18], and has made the improvement. The value of C is equal to $0.86H$. The proportional relationship between the X , Y , and Z coordinates of data points and the loop structure parameters, such as width W , height H and thickness T are shown as Table 1.

TABLE 1 The proportional relationship between data points and the loop structure parameters

Data Points No.	X Coordinate	Y Coordinate	Z Coordinate
P_0	0	0	0
P_1	$0.2500W$	$0.0407H$	$0.1275T$
P_2	$0.3493W$	$0.2035H$	$0.5966T$
P_3	$0.3145W$	$0.3494H$	$0.8901T$
P_4	$0.2398W$	$0.4954H$	$0.9999T$
P_5	$0.1751W$	$0.6486 H$	$0.8930T$
P_6	$0.1488W$	$0.7874H$	$0.6193T$
P_7	$0.2190W$	$0.9230 H$	$0.2396T$
P_8	$0.5000W$	H	0
P_9	$0.7810W$	$0.9230H$	$0.2396T$
P_{10}	$0.8512W$	$0.7874H$	$0.6193T$
P_{11}	$0.8249W$	$0.6486H$	$0.8930T$
P_{12}	$0.7602W$	$0.4954H$	$0.9999T$
P_{13}	$0.6855W$	$0.3494H$	$0.8901T$
P_{14}	$0.6507W$	$0.2035H$	$0.5966T$
P_{15}	$0.7500W$	$0.0407H$	$0.1275T$
P_{16}	W	0	0

3 Three dimensional surface model of yarn

The basic idea of generating yarn model is to combine the yarn cross-section and the centreline into three dimensional yarn, using sweeping method namely to construct the tube surface along the centreline to indicate yarn [25, 26], as shown in Figure 3.



3.1 NURBS CURVES

Because NURBS (Non-Uniform Rational B-Splines) curve can be partially adjusted, and has good geometry characteristic as well as high flexibility, especially to meet the requirements of second order continuous [19-21]. It can simulate the trend of the centreline of the yarn well, so NURBS curve may be used to simulate yarn centreline.

3.1.1 NURBS curves description

NURBS is a mathematical method adopted to describe the free curves and surfaces. A NURBS curve can be formulated by a sub-rational polynomial function [26-28].

$$p(u) = \frac{\sum_{i=0}^n w_i d_i N_{i,3}(u)}{\sum_{i=0}^n w_i N_{i,3}(u)}, i = 0, 1, 2, \dots, n' \tag{1}$$

where w_i is the weight factor, d_i is the control point, and $N_{i,k}(u)$ is the k standard B-spline basis functions.

3.1.2 Control points revolved inversely

According to NURBS method, the control points of curve can be revolved inversely by the given data points and then the shape of the curve can be derived by interpolation [29, 30].

3.1.3 NURBS derivative

The derivative of Formula (1) is:

$$P'(u) = \frac{\sum_{i=0}^n w_i d_i N_{i,k}(u)}{\sum_{i=0}^n w_i N_{i,k}(u)} \tag{2}$$

The key point is its basic function derivative, namely [27],

$$N'_{i,0}(u) = 0, \tag{3}$$

$$N'_{i,k}(u) = \frac{1}{u_{i+k} - u_i} N_{i,k-1}(u) + \frac{u - u_i}{u_{i+k} - u_i} N'_{i,k-1}(u) - \frac{1}{u_{i+k+1} - u_{i+1}} N_{i+1,k-1}(u) + \frac{u_{i+k+1} - u}{u_{i+k+1} - u_{i+1}} N'_{i+1,k-1}(u) \tag{4}$$

Formulae (3) and (4) indicate that the derivative of the basic function is also defined by recursive method [28]. Thus, we can calculate the first derivative of NURBS curves. Similarly, it can also be used to calculate the second derivative.

3.2 THE CENTERLINE OF YARN

3.2.1 Definition of the centreline of yarn

We presume that the centreline of yarn is represented by a space curve C whose parametric equation is determined by the formula [25, 28]:

$$\begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) \end{cases} \quad (0 \leq t \leq 1), \tag{5}$$

If it is represented by vector equation, the result is:

$$r = r(t) = x(t)i + y(t)j + z(t)k, \tag{6}$$

Each function of the formulas has the first and the second order continuous derivative, namely, $r(s) \in C^2$.

3.2.2 Tangent vector of the centreline of yarn.

It is assumed that $r(t_0)$ which is short for point t_0 represents any point on the curve C. The tangent vector at point t_0 of the curve C is:

$$\frac{dr}{dt} = r'(t) = x'(t)i + y'(t)j + z'(t)k, \tag{7}$$

where, $x'(t) = \frac{dx}{dt}, y'(t) = \frac{dy}{dt}, z'(t) = \frac{dz}{dt}, \tag{8}$

According to the arc length parameter method, the standard tangent vector T at point of the curve C is:

$$T = \frac{dr}{ds} = \frac{dr}{dt} \cdot \frac{dt}{ds} = \frac{r'(t)}{|r'(t)|} = t_x(t)i + t_y(t)j + t_z(t)k, \tag{9}$$

Hence,

$$t_x(t) = \frac{dx}{dt} / \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}, \tag{10}$$

$$t_y(t) = \frac{dy}{dt} / \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}, \tag{11}$$

$$t_z(t) = \frac{dz}{dt} / \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}, \tag{12}$$

3.2.3 Normal vector of the centreline of yarn

The unit vice normal vector B at point t_0 of curve C can be written as:

$$B = b_x \cdot i + b_y \cdot j + b_z \cdot k, \tag{13}$$

where the B value can be obtained according to formula as follows [26]:

$$B = \frac{r'(t) \times r''(t)}{|r'(t) \times r''(t)|}, \tag{14}$$

The unit main normal vector N at point t_0 of curve C can be written as:

$$N = \frac{dT}{ds} = n_x i + n_y j + n_z k, \tag{15}$$

where the N value can be obtained according to formula as follows [26]:

$$N = B \times T = \frac{r'(t) \times r''(t) \times r'(t)}{|r'(t) \times r''(t)| \cdot |r'(t)|}, \tag{16}$$

3.2.4 The cross-section of yarn

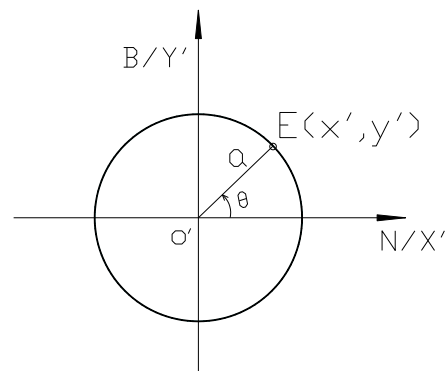


FIGURE 4 Cross-section circle of yarn

In the normal plane that the any point O' of curve C is located, we treat the main normal vector N as X' axis, vice normal vector B as Y' axis, and point O' as the origin point to construct a plane coordinate system. The circle with centre point O' and radius a is the cross-section circle of yarn, which is shown in Figure 4. The parametric equation of any point E on the circle is as follows (θ is a parameter):

$$\begin{cases} x' = a \cos \theta \\ y' = a \sin \theta \end{cases} \quad (0 \leq \theta < 2\pi), \quad (17)$$

According to Formulae (9-12,14,16), in order to get the tangent vector T , the main and vice normal vector N , B of the curve C , the first derivative $r'(t)$ and the second derivative $r''(t)$ are needed to calculate, therefore, the derivative of the centreline of the yarn is equal to the derivative of NURBS curve [27,28].

3.2.5 The surface model of yarn

As the yarn is considered to be formed by the sweeping of cross-section circle along the centreline, and a series of cross-section circulars is distributed on the discrete points of the centreline, we can construct the surface of the yarn on the basis of the adjacent cross-section circles.

According to formula (6), $C:r=r(s)$ is a space curve of arc length parameter s , $\{k(s), \tau(s), T(s), N(s), B(s)\}$ are the Curvature, Torsion and Frenet frame of the curve C respectively [25]. Assuming $K(s) < 1/f$ (f is a positive constant), we create the surface S :

$$\rho(s, \theta) = r(s) + a[N(s)\cos\theta + B(s)\sin\theta]. \quad (18)$$

In the formula, s represents the parameter in the centre line direction, θ represents the parameter of the cross-section circular direction, the surface obtained in this way is usually called tube surface [25]. The tube surface can be used to represent the yarn surface whose position in the world coordinate system OXYZ is shown in Figure 5, where P is a normal plane that any point O' on the centreline of the yarn is located, N is the main normal vector, T is the tangent vector, B is the vice normal vector, OO' is the vector of any point O' on the centre line of yarn, OE is the vector of any point E on the cross-section circle, their vector relationships are as follows:

$$OE = OO' + O'E, \quad (19)$$

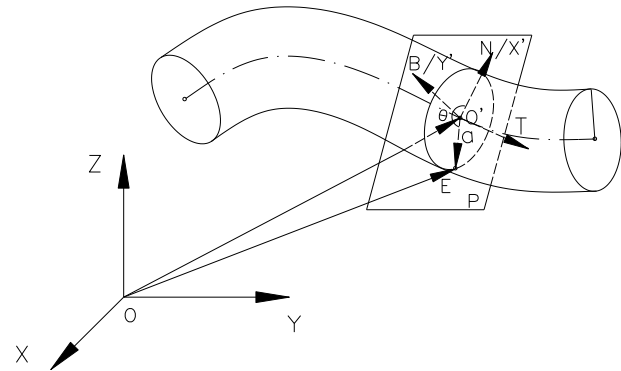


FIGURE 5 Surface model of yarn

In equation (15), $r(s)$ can be obtained according to the geometry model of the centreline of the loop with NURBS function, after its derivative, then according to the equations (11) and (13), $B(s)$ and $N(s)$ can be obtained, thus the coordinate of any point E whose corresponding angles is θ on the cross-section circle with centre point O' and radius a of the curve C . On the basis of this, the wireframe model, and surface model of the yarn can be constructed.

4 Simulation realization and discussion

4.1 SIMULATION STEP

Visual C++ and OpenGL are used to program the surface loop model [26, 31]. Based on the geometric model of the loop and the three dimensional surface yarn modelling, the three dimensional computer simulation of the plain loop can be achieved by programming, the algorithm is as follows:

Step 1 To input the information of the structural parameters of the loop, such as loop width, loop height, thickness, yarn diameter, colour etc.

Step 2 To calculate the coordinates of the coordinate value of data points on the loop unit respectively.

Step 3 To calculate the coordinate value of the data points on the loop and inversely to calculate the coordinate value of the controlling points of NURBS.

Step 4 To calculate the coordinate value of each interpolation point on the loop according to the controlling points by using NURBS.

Step 5 Based on the loop centreline, calculate the coordinates of discrete points at yarn cross-section.

Step 6 Lighting is set in order to increase the three dimensional effect of loop model.

Step 7 Use the drawing function of OpenGL, rendering loop model.

4.2 SIMULATION RESULTS

4.2.1 The centreline of loop model

When setting the parameters of loop, the interpolation points on the loop can be calculated by using NURBS

curves, and then the loop centreline can be obtained with the OpenGL of drawing line function, as shown in Figure 6.

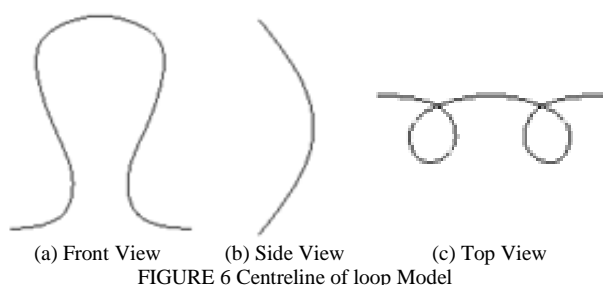


FIGURE 6 Centreline of loop Model

4.2.2 The wireframe of loop model

Based on the loop centreline, the points of cross-section circular are calculated, the wireframe of loop model is drawn with the line function of OpenGL, see Figure 7.

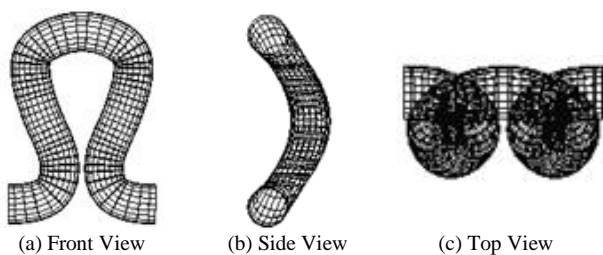


FIGURE 7 Wireframe of loop model

4.2.3 The surface of loop model.

Based on the loop cross-section circle, the surface of loop model is drawn by mesh chips, see Figure 8.

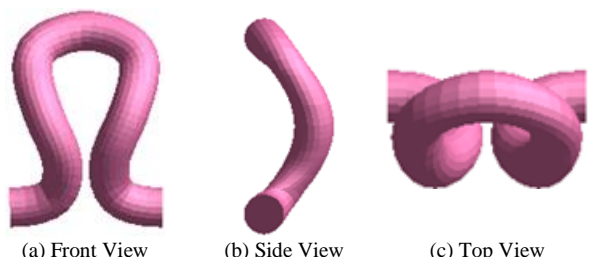


FIGURE 8 Surface of loop model

4.2.4 The plain structure simulation

Comparing the simulation results of the plain structure based on the simulation of the loop with the real fabric, the effect is shown in Figure 9.

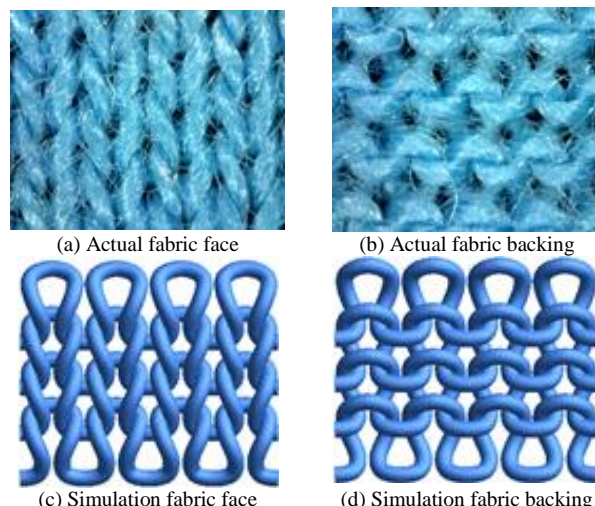


FIGURE 9 Actual fabric and Simulation result

4.3 DISCUSSION

According to Figure 9, the simulation results are relatively close to the actual fabric, it shows that's feasible to create surface model by using the NURBS curve, and has strong adaptability to the change of loop shape. The deficiency is that the use of mathematical tools is difficult.

Through constructing the loop model by surface model, although the process is with more trouble than method of sphere, but change more in its shape in cross-section, can also use the texture mapping technology, and can conveniently set the physical attributes. It is conducive to the development of CAD software.

5 Conclusions

The following conclusions can be drawn from the study: In this work, a three dimensional surface loop model is developed using NURBS curve and surface model of OpenGL mesh chips. This three dimensional model is successfully used to represent plain structure of weft knitted fabrics on the computer. Simulation results of knitted structures show that the NURBS curves and surface modelling by mesh chips have perfect realistic effect. By changing the shape of cross-section, using the surface model by mesh chips, complex cross-sectional shape of yarn can be obtained, for example, the yarn with twist. The yarn surface model by mesh chips is recommended to obtain more realistic representations of knitted structures in further research work in this area. In addition, by using texture mapping, the real yarn images are mapped to the surface of the three dimensional yarn models.

Acknowledgements

This work is supported by Natural Science Foundation of China (No.51303131); Natural Science Foundation of Tianjin, China (No.11JCYBJC264.00).

References

- [1] Deng Z M, Wang L J 2010 *Fibers and Polymers* **11**(3) 531-6
- [2] Kurbak A, Ekmen O 2008 *Textile Research Journal* **78**(3) 198-208
- [3] Chambeilain J 1949 *Hosiery Yarns and Fabrics* City of Leicester College of Technology: Leicester **2** 107
- [4] Peirce F T 1937 The geometry of cloth structure *J. Text. I.: Trans.* **28**(3) 45-96
- [5] Peirce F T 1947 Geometrical principles applicable to the design of functional fabrics *Textile Research Journal* **17**(3) 123-47
- [6] Leaf G A V, Glaskin A 1955 The geometry of plain-knitted loop *J. Text. I.* **46** 587-605
- [7] Leaf G A V Models of plain knitted loop *J. Text. I.* **51** 49-58
- [8] Munden D L 1959 The geometry and dimensional properties of plain-knit fabrics *J. Text. I.* **50** 448-71
- [9] Munden D L 1961 The geometry of a knitted fabric in its relaxed condition *Hosiery Times* **April** 43
- [10] Postle R 1971 Structure shape and dimensions of wool knitted fabrics *Journal of applied polymer science. Applied polymer symposium* **18** 1419
- [11] Demiroz A 1998 *A study of graphical representation of knitted structures* Ph.D. Thesis, University of Manchester Institute of Science and Technology: Manchester
- [12] Kurbak A 1988 Some investigations on the geometric properties of plain knitted fabrics *Tekstil ve Makine* **2**(11) 238
- [13] Kurbak A 1998a Plain knitted fabric dimensions (part I) *Textile Asia* **March** pp 33, 36, pp 41-4
- [14] Kurbak A 1998b Plain knitted fabric dimensions (part II) *Textile Asia* **April** pp 36-44
- [15] Liu S, Long H R 2007 Three-dimensional computer simulation of plain weft knitted fabric *J. Text. Res.* **28**(12) 41-4 (*In Chinese*)
- [16] Meng R J, Fang Y 2007 The modelling analysis of knitting fabric loop construction using NURBS curve *J. Zhejiang Sci-Tech Univ.* **24**(3) 219-24 (*In Chinese*)
- [17] Li Y L, Yang L H, Chen S Y et al. 2013 3D modelling and visual simulation of plain knitted structures *Int. J. Online Eng.* **9**(5) 59-61 (*In Chinese*)
- [18] Li Y L, Yang L H, Chen S Y et al. 2012 3D modelling and simulation of fancy fabrics in weft knitting *Journal of DongHua University (English Edition)* **29**(4) 351-8 (*In Chinese*)
- [19] Goktepe O 2001 Use of non-uniform rational B-spline for three-dimensional computer simulation of warp knitted fabric *Turkey J. Eng. Environ. Sci.* **25** 369-378
- [20] Wang J J, Wang H J, Yin Y 2012 A deformable image registration algorithm using NURBS *Int. J. Digital Content Technol. Appl.* **6**(10) 70-7 (*In Chinese*)
- [21] Chen L J, Zhang X, Li H Y 2012 Application in numerical control system of NURBS curve interpolation *Int. J. Adv. Comput. Technol.* **4**(10) 165-73 (*In Chinese*)
- [22] Alan H W 1999 *3D Computer Graphics with Cd rom (3rd edition)* Addison-Wesley Longman Publishing Co., Inc.: England (*In Chinese*)
- [23] Song G L, Jiang G M 2008 *Knitted Organization and Design* China Textile Press: China (*In Chinese*)
- [24] Long H R 2006 *Knitting* China Textile Press: China (*In Chinese*)
- [25] Su B Q, Liu D Y 1985 *Elementary Differential Geometry* Shanghai Science and Technology Press: Shanghai (*In Chinese*)
- [26] Sun J G, Hu S M 2005 *Basic Tutorial for Computer Graphics* Tsinghua University press: Beijing (*In Chinese*)
- [27] Zhu X X 2000 *Liberal Curve and Surface Modelling Techniques* Science Press: Beijing (*In Chinese*)
- [28] Shi F Z 2001 *Computer-Aided Geometrical Design and Non-uniform Rational B-Splines*, Higher Education Press: Beijing (*In Chinese*)
- [29] Han Q Y, Dong Y F, Shi X H 2007 Research of interpolation on cubic NURBS space curve *Coal Mine Mach.* **28**(1) 44-6 (*In Chinese*)
- [30] Lv D, Tong C M, Deng F S et al. 2006 Calculating control points of cubic NURBS curve *J. Projectiles, Rockets, Missiles Guid.* **26**(4) 357-9 (*In Chinese*)
- [31] Shreiner D 2007 *The Khronos OpenGL ARB Working Group OpenGL Programming Guide, 7th Edition* Addison-Wesley: USA

Authors

Yinglin Li, born in December, 1967, Shandong, China



Current position, grades: Tianjin Polytechnic University, Associate Professor, doctoral candidate
University study: Tianjin Institute of Textile Engineering (Clothing Department), 1986
 B.E. degree in Clothing Engineering, Tianjin Institute of Textile Engineering, Tianjin, China, 1990;
 M.A. degree (Software Engineering), Shandong University, Jinan, China, 2005
Scientific interest: Fabric modeling, Textile and garment enterprises informatization,
Publications: Textile enterprise management informatization, Garment factory design
Experience: Worked in Shandong Province Textile Design Institute (1990.7-2005.12); works in Tianjin Polytechnic University (2006.1-).
Professional Activities and Memberships: China Light Industry Machinery Association garment machinery branch (2013-)

Lianhe Yang, born in July, 1964, Hebei, China



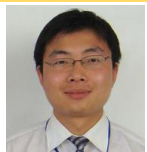
Current position, grades: Tianjin Polytechnic University, Professor, doctoral tutor
University study: Xi'an Electronic and Science University (Faculty of Computer Application), 1982
 B.E. degree in Computer Application, Xi'an Electronic and Science University, Xi'an, China, 1986;
 M.A. degree (Industrial Automation), Tianjin Institute of Textile Engineering, Tianjin, China, 1989;
 Ph.D. degree in Textile Science and engineering, Tianjin Institute of Textile Engineering, Tianjin, China, 1998
Scientific interest: The Mechanical Properties of Composite Materials, Computer Simulation and Computer Aided Design
Publication: Software development tool
Experience: works in Tianjin Polytechnic University (1989.3-).
Professional Activities and Memberships: Tianjin Association of automation, (1999-); Tianjin Association of Quantitative economics (2000-)

Suying Chen, born in May, 1966, Shandong, China



Current position, grades: Qingdao University, Associate Professor
University study: Tianjin Institute of Textile Engineering (Clothing Department), 1985
 B.E. degree in clothing engineering, Tianjin Institute of Textile Engineering, Tianjin, China, 1989;
 M.A. degree (fashion and textile design), Hongkong Polytechnic University, Hongkong, China, 2002
Scientific interest: Clothing Technology Theory, Clothing Construction Design and Human Engineering
Professional Activities and Memberships: Member of China Textile Engineering Society, (2001-)
Experience: works in Qingdao University (1989.7-).

Lei Xu, born in April, 1983, Shandong, China



Current position, grades: Tianjin Polytechnic University, Lecturer & Ph.D
University study: Qingdao University (Faculty of Textile engineering), 1999
 Ph.D. degree in Textile Science and engineering, Tianjin Polytechnic University, Tianjin, China, 2010
Scientific interest: Mechanics and Geography of fiber bundles and textile structures; Intelligent textile and composite structures.
Experience: works in Tianjin Polytechnic University (2011.7-).

Three-dimensional computer simulation of soil nailing support in deep foundation pit

Chang Zhi Zhu^{1,2*}, Quan Chen Gao¹

¹ School of Mechanics & Civil Engineering, China University of Mining & Technology, Beijing 100083, China

² School of Urban and Rural Construction, Agricultural University of Hebei, Baoding 071001, China

Received 1 March 2014, www.tsi.lv

Abstract

The computer application program that is applied based on the finite difference method. By taking the soil nailing support structure in Shijiazhuang city as an example, the three-dimensional computer numerical model of deep foundation pit is set up; the horizontal displacement and the ground settlement of the deep foundation pit are simulated in the process of excavation and support. The simulation result is consistent with the test result. The results show that the deformation behaviour of the deep foundation pit can be analysed by using three-dimensional computer simulation technology in actual project. The method overcomes the deficiency of theoretical analysis method and offers effective guidance for design and construction of foundation pit excavation and support.

Keywords: Three-dimensional computer simulation, Soil nailing support, Deep foundation pit

1 Introduction

Because of its economy, stability and flexibility, the method of soil nailing was widely applied to the foundation pit support [1, 2]. In the process of the design of soil nailing supporting structure, the design method adopted was combination theoretical analysis with experience of practical engineering, but there are a number of assumed conditions in the process of theoretical analysis, the interaction mechanism is very complicated [3, 4] between the supporting structure and soil, these will cause a lot of parameter values that cannot be calculated exactly, which may affect the foundation pit itself and safety of surrounding buildings.

In recent years, with the rapid development of computer simulation technology, many researchers are using this technology in the field of engineering design or scientific research. Thus in order to overcome the deficiency of pure theoretical analysis in the soil nailing structure design, the method of computer simulation technology was introduced in the analysis of the soil nailing structure, the interaction mechanism was simulated by the computer program, and analyse the deformation behaviour in the process of the deep foundation pit excavation and support.

In this paper, taking the soil nailing support structure in Shijiazhuang city as an example, based on finite difference program, the process of the deep foundation pit excavation and soil nailing support was simulated by the numerical analysis model, analyse the law of the horizontal displacement and the ground settlement, the numerical simulation and measured result can be used to analyse the deformation behaviour, the result can be used

to offer effective guidance for design and construction of foundation pit excavation and support.

2 Engineering situation

The foundation pit of a high-rise building was in Shijiazhuang, the depth was 13 m, the total construction area was 62585.46 m², there was a 3-story office building on the top of the foundation pit east. The length of building was about 50 m, the width was 13 m. The distance from the axis of the exterior wall to the edge of foundation pit was about 2 m. Plane arrangement chart of the site was shown in Figure 1.

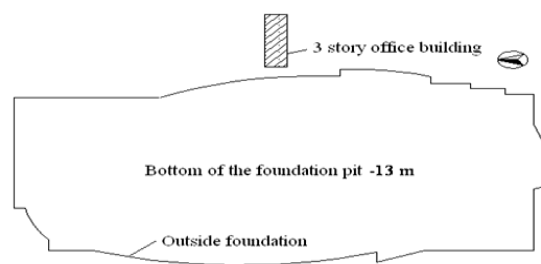


FIGURE 1 Plane arrangement chart of the site

The soil physical and mechanical properties were shown in Table 1. The slope of pit was 90°, soil nails was used in the process of the excavation of foundation pit, the soil nails was eight rows in number, the diameter of hole was 100 mm, and the angle with the horizontal plane was 15°. The horizontal and vertical distance was 1.0 m, the steel's diameter was 22 mm, and the length of the soil nailing was shown in Figure 2.

* Corresponding author e-mail: 13930854216@139.com

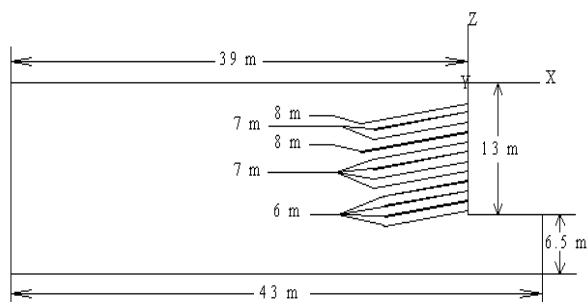


FIGURE 2 The profile of foundation pit support design

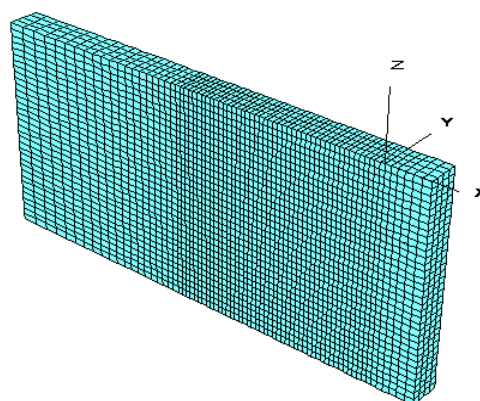


FIGURE 3 Model diagram

3 Calculation model

3.1 THE GEOMETRY MODEL OF FOUNDATION PIT

According to engineering experience, when the soil nails were arranged in the soil at a certain distance along the horizontal and vertical directions, the most dangerous parts of the pit was generally not near the corner, but in of the central part [5, 6] of the long side of the foundation pit. So 3 m wide range soil of the central part of the long side of the foundation pit was simulated by establishing the numerical model, the model geometry was determined by trial calculation, and the specific dimensions were shown in Figure 2. The numerical model was shown in Figure 3. In order to meet accuracy requirements, the regional units were small in the range of soil nail reinforcement regional, and it was a total of 9588 regional units.

3.2 BOUNDARY CONDITIONS OF MODEL

The boundary conditions of the numerical model were these: The bottom surface of the foundation pit was fixed, the displacement of horizontal and vertical direction was restricted; the top surface and air face after excavating were free; the horizontal displacements of other side surfaces were restricted, and vertical displacements were free.

3.3 VALUES OF MATERIAL PARAMETER IN MODEL

The physical and mechanical parameters of the soil were shown in Table 1. Soil nails were stimulated by cable element, and the parameters were shown in Table 2.

TABLE 1 The soil physical and mechanical properties

Soil No.	Soil name	Thickness (m)	γ (kN·m ⁻³)	c(kPa)	ϕ (°)	K(MPa)	G(MPa)	μ
1	fill	0.80	16.0	15.0	10.0	36.25	27.18	0.4
2	silty clay	2.60	19.2	20.0	17.6	8.02	3.70	0.3
3	loess-like silt	1.40	19.6	9.6	23.4	6.75	4.05	0.25
4	silty clay	3.20	19.2	22.0	22.8	10.14	4.68	0.3
5	fine sand	1.20	18.0	0.0	28.0	50.56	30.33	0.25
6	medium sand	6.10	18.0	0.0	33.0	50.33	35	0.25
7	fine sand	3.2	18.6	0.0	36	54.3	36	0.24
8	silty clay	2.20	19.4	25.0	21.8	12.14	5.23	0.29
9	coarse sand	1.2	19.6	0.0	26.0	54.2	35.7	0.23
10	silty clay	8.20	19.9	20.0	20.0	12.03	5.55	0.3

TABLE 2 The calculation parameter values of soil nails

Elastic modulus of steel bar (kPa)	Tensile strength of steel bar (MPa)	Density of steel bar (kg·m ⁻³)	Mortar stiffness (kPa)	Mortar cohesion (kPa)	Mortar friction angle (°)	Cross sectional area of soil nailing (m ²)
2.0e8	335	7800	185.1	5.4e5	25	2.011e-4

3.4 NUMERICAL SIMULATION

In the numerical analysis model, the constitutive model of the soil was non-linear elastic-plastic, the failure criteria was Mohr-Coulomb criteria [7], It can reflect the characteristics of the soil well [8], due to the deformation of soil under the action of soil nails and the tension of soil nail were closely related with the construction

process [9], so the process of numerical simulation was accordance with the actual construction process rigorously, concrete steps are as follows:

- 1) The model was calculated on the condition of gravity stress, and the initial stress state of soil was obtained.
- 2) Initialize the displacement, velocity, boundary conditions, and apply external load, calculate the stress.

3) Simulated excavation, after the first step excavation, the balance calculation was done, and then soil nails were driven into the soil, calculating again, until the calculation precision to meet the requirements for the next step, and so on, until the end of excavation and support process.

4 Numerical results and analysis

Figure 4 to Figure 7 were the change curve of horizontal displacement with the depth of the excavation of the excavation surface and the positions away from the excavation surface were 3 m, 6 m, and 9 m, respectively. From the figures, we can see, with increasing of excavation depth, the lateral restraint of slope was gradually lifted, on the action of gravity stress and external loads, the horizontal displacement of the soil increased with the number of excavation steps increased at the same location of slope. After every step of the excavation is completed, of the horizontal displacement of the slope soil decreased with the soil depth increased. By comparing Figure 4 to Figure 7, the horizontal displacement of the slope that was at the same the depth of soil decreased with the distance from the excavation surface increased.

Figure 8 was the change curve of the settlement of the top of slope with the change of the number of excavation steps. Moreover, it shows that in the early stage of excavation, as stress redistribution of the soil caused by the excavation, there was a certain amount of resilience. Therefore, it showed a range of rising of the excavation surface. With the excavation depth increases, the settlement of the slope soil increased on the action of gravity stress and external loads. The settlement of excavation surface was the largest, and it decreased with the distance from the excavation surface increased.

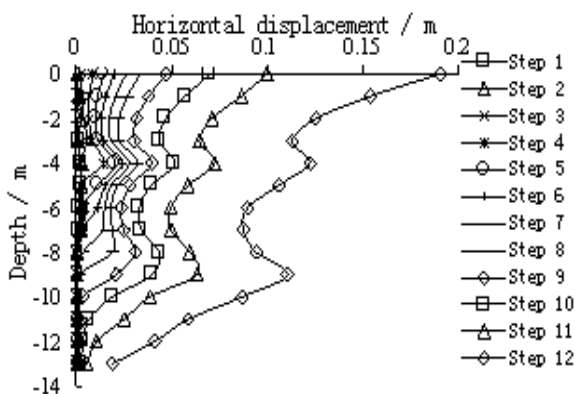


FIGURE 4 The change curve of horizontal displacement with the depth of the excavation surface.

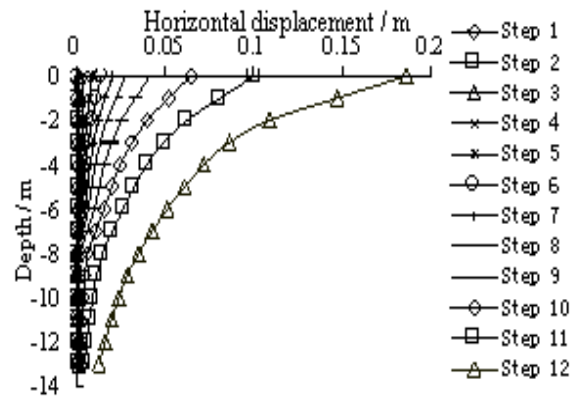


FIGURE 5 The change curve of horizontal displacement with the depth of the excavation that the position away from the excavation surface was 3 m

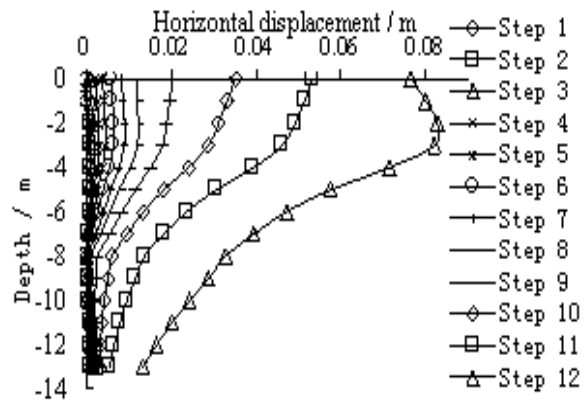


FIGURE 6 The change curve of horizontal displacement with the depth of the excavation that the position away from the excavation surface was 6 m

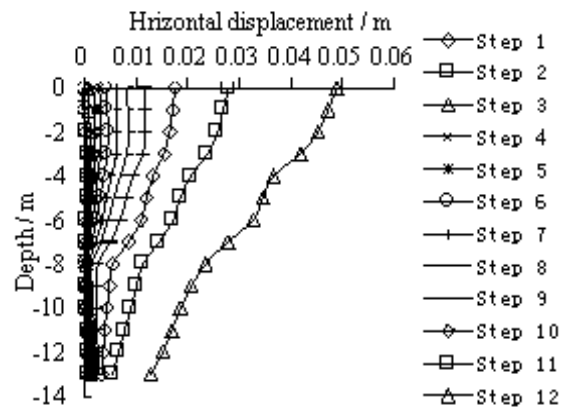


FIGURE 7 The change curve of horizontal displacement with the depth of the excavation that the position away from the excavation surface was 9 m

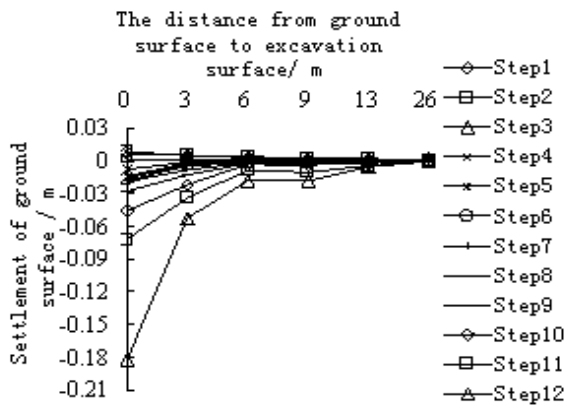


FIGURE 8 The change curve of the settlement of the top of slope with the change of the number of excavation steps

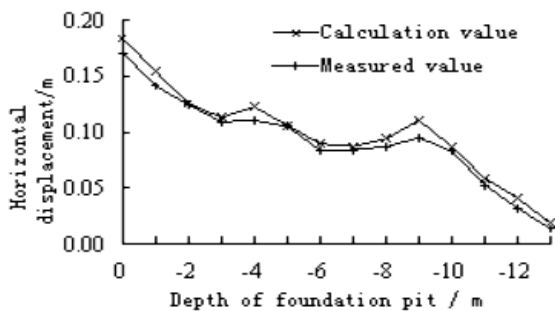


FIGURE 9 The comparison curve of the displacement of calculation value and measured value of excavation surface

Figure 8 shows that the settlement of the top of the slope of position that the distance from the excavation surface was 9 m increased, and this phenomenon was more obvious when the excavation for the tenth and eleventh step by step. The reason was that there was a building on the top of the excavation surface of the slope, and on the action of the load of building, the settlement of the surface increased.

Figure 9 is the change curve of the simulated values and the measured values of horizontal displacement with the depth of the foundation pit after the excavation completed, the simulation result was consistent with the test results. Whether calculated or measured horizontal displacement were large at the depth 9 m, Mainly

References

[1] Cheng Z Y, Cui J H 2000 *Applications of Soil Nailing in the Pit Engineering*, 2nd edition, China Building Industry Press: Beijing (In Chinese)
 [2] Gong X N 1998 *Manual of Deep Foundation Pit Design and Construction*, 1st edition, China Building Industry Press: Beijing (In Chinese)
 [3] Zeng X M, Lin G 2002 Experiment of soil nailing reinforcing mechanism in soft soil slope *Chinese J. Rock Mech. Eng.* **21**(3) 429-33 (In Chinese)
 [4] Duan T 2005 Soil anchor connected bracing for deep pit of foundation, Master's Thesis, *Wuhan Univ.*: Wuhan, China (In Chinese)

because the soil were fine sand or medium sand, this type of the soil conditions were not conducive soil nailing to play the role of supporting structure, so that the horizontal displacement is too large.

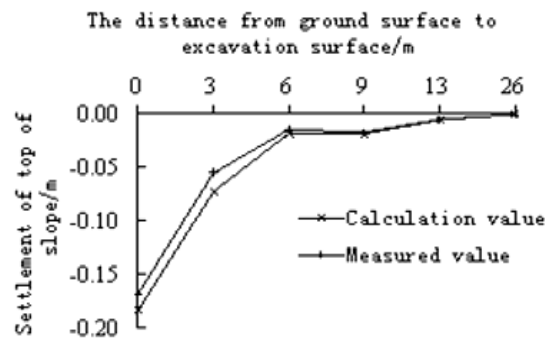


FIGURE 10 The comparison curve of displacement calculation value and measured value of ground surface



Figure 10 is the related curve of the simulated values and the measured values of the settlement of the top of the slope with the distance away from the excavation surface of the foundation pit after the excavation completed. It shows that the simulation results were consistent with the measured ones. The final calculation value of settlement is 0.183 m, and the measured one is 0.168 m, the relative error is only 8.2%.

5 Conclusions

The following conclusions can be drawn from the study:

1. Based on practical engineering, a numerical analysis model was established by the finite difference program, and the process of the deep excavation and support was simulated by computer, the horizontal displacement of the slope and settlement of the top of the slope are obtained, the results shows that there not only has the same regularity but also good agreement between calculated and measured values.
2. The fine difference program can be used to the simulation of the excavation and support of deep foundation pit, and it will provide the basis for the design and construction of practice project.

[5] Liu J G, Zeng Y W 2006 Application of FLAC3D to simulation of foundation excavation and support *Rock Soil Mech.* **27**(3) 505-8 (In Chinese)
 [6] Li X K 2002 The study on soil nailing based on soil arching action, Master's Thesis *Central South Univ.*: Changsha, China (In Chinese)
 [7] Peng W B 2008 *FLAC3D Practical Guide* 1st edition, China Machine Press: Beijing (In Chinese)
 [8] Wang G G, Du M F 2000 Large deformation analysis for braced excavation *Chinese J. Rock Mech. Eng.* **19**(4) 509-12 (In Chinese)
 [9] He R L, Zhang P 2007 Numerical analysis of the reinforcement of soil-nailing support *J. Human Univ.* **34**(1) 14-8 (In Chinese)

Authors	
	<p>Changzhi Zhu, born in April, 1980, Zunhua county, Hebei province, China</p> <p>Current position, grades: Work in urban and rural construction institute, Agricultural University of Hebei, and doctoral candidate in geotechnical engineering, China University of Mining & Technology</p> <p>University studies: Urban and rural construction institute (Faculty of civil engineering, 1999), Master of engineering degree in structural engineering, agricultural university of Hebei, Baoding, China, 2003; Doctoral candidate in geotechnical engineering, China University of Mining & Technology, (2012-)</p> <p>Scientific interest: Computer simulation of geotechnical engineering</p> <p>Publications: Published 7 papers</p> <p>Experience: Member of Chinese Sub-Society for Soft Rock Engineering & Deep Disaster Control, (2004-)</p>
	<p>Quanchen Gao, born in November, 1957, Wanrong county, Shanxi province, China</p> <p>Current position, grades: Professor of China University of Mining & Technology</p> <p>University studies: School of Mechanics & Civil Engineering (Faculty of underground engineering, 1977); Master of engineering degree in mine construction engineering, China University of Mining & Technology, Beijing, 1984; Doctor of engineering degree in mine construction engineering, China University of Mining & Technology, Beijing, 1988</p> <p>Scientific interest: Deep foundation pit support engineering</p> <p>Publications: more than 20 papers</p> <p>Experience: Vice Chairman of Blasting Professional Committee of China Coal Society (2006-); Director of China Society of Engineering Blasting (2003-)</p>

Organizations learning mechanism in the cyber society ecology system: an agent-based simulation

Bing Bai¹, Xiuquan Deng^{2*}, Zhiqiong Guo¹, Dehua Gao²

¹ Business School, Jiangsu Normal University, Xuzhou 221116, China

² School of Economics and Management, Beihang University, Beijing 100191-China,

Received 1 March 2014, www.tsi.lv

Abstract

In this paper, we try to focus on the cyber society ecology system, which is a naturally occurring complex system to a certain stage of networks economic development. Based on the multi-agent simulation methodology, this paper analyses adaptive learning among organizations of cyber society ecology system, and then designs interaction rules of agents and simulation parameters, and finally the simulation results are analysed.

Keywords: Cyber Society Ecology System, Adaptive Learning, Agent-Based Simulation

1 Introduction

The Cyber Society is a new social form derived from the emergence of computer and Internet technology, which serves as the representatives of the information technology. From the ecology perspective, some scholars proposed the concept of “Cyber Society Ecosystem” [1]. The ecosystem of cyber society is a whole within a certain time and space, consisting of various ecological groups, resources and environment in the cyber society, which is a naturally occurring complex system to a certain stage of networks economic development [2]. Comprehensively understanding the feature of the cyber society ecosystem, revealing its internal evolution mechanism has an important theoretical and practical significance for promoting rapid development of the internet [3]. This paper, based on complex adaptive system theory and ecology theory, uses agent-based modelling and simulation method to study organizations learning mechanism in the cyber society ecology system, so as to propose a new direction for understanding and studying the formation and evolution of cyber society ecosystem. There are many agents in the cyber society, such as the internet users and the virtual enterprise etc. These are parts of the cyber society ecosystem, as the biological organisms in the biological ecosystem, which makes the cyber society ecosystem having the same adaptive evolution process as the biological systems.

The structure of the paper is as follows: section 2 analyses the theory and methodology; section 3 introduces the model and simulation process; section 4 analyses the simulation results; and section 5 discusses the conclusions.

2 Theory and methodology

2.1 CAS THEORY

The complex adaptive system theory is proposed by Holland [4], who is one of the members of SFI School, which is focus on studying complex science. Its core idea is “Adaptability creating Complexity” that the evolution of system benefited from the “living” agent. In order to adapt to the environment or to win the right to survive, agents will adjust their behaviours constantly according to the external environment and other agents [5]. Complex adaptive system theory provides a new perspective to explore the complexity of organization such as the evolution of the cyber society ecosystem.

2.2 AGENT-BASED SIMULATION IN SOCIAL AND ORGANIZATIONAL THEORY

Generally speaking, the social and economic phenomena have complex and dynamic characteristics. These lead to the conclusion that traditional tools, such as mathematics and experimental techniques etc., often do not provide adequate descriptions or explanations of these complex phenomena. While the computer-based simulation method, especially the agent-based simulation, has a unique advantage in exploring the complex and dynamic organizational and social phenomena, the method has been widely used in the past few years and earned a wide acceptance [6, 7]. Agent-based simulation takes the systems’ macro-phenomena as the results of interactions within the micro-level actions [7]. By defining the agents and their interactions rules, as well as their environments, simulation models can reveal some mechanisms of these macro-phenomena, and may even be used for theory

* *Corresponding author* e-mail: dengxiuquan@buaa.edu.cn

discovery and hypothesis generation. In conclusion, agent-based simulation is an effective method to study the 'emergence' of macro level phenomena from micro level actions [7]. Consequently, this paper studies the adaptive learning among organizations of cyber society ecosystem by using agent-based simulation methodology.

2.3 THE CYBER SOCIETY ECOSYSTEM AS CAS

The cyber society ecosystem exchanges massive information, matter and energy with external system, and lies in a non-equilibrium status. The exchanging process is complex adaptive process which is from disorder to order, and then from one order to another order. The cyber society ecosystem will transmit information to internal system through a feedback mechanism, and evolve by the driven force in the internal system. The cyber society ecosystem acquires the ability to adapt to the environment by self-organization and restructuring. Consequently, the cyber society ecosystem is a representative complex adaptive system, which ties the elements of the network society and constraints spontaneous formation.

3 Model Design and Simulation

3.1 MODEL ASSUMPTIONS

In order to simplify the analysis, we need to statement some assumptions and confirm the simulation model boundary. The assumptions of the simulation model are followed:

(1)The agents of the simulation model are the organizations of cyber society ecology. We premise that the organizations are homogeneous in the cyber society ecology.

(2)The agents are bounded rationality. The agents cannot find the optimal decision in the beginning. They only have limit knowledge, and improve their decision through constantly continuous learning.

(3)Using random values to replace the parameters, which cannot be measured and the information, which cannot be estimated accurately [8].

3.2 DEFINING THE AGENT AND IT'S ATTRIBUTE

The agents in the model are the abstract of virtual organization, and have following attributes:

(1)Cyber Society Space. Cyber Society Space is the position of organizations or agent in the grid, separately locates the expression through X-axis and the Y-axis.

(2)Boundary Value. Boundary value reflects the intensity degree of mutual learning among virtual organizations in the cyber society ecosystem. The higher the boundary value is, the greater the degree of mutual learning. The value is 0 that means there is not learning relationship from each other.

(3) Agent's Degree. Agent's Degree is the number of edges, which the agent owns in the position. The more the number, the more linkages between the agent and other agents, otherwise there is less relationships.

(4) Absorption Capability. Absorptive capability is the agent's ability to access knowledge in his location by learning from the agents in other locations. The stronger absorptive capability is, the greater acquiring knowledge, and vice versa. Absorption capability is measured by the absorption capability coefficient in the simulation model.

(5) Knowledge Stock. Knowledge stock reflects the knowledge quantity, which the agent has in this location. The more the knowledge stock is, the higher the learning probability.

3.3 ADAPTIVE LEARNING PROCESS OF AGENT

Adaptive learning process of the cyber society organization is the process to exchange knowledge among members. From the multi-agent system perspective, we regard the members of cyber society ecosystem as the agents of system, and then the learning process of the system is the flow of knowledge between the agents. The essential process of knowledge flowing is that the agents transfer knowledge to their "neighbouring agent", and then these neighbouring agents transfer knowledge to their "neighbouring agent" too. Under circumstance, the flowing of knowledge will occur on the network, which is composed of by neighbouring agents. This network has shown significant characteristics of small-world networks. Therefore, this paper analysed the adaptive learning process of the members of cyber society ecological system according to the relevant outcomes of the small world network theory.

According to the characteristics of small-world networks, we mapped the agents as the node of network, and used $(N=i, i=1,2,3,\dots,n)$ to stand for the set of agents in the cyber society ecosystem. The N includes a series of channels of knowledge flowing, and each channel is connected to two or more network nodes, which express the learning relationship among the agents. General speaking, for the agent i which is belong to N , has k_i boundaries that connected to other agents, and we call the k_i agent as the neighbouring agents of agent i . We use the set of $N(G,i)$ to stand for all the neighbours of agent i . That's to say, if agent j is belong to $N(G,i)$, the agent j is the neighbours of agent i and the node i and j will become a boundary. The value of boundary reflects the strength of the degree of connection between agents. The value of boundary is larger, and the degree of connection between agents is stronger. If the value is zero, it represents that there are no learning connection between agent i and agent j . The degree of the node indicates the number of boundaries of the agent i .

The learning rule is the system randomly selected an agent k from the small word network N . Then the agent k gain the set of neighbour $N(G,k)$ by the boundary value. In each period, the agent k searches for each neighbour,

and compares the size of each neighbour knowledge stock with it owns. If the neighbouring agent's j knowledge about one type is more than the agent's k what's more, the agent j's absorption capability is the most of all the neighbouring agents, the agent k will learn this type knowledge from the agent j. The value of Absorption capability determines the size of the knowledge acquired in the network.

3.4 IMPLEMENTATION OF SIMULATION

Swarm is the widely used agent-based simulation platform, so we realize this simulation program under the Swarm-2.2-java and the jdk1.6.0_10 development package. The main program documents include Enterprise.java, CyerSocietySpace.java, Adaptivelearning Mechanism.java, ModelSwar-m.java, ObserverSwarm. -java, StartSimulation. java and so on.

3.5 VERIFICATION AND VALIDATION OF THE SIMULATION MODEL

Before running the model, we need to test the program. The testing procedure includes validation and a detailed verification process [9]. The way verification is perceived and deployed depends on the model design methodology and the paradigm utilized for its representation [10]. For the purpose of model validation, we have opted to analyse each agent separately and verify to what extent the results of the experiments correspond to real world data [10]. This paper adopts the unit testing method for verification of the program.

The testing results verify our assumptions. According to the model proposed by Sargent [11], we test the validation according to four components: conceptual model validity, computerized model verification, operational validity, and data validity. The results show that the model's output is very close to the real world. These results indicate that the model's response accuracy is within an acceptable range for its intended purpose and therefore the model is valid.

4 Analysis of the Simulation Result

4.1 MODEL INITIALIZING

In order to make the simulation results reasonable, we should initialize the variables which we statement at the simulation experiment beginning. The generation of small world network is based on the WS model, which is proposed by Watts and Strogatz [12]. According to Watts and Strogatz 's point, we may get small world network by denoting the reconnection probability whose value is 0.01. Through a number of tests of the model, we repeated modifications and corrections for the program, and obtained the reasonable value of the related parameters. The interface of initializing parameters is following:

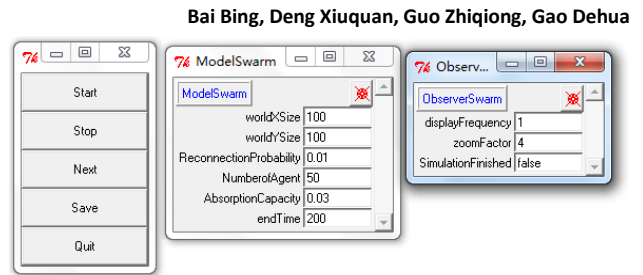


FIGURE 1 The Interface of Initializing Parameters

4.2 ANALYSIS OF THE SIMULATION RESULT

From simulation schedule 1 to 100, the change in the distribution of cyer society space and the learning net are following:



FIGURE 2 Adaptive Learning Network

The Figure2 shows the complicated knowledge flow network forming through long practical learning among interior organizations in the cyber society ecosystem system. With the simulation time going on, the knowledge flow network is becoming more and more complex. The line among agents represents the relevance of knowledge flowing. The more the lines, the more the knowledge are absorbed. By adjusting model parameters, when setting the knowledge stocks value is very high we can find the following findings as shown below. Adjusting the various parameters, we obtain the following simulation results:

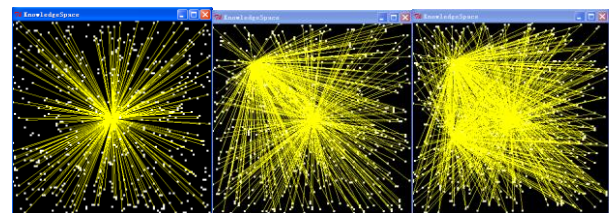


FIGURE 3 The Concentration of Adaptive Learning

The above Figure3 shows when the agent's knowledge stocks in the cyber society ecosystem system is higher, the adaptive learning network presents centre salient features, and learning among agents appears agents keeping higher knowledge stores flow and spread outwards.

The simulation results did just preliminary finish the simulation on learning process and knowledge flowing among organizational members in the cyber society ecosystem system. The problems of relevance about how to measure knowledge flow still need to further research and develop.

5 Conclusion and Discussion

This paper, using agent-based simulation, studied learning process of organization in the cyber society ecology from the complex adaptive system perspective. Through analysis of the simulation result, we draw the following conclusions:

(1) The cyber society ecology system is a complex adaptive system, and has complex adaptive characters. Its complexity is reflected in the structure of cyber society ecology system. The structure of cyber society ecology system includes internet users, virtual enterprises and so on. They all have initiative and adaptive, like biological organisms in the biological ecosystem.

(2) Adaptive learning among internal organizations in the cyber society ecology is an important driving force for the evolution of the cyber society ecology system. Adaptive learning spurs the cyber social to develop.

(3) The network group with subjective initiative, such as internet users and virtual enterprises etc., are part of

organisms in all elements of the cyber society ecology system, as living organisms in the biological ecosystem. This makes the formation for the self- organization process possible in the cyber society ecology system like the biological system.

These results are helpful in understanding the adaptive learning process among agents in the cyber society ecology system, and make a base for continued modelling and simulation of the cyber society ecology system evolution.

Acknowledgments

The research work in this paper is supported by the grants from National Natural Science Foundation of China (No. 71302188), the grants from Philosophy and Social Science Fund of Colleges and Universities in Jiangsu Province(No. 2013SJB6300032), and Beijing Natural Science Foundation (No. 9142013).

References

- [1] Ji Zhen Zhang, Run Tong Zhang 2008 *Ecosystem of Cyber-society* Electronic Industry Press of China (In Chinese)
- [2] ShanWang 2010 *Research on the Evolution of the Ecosystem of Cyber society Based on Self-Organization Theory* BeiJing Jiaotong University, Dissertation (In Chinese)
- [3] Xiaolan Guan 2011 *Research on the Formation Mechanism of Ecosystem of Cyber-society* Beijing Jiaotong University, Dissertation (In Chinese)
- [4] Holland J H 1995 *Hidden Order: How Adaptation Builds Complexity* Addison Wesley Publishing Company
- [5] Guoling Lao, Luyuan Xiao, Rong Zhou 2008 CAS-based Enterprise Knowledge Sharing Modeling and Simulation *Wireless Communications, Networking and Mobile Computing, WiCOM 08. 4th International Conference*, July pp 1-4
- [6] Axelrod R 1997 *The Complexity of Cooperation: Agent-based Models of Competition and Collaboration* Princeton University Press, New Jersey
- [7] Gilbert N 2004 *Agent-based Social Simulation: dealing with Complexity* <http://cress.soc.surrey.ac.uk/web/resources/ABSS%20-%20dealing%20with%20complexity-1-1.pdf>
- [8] Xiuquan Deng, Haorun Huang 2008 Research on the Concept Model of Enterprise Capability Multi-Agent Simulation Based on CAS *The 38th International Conference on Computers and Industrial Engineering* pp 818-828
- [9] Cil I, Mala M 2010 A multi-agent architecture for modelling and simulation of small military unit combat in asymmetric warfare *Expert Systems with Applications* **37** 1331-1343
- [10] Levent Yilmaz 2006 Validation and verification of social processes within agent-based computational organization models *Computational & Mathematical Organization Theory* **12** 283-312
- [11] Sargent R G 1999 Validation and verification of simulation models *In P. A.Farrington, H. B. Nembhard, D. T. Sturrock, & G. W. Evans (Eds), Proceedings of the winter simulation conference, Phoenix, AZ, USA* 39-48
- [12] Watts D J, Strogatz S H 1998 Collective Dynamics of 'Small-world' Networks *Nature* **393**(6684) 409-410

Authors



Bing Bai, born on January 28, 1979, Xuzhou, China

Current position, grades: Doctor of Management Science and Engineering, associate professor and master supervisor in Jiangsu Normal University

University studies: Management Science and Engineering in Beihang University

Scientific interest: Organization Strategy; Cyber Society Ecology System

Publications: 22 Papers

Experience: Dr.Bai is an associate professor in the Business School at Jiangsu Normal University in China. He received his Doctor degree from Beihang University (Beijing University of Aeronautics and Astronautics) in 2012. He majors in management science and engineering. His current research interest focuses on management information system and organization strategy from agent-based modelling and simulation perspective.



Xiuquan Deng, born on April 18, 1979, Beijing, China

Current position, grades: Doctor of Management Science and Engineering, associate professor and master supervisor in Beihang University

University studies: Management Science and Engineering in Beihang University

Scientific interest: organization strategy; industry engineering

Publications: 36 Papers

Experience: Xiuquan Deng is an associate professor in the School of Economics and Management, Beihang University, China. His research interest includes the organization strategy, complex system simulation, and industry engineering.



Zhiqiong Guo, born on July 26, 1977, Xuzhou, China

Current position, grades: Master of Resource Economics and management, assistant professor in Jiangsu Normal University

University studies: Resource Economics and management in Shandong University of Science and Technology

Scientific interest: Strategy management

Publications: 8 Papers

Experience: Ms.Guo is an assistant professor in the Business School at Jiangsu Normal University in China. She received his Master degree from Shandong University of Science and Technology in 2010. She majors in Resource Economics and management. Her current research interest focuses on strategy management.



Dehua Gao, born on August 16, 1982, Beijing, China

Current position, grades: a Ph.D student in the School of Economics and Management, Beihang University.

University studies: Management Science and Engineering in Beihang University

Scientific interest: complex system modeling and simulation

Publications: 18 Papers

Experience: Mr.Gao is a Ph.D student in the School of Economics and Management, Beihang University. He majored in Management Science and Engineering. His current research interest focused on Multi-agent Simulation in Enterprise capabilities

The approach of fixed asset management based on the shortest path

Xing-bin Ma^{1*}, Cui-pin Jiang²

¹ Department of Finance and Economics Shandong university of science and technology, Jinan Shandong, 250031

² Department of Public Courses Shandong University of Science and Technology, Jinan Shandong, 250031

Received 1 January 2014, www.tsi.lv

Abstract

We often meet with shortest path problem in National Undergraduate Mathematical Contest and practical life. The definition of shortest path problem is introduced, Dijkstra algorithm and 0-1 Programming Method to solve the shortest path problem are given. A practical problem is given and is calculated by these two methods.

Keywords: mathematical modelling, shortest path problem, Dijkstra algorithm, 0-1 Programming

1 Introduction

The shortest path is a very important problem in the graph theory. Many practical problems can be translated into the shortest path problem or into the sub-problem of it, so it is often met with in National Undergraduate Mathematical Contest in Modelling. The Shortest Path Problems are usually used in GISs, [1] and in facility location problems, [2] and in the project design of the laying the pipeline system. How to look for shortest path is the key to solve the intelligence traffic [3]. The Dijkstra algorithm is widely considered to be the excellent algorithm to solve the shortest path in graph theory [4, 5]. And, Moto, et al proposed the method that how to find the shortest path in a certain period of time based on The Dijkstra algorithm [6]. On this basis, someone try to find the method for the extend shortest path [7]. The equipment update timing selection can be generalized into the Shortest Path Problem, as a consequence, these have a practical significance to get hold the method to solve this kind of problem.

2 The Shortest Path Problem and it's solving method

2.1 DEFINITION

The shortest path problem is that to find a path from v_s to v_t in the weighed direct graph (the weight can be the length of the path, or the cost depending on require of specific questions), and the path that all the weighting sum of the arc is minimum number is called be shortest path from v_s to v_t , the weighting sum of the arc is called the distance from v_s to v_t .

2.2 THE SOLVING DIJKSTRA ALGORITHM ON SHORTEST PATH [8]

The Dijkstra algorithm applies to solve the shortest path problem in the condition that the weights w_{ij} of all arc (v_i, v_j) are bigger than 0, so the Dijkstra algorithm can be called double labelling method also. Double labelling method is that assign the point v_j two label (l_j, k_j) , the first label l_j mean the length of shortest path and the second k_j mean a subscript of adjacent points front v_j on shortest path from v_s to v_j , then we can find the shortest path from v_s to v_j and the shortest distance from v_s to v_j .

The following is the concrete steps of the Dijkstra algorithm. First, to assign the start point v_1 as label $(0, s)$ that mean the distance is 0 from v_1 to v_1 . Second, to suppose a set of labelled points I and a set of unlabelled points J , a set of arcs $\{(v_i, v_j) | v_i \in I, v_j \in J\}$, the arcs in this set are the arcs from the labelled points to the unlabelled points. Last, the calculations are finished if the arc set is empty, the shortest length is l_t if v_t is labelled (l_t, k_t) and the shortest path from v_1 to v_t can get by trace-back from k_t to v_1 , the shortest path is non-existent if v_t are never labelled. If the arc set is nonempty, we can work out $s_{ij} = l_i + w_{ij}$ corresponding to every arc (v_i, v_j) . There should be the arc that have minimum value in all s_{ij} and this arc is supposed as (v_c, v_d) , the ending point v_d of this arc as (s_{cd}, c) . Return to the second step. If

*Corresponding author e-mail: maxingbin@126.com

there are many arcs that have the minimum value s_{ij} in third step, we can choose any one to label the end points of these arcs, or to label on every one also.

2.3 THE SOLVING SHORTEST PATH PROBLEM ON THE 0-1 PROGRAMMING THEORY [9]

Suppose v_1 as start point and v_n as end point. Introduce 0-1 decision variable x_{ij} ,

$$x_{ij} = \begin{cases} 1, & \text{if the arc}(v_i, v_j) \text{ on the shortest path} \\ 0, & \text{if the arc}(v_i, v_j) \text{ not on the shortest path} \end{cases}$$

The arc of all ones that start from point $v_i (1 < i < n)$ must be on shortest path if $\sum_{j=1}^n x_{ij} = 1$ for any vertex $v_i (1 < i < n)$, that is to say, this vertex must be on the shortest path, and the arcs from others vertex to this must be on the shortest path, so $\sum_{j=1}^n x_{ji} = 1$, the vertex $v_i (1 < i < n)$ is not on the shortest path if $\sum_{j=1}^n x_{ij} = 0$, so

there will be $\sum_{j=1}^n x_{ji} = 0$; Combining the above two cases,

we can get $\sum_{j=1}^n x_{ij} = \sum_{j=1}^n x_{ji}, 1 < i < n$. There must be

$$\sum_{j=1}^n x_{1j} = 1 \text{ for start point } v_1 \text{ and there must be } \sum_{j=1}^n x_{jn} = 1$$

for end point v_n . The value of the objective function that sum up weight of every arcs on shortest path is minimum, so 0-1 Programming Model that solve the shortest path problem is as: objective function: $\min z = \sum_{(v_i, v_j) \in E} w_{ij} x_{ij}$ (E

is a set of all arc in chart)

$$\begin{cases} \sum_{(v_i, v_j) \in E} x_{ij} = \sum_{(v_i, v_j) \in E} x_{ji}, 1 < i < n \\ \sum_{(v_1, v_j) \in E} x_{1j} = 1, \sum_{(v_j, v_n) \in E} x_{jn} = 1 \\ x_{ij} = 0 \text{ or } 1 \end{cases}$$

3 The application of shortest path problem

3.1 EQUIPMENT REPLACEMENT PROBLEM

There is a machine that can work for 4 years continuously, or can be sold at the end of year and buying a new one instead. As following, we have known the price of new machine at beginning of the year and the sold price of each machine that have different enlistment age. The operation costs and maintenance costs of the new machine is 3000 Yuan; and the cost of operation and maintenance of the machine that use 1 to 3 years in each year are 8000yuan, 15000yuan, 20000yuan respectively.

How make sure the machine optimal update strategy to attain a least cost that sum up purchase and replacement and maintenance within four years?

TABLE 1 The sold price of the machine at every year

Year j	1	2	3	4
Purchase price at beginning year	2.5	2.6	2.8	3.1
reduced price at end year j	2.0	1.6	1.3	1.1

3.2 TO BE TRANSLATED INTO THE SHORTEST PATH PROBLEM

Equipment replacement can be translated into the shortest path problem, as follow figure 1. The point v_i means that “we purchase a new machine at beginning of the year i ”, the point v_5 means the end of the fourth year. We will draw arcs from v_i to v_{i+1}, \dots, v_5 respectively, the arc (v_i, v_j) means purchase new machine at beginning year i , has been used until the beginning of year j , that is the end of year $j-1$. The weight of the arc (v_i, v_j) is total cost including the purchase cost and the maintenance cost from beginning of year i to the end of year $j-1$ and to subtract the residual value of the equipment at the end of year $j-1$. Example, the weight of arc (v_1, v_5) is that the purchase cost 25000 Yuan at first year plus maintenance cost from beginning of first year to the end of fourth year $3000+8000+15000+20000=46000$.

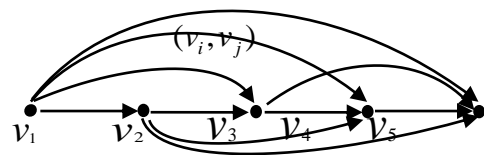


FIGURE 1 Reducing practical problem to Shortest Path Problem

TABLE 2 Weight w_{ij} of all arc (v_i, v_j) /wan Yuan

w_{ij}					
j	1	2	3	4	5
i					
1	-	0.8	2	3.8	6
2	-	-	1.3	2.4	4.1
3	-	-	-	1.8	2.8
4	-	-	-	-	2.3
5	-	-	-	-	-

Moreover, subtract deduced cost 11000 Yuan at end of fourth year, got 60000 Yuan. The weights w_{ij} of all arcs (v_i, v_j) have been shown in tab.2 as following:

So, to find a shortest path from v_1 to v_5 , we can attain the optimal equipment replacement strategy, which is the least cost summing up the purchase and the replacement and the maintenance within four years.

3.3 USING DIJKSTRA ALGORITHM TO SOLVE THE SHORTEST PATH PROBLEM

The Dijkstra algorithm can be used to solve the shortest path as follow:

(1) The point v_1 is labelled as $(0, s)$, there is $I = \{v_1\}$, $J = \{v_2, v_3, v_4, v_5\}$, the set of arc is $\{(v_i, v_j) | v_i \in I, v_j \in J\} = \{(v_1, v_2), (v_1, v_3), (v_1, v_4), (v_1, v_5)\}$, and $s_{12} = l_1 + w_{12} = 0 + 0.8 = 0.8$, $s_{13} = l_1 + w_{13} = 0 + 2 = 2$. Similarly, $s_{14} = 3.8$, $s_{15} = 6$, $\min(s_{12}, s_{13}, s_{14}, s_{15}) = s_{12} = 0.8$, the end point v_2 of the arc (v_1, v_2) is labelled as $(0.8, 1)$.

(2) There is $I = \{v_1, v_2\}$, $J = \{v_3, v_4, v_5\}$, set of arc is $\{(v_1, v_3), (v_1, v_4), (v_1, v_5), (v_2, v_3), (v_2, v_4), (v_2, v_5)\}$, and $s_{23} = l_2 + w_{23} = 0.8 + 1.3 = 2.1$, $s_{24} = l_2 + w_{24} = 0.8 + 2.4 = 3.2$, $s_{25} = l_2 + w_{25} = 0.8 + 4.1 = 4.9$, $\min(s_{13}, s_{14}, s_{15}, s_{23}, s_{24}, s_{25}) = s_{13} = 2$, the end point v_3 of the arc (v_1, v_3) are labelled $(2, 1)$

(3) there is have $I = \{v_1, v_2, v_3\}$, $J = \{v_4, v_5\}$, set of arc is $\{(v_1, v_4), (v_1, v_5), (v_2, v_4), (v_2, v_5), (v_3, v_4), (v_3, v_5)\}$, and $s_{34} = l_3 + w_{34} = 2 + 1.8 = 3.8$, $s_{35} = l_3 + w_{35} = 2 + 2.8 = 4.8$, $\min(s_{14}, s_{15}, s_{24}, s_{25}, s_{34}, s_{35}) = s_{24} = 3.2$, the end point v_4 of the arc (v_2, v_4) are labelled $(3.2, 2)$.

(4) there is have $I = \{v_1, v_2, v_3, v_4\}$, $J = \{v_5\}$, set of arc is $\{(v_1, v_5), (v_2, v_5), (v_3, v_5), (v_4, v_5)\}$, and $s_{45} = l_4 + w_{45} = 3.2 + 2.3 = 5.5$, $\min(s_{15}, s_{25}, s_{35}, s_{45}) = s_{35} = 4.8$, the end point v_5 of the arc (v_3, v_5) are labelled $(4.8, 3)$.

So, the length of shortest path is 4.8 from v_1 to v_5 and the shortest path is $v_1 \rightarrow v_3 \rightarrow v_5$, that is the cost of this layout that we will purchase a new machine at the beginning first year and deal with it at the end of second year along with purchase a new machine at the beginning third year and deal with it at the end of fourth year is 48000 yuan.

3.4 0-1 PROGRAMMING AND LINGO PROGRAM

Introduce 0-1 decision variable x_{ij} [10],

References

[1] WANG Jie-chen, MAO Hai-cheng, YANG De-zhi 2000 United structure of point: Arc for network graph and It's application in GISs shortest path searching *Science of Surveying and Mapping* 29(1) 47- 51
 [2] Arya V, Garg N, Khandekar R, et al. 2004 Local search heuristics for p- median and facility location problems *SIAM Journal on computing* 33(3) 544- 62
 [3] Li Jingjing, Lin Hong 2009 Analysis of Shortest Path Algorithms in Intelligent Traffic System *Computer & Digital Engineering* 37(4) 24-7

$$x_{ij} = \begin{cases} 1, & \text{if the arc}(v_i, v_j) \text{ on the shortest path} \\ 0, & \text{if the arc}(v_i, v_j) \text{ not on the shortest path} \end{cases}$$

Therefore, 0-1 Programming model to solve the shortest path problem is as:

Objective function $\min z = \sum_{(v_i, v_j) \in E} w_{ij} x_{ij}$ (E is a set of all

arc in chart) $s.t. \begin{cases} \sum_{(v_i, v_j) \in E} x_{ij} = \sum_{(v_i, v_j) \in E} x_{ji}, 1 < i < n \\ \sum_{(v_i, v_j) \in E} x_{ij} = 1, \sum_{(v_j, v_i) \in E} x_{ji} = 1 \\ x_{ij} = 0 \text{ or } 1 \end{cases}$

The answer can be found by using LINGO Program as following:

```

model:
sets: years/v1,v2,v3,v4,v5/;
roads(years,years)/
v1,v2 v1,v3 v1,v4 v1,v5 v2,v3 v2,v4 v2,v5 v3,v4
v3,v5 v4,v5/:W,X;
endsets
data:
w=0.8 2 3.8 6 1.3 2.4 4.1 1.8 2.8 2.3;
enddata
N=@SIZE(YEARS);
MIN=@SUM(roads:W*X);
@for(years(i)|i #GT# 1 #AND# i #LT# N:
@SUM(roads(i,j):X(i,j))=@SUM(roads(j,i):X(j,i)));
@SUM(roads(i,j)|i #EQ# 1:X(i,j))=1;
@SUM(roads(i,j)|j #EQ# N:X(i,j))=1;
end
    
```

The answer made by using computer is the same with using Dijkstra algorithm; the computer is more efficient than Dijkstra algorithm. Its drawback is ask for a shortest path from start point to end point are appointed and we must change procedure and recalculate if change start point.

Acknowledgment

This work was supported by Natural Science Foundation of China, under the Grants 71071089; supported by Science and Technology Development Planning of Shandong, under the Grants 2012RKB01457.

[4] WANG Feng, et al. 2006 Application of Dijkstra and Dijkstra-based N-shortest paths algorithm to intelligent transportation systems *Journal of study on the application of computer* 2006(9) 203-8
 [5] CHAI Deng-feng, ZHANG Deng-rong 2002 Algorithm and its application to N shortest paths problem *Journal of Zhejiang University (Engineering Science)* 36(5) 531-4
 [6] NOTO Masato, SAT OHiroaki Sato 2000 A method for the shortest path search by extended Dijkstra algorithm *IEEE International Conference on Systems, Man, and Cybernetics, 2000* 3 2316-20

- [7] ZHENG Yu-xi, LI Jiang, JIANG Qian-lin 2004 New algorithm for finding extended shortest path based on switching function *Journal of Zhejiang University (Engineering Science)* 38(3) 322-4
- [8] Han Botang 2004 *Operation research or management science* Beijing: Higher Education Press 211-30
- [9] Yuan Xinsheng, Shao Dahong, Yu Shilian 2007 *The application of LINGO and Excel on mathematical modelling* Beijing: Science Press 72-80
- [10] Han Zhonggeng 2009 *The methods and application on mathematical modelling (the second edition)* Beijing: Higher Education Press 2009(6) 267-91

Authors	
	<p>Xingbin Ma, born on December 1, 1965, Jinan, China</p> <p>Current position, grades: master of management science, vice professor of Shandong university of science and technology University studies: Management College of Xi'an Jiaotong University Scientific interest: Management science and engineering Publications: 30 papers Experience: My undergraduate major is Mining Engineering in Shandong University of science and technology, and read in the major of the management science and engineering in master grade. I studied the doctoral course in the management college of Xi'an Jiaotong University. Now, I am a vice professor in the college of economic and finance of Shandong university of science and technology. The interest is management science and engineering. By now, more than 30 papers have been published. Email: maxingbin@126.com</p>
	<p>Cuiping Jiang, born in 13 June, 1987, Shandong, China</p> <p>Current position, grades: Master of Management Science, Teaching assistant in Shandong University of Science and Technology University studies: Management Science, in Shandong University of Science and Technology Scientific interest: System complexity of Organization Publications: 3 Papers have been published.</p>

Correlation between austempering parameters and hardness of austempered ductile iron based on artificial neural network

Xuhong Guo*

School of Mechanical and Electric Engineering, Soochow University, Suzhou- Jiangsu, China

Received 1 March 2014, www.tsi.lv

Abstract

Mechanical properties of ADI mainly depend on the austempering parameters, which include austenitizing temperature and time, austempering time and temperature, apart from chemical composition, alloyed elements and casting parameters. In this paper, an investigation has been conducted on the prediction model of mechanical properties of ADI between austenitizing temperature and time, austempering temperature and time as inputs and Vickers hardness of samples after austempering as the outputs based on artificial neural network. There are two types of data of the model: training data and testing data. The former data come from the published literature and 12 experimental data used for network testing. The research results of the model shows that the predicted values approach to the measured data in most of the testing samples and the maximum margin of error between experimental and predicted data is 4.682%.

Keywords: Austempered Ductile Iron, Artificial Neural Network, Austempering Parameters, Mechanical Properties, Prediction Model

1 Introduction

In recent years, austempered ductile iron (ADI) has been attracted much attention due to its exceptional combination of strength, hardness, impact resistance and ductility [1, 2, 3]. The attractive properties are related to its unique microstructure known as ausferrite, which is a mixture of bainitic ferrite (α) and high carbon austenite (γ_{HC}) [4, 5, 6]. Hence, ADI has been applied in many areas, such as automobile, aerospace, agricultural machinery, railroads and so on. S.F. LIU et al [7] indicated that the Mn-Cu alloyed ADI standard sample could reach European standard EN1564-97/EN-C15-1000-5 and could replace 20CrMnTi forged steel for manufacturing the EQ140 helical bevel gears. J.F. Dias [8] investigated reducing the austempering time increased the fatigue life and did not affect the mechanical properties or the rate of fatigue crack propagation, and could used for mechanical parts with stress riser details.

Mechanical properties of ADI mainly depend on the austempering parameters, which include austenitizing temperature and time, austempering time and temperature, apart from chemical composition, alloyed elements and casting parameters. It has been investigated [9] that the morphology of bainite and the fraction of retained austenite depend largely on the austempering temperature, which is one of the most important factors to affect the mechanical properties of ADI. The microstructure of ADI is mixture of finer ferrite and retained austenite when austempered at lower temperatures (230~350°C) and these results in higher tensile strength and hardness but with lower ductility. On

the other hand, it is compound of coarser or feathery ferrite and austenite when austempered at higher temperatures (350~400°C) and this leads to lower tensile strength and hardness but with higher ductility.

It is known that the toughness of a material decreases as its strength increases. To improve the toughness of ADI meanwhile without reducing the strength, researchers presented a new and improved isothermal heat treatment to meet this demand. Susil K. Putatunda [10, 11] proposed a novel and innovative concept of two-step austempering process, which was used to produce ADI with simultaneous high yield strength and fracture toughness. Generally, two-step austempering process acquired higher combination of hardness and ductility compare with the single-step. This is mainly due to that, the former process obtained finer microstructure than the latter [12].

Alloyed elements such as nickel, copper and boron are usually added to strengthen the mechanical properties of ADI. Alloying of ADI with Ni and Mo increased its fracture toughness but decreased its tensile strength and hardness in the heat treated condition [13]. The Cu-alloyed ADI has better impact toughness and fracture toughness than does the unalloyed one because of copper increases the retained austenite content in ADI [14]. With higher boron content, the mechanical properties of CADI such as hardenability and toughness are found to decreases [15].

A significant number of studies have been carried out on the mechanical properties of ADI, but few studies have examined the correlation between the mechanical properties of ADI and influencing factors, such as

*Corresponding author e-mail: huangxing8677@163.com

austempering parameters, the percentage of chemical composition and alloyed element and so on. M.A.Yescas [16] has been estimated of the fraction of retained austenite in ADI as a function of their chemical composition (C, Mn, Si, Ni, Mo, Cu) and the austempering parameters. H.PourAsiabi [17] has been developed a multi-layer perceptron artificial neural network model, which used Mo%, Cu%, austempering time and temperature as inputs and the Vickers hardness of samples after austempering as the output. However, very little is currently available in literature on the mechanical properties of unalloyed ADI with low manganese content.

However, very little information is currently available in literature on the mechanical properties of unalloyed ADI with austempering parameters. Specifically, the non-linear correlation of austenitizing temperature and time, austempering temperature and time as the inputs variable and the hardness of unalloyed ADI as the output are not clearly established. The main purpose of this work is to investigate the mechanical property prediction model between austenitizing temperature and time, austempering temperature and time as inputs and Vickers hardness of samples after austempering as the output based on artificial neural network.

2 Experimental details

The chemical composition of as-cast ductile iron in weight percent is reported in Table 1. The microstructure of as-cast ductile iron, which used in austempering experiment is shown in Fig. 1. The rate of graphite spheroidization is above 90% (level 3) and the diameter of the graphite nodule reach level-6. As-cast ductile iron is a mixture of pearlite, ferrite (about 40%) and graphite nodule.

TABLE 1 Chemical composition of as-cast ductile iron (wt%)

C	Si	Mn	P	S	Mg	Re	Al	Ti
3.65	2.78	0.32	0.026	0.009	0.042	0.025	0.015	0.002

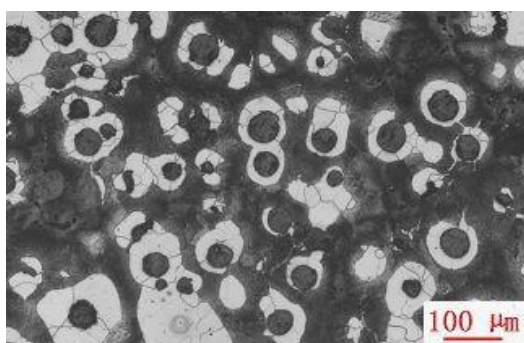


FIGURE 1 Microstructures of as-cast ductile iron

Austempering process was conducted by first austenitizing the as-cast ductile iron by heating to 900°C in an induction furnace and holding the temperature at that level for 2h. Then quickly quenched in a salt bath down to the austempering temperature at 400°C, 350°C,

300°C and 250°C, in which a holding time was chosen as 1h, 2h and 3h, respectively, for each austempering temperature. Finally, the samples were immediately cooled in air to room temperature. 12 samples with different heat treating cycles were obtained and the detailed austempering parameters used for this study is shown in Fig. 2.

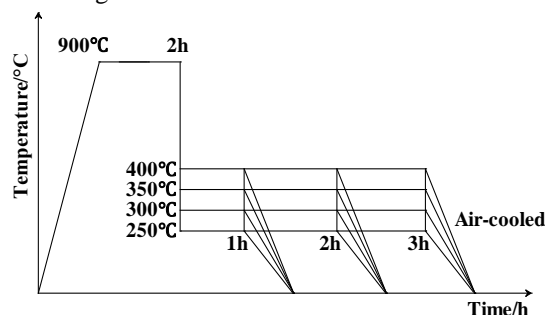


FIGURE 2 Austempering parameters schematic

The hardness of the samples after the austempering process was evaluated using by HB-3000 Brinell hardness tester, under 3000 kgf loading according to GB/T 231.1-2009. Three times were taken for each specimen and then the results were averaged.

3 Artificial neural network

Artificial neural network (ANN) is a mathematical model that can learn and generalize the things learned, especially suitable for non-linear properties from input to output [18]. Therefore, there is growing interest for development of intelligent dynamic systems based on practical data. Back propagation (BP), which is one of the most famous training algorithms for multilayer perceptions, is a gradient descent technique to minimise the error for particular training pattern [19].

The neuron is the basic part of artificial neural network. In general, each input is multiplied by its related weight and add together, plus the threshold value and then cross through activation functions to produce the outputs [20]. Fig. 3 shows the data processing in a neural network cell and the output of the neuron as Eq.1.

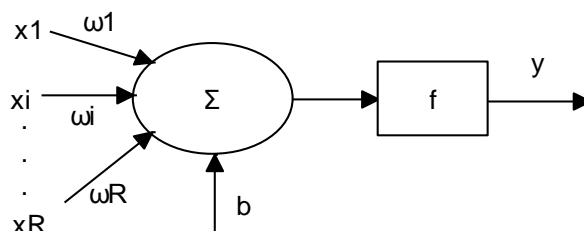


FIGURE 3 The data processing in a neural network cell

$$y = f\left(\sum_{i=1}^R x_i w_i + b\right), \tag{1}$$

where x_i is normalized input variable, w_i is the weight of that variable, b is the threshold value, R is the number of the input variables, f is the activation

function of the cell and y is the network output of the cell.

Generally, the activation function is one of the most parts in neural network modeling, and the function is continues and derivative. Mean Square Error (MSE) as statistical criterion are utilized to evaluate results accuracy according to following equation:

$$MSE = \frac{1}{n} \sum_{j=1}^n (t_j - o_j)^2, \tag{2}$$

where t_j and o_j is target and predicted values respectively, n is the number of the network outputs [21, 22].

A typical BP-ANN model include three layers: input layer, one hidden layer and output layer. Full connection occurs among neurons belonging to each layer, while no connections exist among neurons belonging to the same layer. The weight of each layers can be adjust through network training. Generally, there are two types of data of the model: training data and testing data. 121 groups of data [1, 8, 9, 11, 12, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32], which come from the published document, are designed for model training in this research. 12 experimental data used for network testing. To remove the order of magnitude from different input and output parameters, the data are disposed by using normalized method, which is set data between [0 1] (Eq. 3). After network testing is done, the outputs were reverse normalized (Eq. 4).

$$x_k = \frac{x_k - x_{\min}}{x_{\max} - x_{\min}}, \tag{3}$$

$$x_k = x_{\min} + (x_{\max} - x_{\min})x_k, \tag{4}$$

where x_{\min} is the minimum value of each column, x_{\max} is the maximum value of each column.

The schematic structure of this designed neural network is shown in Fig. 4. A BP-ANN model was used with austenitizing temperature and time, austempering temperature and time as inputs and the Rockwell hardness of ADI as the output of the model. The number of hidden layer selected 11 neurons on the basis of the previous empiric formulas. The activation function of the hidden layer uses a hyperbolic tangent sigmoid (tansig), while the output layer uses a linear (purelin) transfer function.

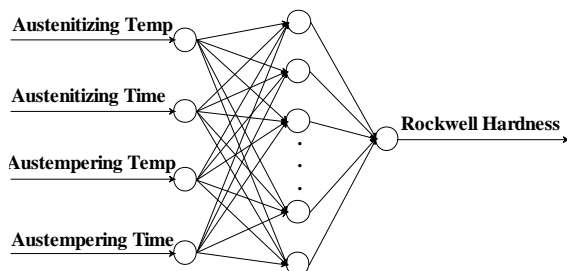


FIGURE 4 The schematic structure of this designed neural network

4 Results and discussion

The variation of hardness as a function of austempering temperature for austempering time 60, 120 and 180 min is shown in Fig. 5. It can be found that increase austempering temperature leads to a linear decrease in hardness at all times. It is mainly due to that the microstructure of ADI is mixture of finer ferrite and retained austenite when austempered under 350 °C, and this results in higher tensile strength and hardness but lower ductility. On the other hand, it has a compound of coarser ferrite and austenite when austempered over 350 °C, and this reduces to the yield and hardness strengths but with higher ductility.

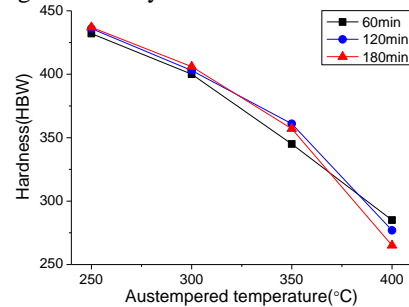


FIGURE 5 Influence of austempering temperatures on hardness at different austempering time

It is clearly observed that the hardness of ADI slightly change at different austempering time under the same temperature. Untransformed supercooling austenite, which under compressive stress gradually lose change activity and then transform into retained austenite when the austempering time more than 60 minute. On the other hand, the rate of bainite transformation increasingly slow and basically finished as the time over 60 minute [33]. It is apparently state that the austempering temperature has greatly impact on the hardness of ADI and the time has little or no effect on it.

The mean square error (MSE) value during network training is shown in Fig.6. It is evidenced that the training step have significant effect on the error of the network, in which MSE firstly decreases with the training step ranging from 0 to 300 epochs, and then it keep constant with further increasing training step. It is obviously found, that the order of magnitudes of the error approached to 10^{-3} , which indicates the desired network error achieved after training.

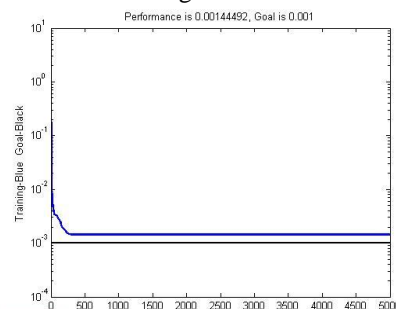


FIGURE 6 The variation of mean squared error (MSE) with number of epochs

The detailed mathematical equation of the neural network model is available after the network training and the formula of the model is as follows (Eq. 5):

$$T = f_2[w_2 f_1(w_1 P + b_1) + b_2], \tag{5}$$

where P is input variable, T is output variable, f_1 is the activation function from input layer to hidden layer, f_2 is the activation function from hidden layer to output layer, w_1 is the weight matrix form input layer to hidden layer, w_2 is the weight matrix form hidden layer to output layer, b_1 is the threshold matrix from input layer to hidden layer, b_2 is the threshold matrix from hidden layer to output layer. The detailed values of the parameters are as follows:

$$w_1 = \begin{bmatrix} 28.38 & -42.34 & -1.08 & 2.21 \\ 1.22 & 2.06 & -4.03 & 0.06 \\ 63.16 & 100.30 & -7.83 & 27.19 \\ 1.24 & 2.14 & -3.46 & 0.11 \\ 0.44 & 1.79 & -5.69 & -0.11 \\ -2.97 & -6.48 & 7.25 & -0.99 \\ 7.04 & 24.68 & 1.09 & -2.23 \\ 2.08 & -5.33 & 0.56 & -105.96 \\ 19.25 & 33.10 & -23.00 & 0.88 \\ -0.89 & -1.86 & 5.00 & 0.05 \\ 2.73 & 5.80 & -6.47 & 0.84 \end{bmatrix}, \quad b_1 = \begin{bmatrix} -12.71 \\ 0.14 \\ -73.93 \\ 0.00 \\ 1.25 \\ -1.34 \\ -5.01 \\ 33.48 \\ -24.91 \\ -0.70 \\ 1.12 \end{bmatrix},$$

$$w_2 = [169.95 \quad -85.10 \quad -0.07 \quad 46.94 \quad -27.58 \quad -8.74 \quad 170.03 \quad -0.08 \quad 0.92 \quad -68.54 \quad -11.18]$$

$$b_2 = [1.05], \quad f_1(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}, \quad f_2(x) = x.$$

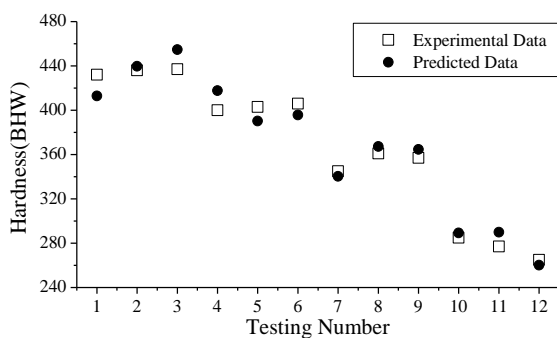


FIGURE 7 Comparison of experimental data and predicted data for testing number after training

After network training, 12 experimental data used for network testing and get the predicted results. Then the outputs of the samples converted to the value of Brinell hardness by the way of reverse normalized. The testing results after transformation are shown in Fig. 7. The horizontal axle shows testing number, while the vertical axle illustrates the hardness of ADI of measured value and predicted value. In the previous diagram, the hollow squares and solid points represent the experimental data and prediction data, respectively. It can be obviously see that the predicted values approach to the measured data in most of the testing samples and indicated that the desired accuracy of the network model can be achieved. Therefore, the model based on artificial neural network can be applied in actual production with good accuracy. Table 2 are listed the absolute error of experimental and predicted value of the testing samples. It can be clearly see from the table that the maximum margin of error is 4.682%.

TABLE 2 Absolute errors of experimental data and predicted data for testing samples

	Experimental data (HBW)	Predicted data (HBW)	Absolute errors (%)
1	432	412.905	4.420
2	436	439.636	-0.834
3	437	454.745	-4.061
4	400	417.652	-4.413
5	403	390.215	3.173
6	406	395.656	2.548
7	345	340.374	1.341
8	361	367.252	-1.732
9	357	364.698	-2.156
10	285	289.265	-1.497
11	277	289.970	-4.682
12	265	260.247	1.794

5 Conclusions

From the study, the following conclusions can be drawn:

1. Increasing austempering temperature leads to a linear decreasing in hardness at the austempering times. The hardness of ADI slightly changed at different austempering time under the same temperature.
2. The detailed mathematical equation of the neural network model is available after the network training and the formula of the model is $T = f_2[w_2 f_1(w_1 P + b_1) + b_2]$. The detailed values of the parameters refer to the paper.
3. The predicted values approach to the measured data in most of the testing samples and indicated that the desired accuracy of the network model can be achieved. Therefore, the model based on artificial neural network can be applied in actual production with good accuracy.
4. The maximum margin of error between experimental and predicted data is 4.

References

- [1] Han J M, Zou Q, Barber G C, Nasir T, Northwood D O, Sun X C, Seaton P 2012 Study of the effects of austempering temperature and time on scuffing behavior of austempered Ni–Mo–Cu ductile iron *Wear* **290-291** 99-105
- [2] Eric O, Sidjanin L, Miskovic Z, Zec S, Jovanovic M T 2004 Microstructure and toughness of Cu Ni Mo austempered ductile iron *Materials Letters* **58** 2707-11
- [3] Uma T R, Simha J B, Murthy K N 2011 Influence of nickel on mechanical and slurry erosive wear behaviour of permanent moulded toughened austempered ductile iron *Wear* **271** 1378-84
- [4] Magalhaes L, Seabra J 1998 Wear and scuffing of austempered ductile iron gears *Wear* **215** 237-346
- [5] Greno G L, Otegui J L, Boeri R E 1999 Mechanisms of fatigue crack growth in Austempered Ductile Iron *International Journal of Fatigue* **21** 35-43
- [6] Sohi M H, Karshenas G, Boutorabi S M A 2004 Electron beam surface melting of as cast and austempered ductile irons *Journal of Materials Processing Technology* **153-154** 199-202
- [7] Liu S F, Chen Y, Chen X, Miao H M 2012 Microstructures and mechanical properties of helical bevel gears made by Mn-Cu alloyed austempered ductile iron *Journal of Iron and Steel Research* **19(2)** 36-42
- [8] Dias J F, Ribeiro G O, Carmo D J, Vilela J J 2012 The effect of reducing the austempering time on the fatigue properties of austempered ductile iron *Materials Science & Engineering A* **556** 408-13
- [9] Liu S F 1999 The influence of heat treatment on the microstructures and properties of unalloyed austempered ductile iron *China Foundry Machinery & Technology* **3** 17-8
- [10] Putatunda S K 2001 Development of austempered ductile cast iron (ADI) with simultaneous high yield strength and fracture toughness by a novel two-step austempering process *Materials Science and Engineering A* **315** 70-80
- [11] Yang J H, Putatunda S K 2004 Improvement in strength and toughness of austempered ductile cast iron by a novel two-step austempering process *Materials and Design* **25** 19-230
- [12] Li X F, Yu J, Liu L J, Liu H M, Zu F Q 2008 Effect of two-step austempering process on microstructure and properties of austempered ductile iron *Transactions of Materials and Heat Treatment* **29(2)** 82-5
- [13] Elsayed A H, Megahed M M, Sadek A A, Abouelela K M 2009 Fracture toughness characterization of austempered ductile iron produced using both conventional and two-step austempering processes *Materials and Design* **30** 1866-77
- [14] Hsu C H, Lin K T 2011 A study on microstructure and toughness of copper alloyed and austempered ductile irons *Materials Science and Engineering A* **528** 5706-12
- [15] Peng Y C, Jin H J, Liu J H, Li G L 2011 Effect of boron on the microstructure and mechanical properties of carbidic austempered ductile iron *Materials Science and Engineering A* **529** 321-5
- [16] Yescas M A, Bhadeshia H K D H, MacKay D J 2001 Estimation of the amount of retained austenite in austempered ductile irons using neural networks *Materials Science and Engineering A* **311** 162-73
- [17] PourAsiabi H, AmirZadeh Z, BabaZadeh M 2012 Development a multi-layer perceptron artificial neural network model to estimate the Vickers hardness of Mn-Ni-Cu-Mo austempered ductile iron *Materials and Design* **35** 782-9
- [18] Ghaisari J, Jannesari H, Vatani M 2012 Artificial neural network predictors for mechanical properties of cold rolling products *Advances in Engineering Software* **45** 91-9
- [19] Genel K 2004 Application of artificial neural network for predicting strain-life fatigue properties of steels on the basis of tensile tests *International Journal of Fatigue* **26** 1027-35
- [20] Gencil O, Kocabas F, Gok M S, Koksal F 2011 Comparison of artificial neural networks and general linear model approaches for the analysis of abrasive wear of concrete *Construction and Building Materials* **25** 3486-94
- [21] Calcaterra S, Campana G, Tomesani L 2000 Prediction of mechanical properties in spheroidal cast iron by neural networks *Journal of Materials Processing Technology* **104** 74-80
- [22] Davim J P, Gaitonde V N, Karnik S R 2008 Investigations into the effect of cutting conditions on surface roughness in turning of free machining steel by ANN models *Journal of Materials Processing Technology* **205** 16-23
- [23] Salman S, Findik F, Topuz P 2007 Effects of various austempering temperatures on fatigue properties in ductile iron *Materials and Design* **28** 2210-4
- [24] Kowalski A, Tybulczuk J 2002 Influence of heat treatment parameters on the structure and property of Ni-Cu alloyed ADI *Foundry* **151(11)** 698-700
- [25] Peng Y C, Jin H J, Liu J H, Li G L 2012 Influence of cooling rate on the microstructure and properties of a new wear resistant carbidic austempered ductile iron (CADI) *Materials Characterization* **72** 53-8
- [26] Lin C K, Lai P K, Shih T S 1996 Influence of microstructure on the fatigue properties of austempered ductile irons - High-cycle fatigue *International Journal of Fatigue* **18(5)** 297-307
- [27] Wei D Q 2005 Effect of alloying elements on structure and mechanical properties of bainite ductile Iron in the step austempering in machine oil at room temperature *Materials for Mechanical Engineering* **29(12)** 29-32
- [28] Sohi M H, Ahmadabadi M N, Vahdat A B 2004 The role of austempering parameters on the structure and mechanical properties of heavy section ADI *Journal of Materials Processing Technology* **153-154** 203-8
- [29] Wang W, Guo X H, Wang C H, Wang H 2010 Research on the influence of heat treatment parameter on the mechanical property of austempered ductile iron *International Conference on Mechanic Automation and Control Engineering*, 26-28 June 2010 5918-21
- [30] Liu G J, Wang T X 2001 Effect of isothermal quenching temperature on mechanical properties of austempered ductile iron *Foundry* **50(9)** 567-9
- [31] Podgornik B, Vizintin J, Thorbjornsson I, Johannesson B, Thorgrimsson J T, Celis M M, Valle N 2012 Improvement of ductile iron wear resistance through local surface reinforcement *Wear* **274-275** 267-73
- [32] Kim Y J, Shin H, Park H, Lim J D 2008 Investigation into mechanical properties of austempered ductile cast iron (ADI) in accordance with austempering temperature *Materials Letters* **62** 357-60
- [33] Zhu J W 2002 The influences of austempering process on the mechanical properties of nodular graphite cast iron *Journal of Jilin Institute of chemical Technology* **19(1)** 51-3

Authors



Guo Xuhong, born on June 22, 1963, Gansu, China

Current position, grades: professor in Soochow University

University studies: Department of Mechanical Engineering in Lanzhou university of technology

Scientific interest: Cutting mechanism, precision finishing, Process parameters optimization.

Publications: 2 Patents, 40 Papers

Experience: 1998 - present, working in school of mechanical and electric engineering, Soochow University. 1997 - 1998, further studying in Sunny electronic (Nagano, Japan) Co. 1996 - 1997, teaching in Department of Mechanical Engineering of Soochow silk engineering college. 1990 - 1995, teaching in process department of Hubei University of Automotive technology. 1987 - 1990, Studying for a master's degree in Metal processing laboratory of South China University of Technology. 1983-1987, working in Hubei University of Automotive technology. 1979 - 1983, studying for a bachelor's degree, majoring in manufacturing technology and equipment in Lanzhou university of technology.

Firefly algorithm for training the radial basis function network in ultrasonic supraspinatus image classification

Chih-Feng Chao, Ming-Huwi Horng*

Department of Computer Science and Information Engineering, National PingTung Institute of Commerce, Pingtung, Taiwan

Received 1 March 2014, www.tsi.lv

Abstract

The physicians observed the echo-texture and the shape changes of supraspinatus to decide the severity of rotator cuff disease in the clinical standard ultrasound examination. It is not reliable because the accuracy of visual observation depends on the experience of physicians. This article proposes a new algorithm called **Firefly RBF** network to training the radial basis function neural network by applying the firefly algorithm for classifying the different supraspinatus disease groups that are normal, tendon inflammation, calcific tendonitis and tendon tears of the ultrasound supraspinatus images based on the texture analysis technology. The texture features are generated from four methods those are the grey-level co-occurrence matrix, the texture spectrum, the fractal dimension and the texture feature coding method to analyse the tissue characteristic of supraspinatus. The F-score measurement are used to select powerful features those are generated from the four texture analysis methods for comparison in the training stage, meanwhile, the proposed Firefly RBF network is used to discriminate test images into one of the four disease groups in the classification stage. Experimental results showed that the percentage of correct classification was more than 93.7% that is superior to other methods in the classification of ultrasonic supraspinatus images.

Keywords: Radial basis function network, Firefly algorithm, Ultrasonic supraspinatus image, Texture analysis

1 Introduction

The injuries of the supraspinatus of the rotator cuff muscles, such as tendon inflammation, calcific tendonitis and tendon tears always cause in the human shoulder pained. For this reason, clinical physicians routinely observe the disease symptoms of supraspinatus using the ultrasound imaging examination. Recently, the Neer's classification system [1] has become a popular method that separates different diseases of supraspinatus into three stages in clinical diagnosis. The inflammation manifestations, such as edema or hemorrhage, usually exhibit in supraspinatus of Stage 1. The supraspinatus of Stage 2 is more serious and considered irreversible. Fibrosis and calcification always appear. Stage 3 generally involves a tendon rupture or tear. If the tendon rupture arises in the supraspinatus, it may require further repair. In clinical, ultrasonic examination has proved to be useful diagnostic tool in patients with shoulder pain and/or limited range of motion. Iagnocco et al. [2] explained how to use the ultrasonography to the careful qualitative assessment of a wide range of changes of different anatomic structures of rotator cuff tendons such as tendonitis, tendon tears and calcific deposits. Arsian et al. [3] used the 7.5-MHz linear-array transducer to grab images of patients with physical examination suggestive rotator cuff injury under longitudinal view. This study demonstrated that the bursa fluid and biceps effusion were high correlated with symptoms of rotator cuff injury. Chiou et al. [4] proposed the diagnosis criteria for

classification of full/partial supraspinatus tears of patients with shoulder pain. The sensitivity and specificity of experiments in applying the diagnosis criterion are 0.98 and 0.87. Marcello et al. [5] used criterion that are (a) one or more cuff tendon(s) was not visible, (b) focal non visibility of one the tendons, and (c) defect of well defined discontinuity of the tendons to diagnose of the rotator cuff tear. AI-Shawi et al. [6] proposed a study on the detection of full thickness rotator cuff tears using the ultrasound. The accuracy of the detection of large and massive tears, moderate tears and small tears are 96.5%, 88.8% and 91.6%. Up to now, most studies still used a subjective evaluation in terms of a visual inspection of the images or measurements of the muscle disorders, however, the accuracy of diagnosis by subjective evaluation always depends on the experiences of clinical physician, and therefore the quantitative evaluation is required.

The quantitative evaluation of tissue characteristics of supraspinatus is an important issue in clinical diagnosis that measured the distribution of the grey scales of the pixel in the areas of supraspinatus in the ultrasonic image to describe tissue characteristics. Nielsen et al. [7] proposed a method, a so-called "blob analysis", related to the higher order grey-level statistics of the image for quantitative ultrasound tissue characterization to measure the discrepancy of supraspinatus muscle and the right vastus lateralis muscle. They found that the first order histogram features are effective in classification the shoulder and thigh muscles. It is deficient to discuss the

* *Corresponding author* e-mail: mh.horng@msa.hinet.net

classification of other impingement syndromes such as inflammation and tendon calcification of the rotator cuff. In our past studies, we proposed a comparative article of using the various multi-class support vector machines to classify ultrasonic supraspinatus images into the four disease groups [8]. The results reveal that the fuzzy support vector machine is the most powerful for classifying the ultrasonic supraspinatus images. The classification rate by using the fuzzy support vector machine can achieve 90%. Furthermore, the comparisons of the maximum likelihood, error-correcting output code, fuzzy SVM and the multi-class radial basis function network methods had discussed in classification of supraspinatus images [9]. It concluded the radial basis function network has better performance than the fuzzy SVM method, however, its correct classification rate is only 92.6% that is not able to satisfy in the requirement of clinical diagnosis.

The firefly algorithm is a new swarm-based approach for optimization, in which the search algorithm is inspired by social behaviour of fireflies and the phenomenon of bioluminescent communication. There are two important issues in the firefly algorithm that are the variation of light intensity and formulation of attractiveness. Yang [10] that simplifies the attractiveness of a firefly is determined by its brightness which in turn is associated with the encoded objective function. The attractiveness is proportional to their brightness. Furthermore, every member x_i of the firefly swarm is characterized by its bright I_i , which can be directly expressed as an inverse of a cost function for a minimization problem. Lukasik & Zak [11] applied the firefly algorithm for continuous constrained optimization. Yang [12] compared the firefly algorithm with the other meta-heuristic algorithms such as genetic and particle swarm optimization algorithms in the multimodal optimization. These works had the same conclusions that the algorithm applied the proposed firefly algorithm is superior to the other existing meta-heuristic algorithms. In this paper, we developed a new RBF neural network classifier called the **Firefly RBF** (i.e., the **Firefly** algorithm to training the **radial basis function**) neural network that is used to classify the four disease groups of the ultrasonic supraspinatus images. In experiments, the results of Firefly RBF network are compared with the above-mentioned other four methods. Experimental results revealed that the proposed Firefly RBF neural network is superior to those of the methods. Its correct classification rate can reach 93.75% that is very close to the diagnostic reports of clinical physicians with the rich experiences. The remaining of this paper is organized as follows. The feature extraction and selection are summarized in Section 2. Section 3 describes how to apply the firefly algorithm for training the radial basis function network in the classification of the ultrasound supraspinatus images. Some experimental results of classifying supraspinatus images are discussed in Section 4, and finally, the conclusions are presented in the Section 5.

2 Feature extraction and selection

The texture-based measurement had been applied to ultrasound images for diagnosing diseases over a decade. Horng et al. [9] compared texture descriptors that include the grey-level co-occurrence matrix (GLCM), the fractal dimension (FD), the texture spectrum (TS), the statistical feature matrix, and the texture feature coding method (TFCM) in the classification of the chronic liver diseases. In this work, it found that features generated from the grey-level co-occurrence matrix and texture feature coding method were effective for classifying the three liver states that are normal liver, hepatitis and cirrhosis. Another work [6] used the grey-level co-occurrence matrix (GLCM), the fractal dimension (FD), the texture spectrum (TS), and the texture feature coding method (TFCM) to extract features for classifying the ultrasonic supraspinatus images based on the characteristics of echo-texture. In summary, each region R of the supraspinatus image can extract 80 features that are generated from the above-mentioned four texture analysis methods. In these features, 56 features were generated from GLCM, eight features from TS, two from FD and the others from TFCM.

Feature selection has become the focus of much research in the area of application for which datasets with tens or hundreds of thousands of features are available. The universal algorithms of feature selection are often divided into two groups that are wrapper and filter approaches [13]. The wrapper model consists of two phase that are feature subset selection phase, and learning and testing phase. The feature subset selection selects the best subset using a classifier's accuracy as a criterion. The learning and testing phase provides a classifier that is learned from the training data with the best feature subset and is tested on test data. Filter approach is built on the intrinsic properties of the data, not on a basis of particular classifier. A filter model of feature selection also consists of two phases, that are one is feature selection that uses some measures such as F-score measurement or mutual information as search criteria, another phase is that the classifier is learned on the training data with the selected features. The F-score measures had been reliable than the mutual information method [9], so in this paper F-score measure is adopted as the search criteria to search powerful features from the those extracted from above-mentioned four texture analysis methods.

The feature ranking approaches use a principal or auxiliary mechanism to select the best feature set for classification. Because of their simplicity and scalability, the approaches have been widely applied. F-score ranking method is one of the feature ranking approaches. The larger the F-score measure of feature is, the more likely this feature is more discriminative. Given training features x_k , $k=1, 2, \dots, n$ if the number of positive and negative instances are n_+ and n_- , respectively, then the F-score of the i^{th} feature is defined as follows:

$$F_score(i) = \frac{(\bar{x}_i^{(+)} - \bar{x}_i)^2 + (\bar{x}_i^{(-)} - \bar{x}_i)^2}{\frac{1}{n_+ - 1} \sum_{k=1}^{n_+} (x_{k,i}^{(+)} - \bar{x}_i^{(+)})^2 + \frac{1}{n_- - 1} \sum_{k=1}^{n_-} (x_{k,i}^{(-)} - \bar{x}_i^{(-)})^2}. \quad (1)$$

In Eq. (1), the \bar{x}_i , $\bar{x}_i^{(+)}$ and $\bar{x}_i^{(-)}$ are the average of the i^{th} feature of the whole, positive and negative data sets, respectively; $\bar{x}_i^{(+)}$ is the i^{th} feature of the k^{th} positive instance, and $x_{k,i}^{(-)}$ is the i^{th} feature of the k^{th} negative instance. In experiments, we calculate the average of F-score measure that is obtained by computing between two different groups in order to analyse the discrimination of each texture feature.

3 Training RBF network by using the firefly algorithm

3.1 RADIAL BASIS FUNCTION NETWORK

The radial basis function network is a popular type of network that is very useful for pattern classification. A radial basis function (RBF) network is considered a special three-layered network shown in Fig 1.

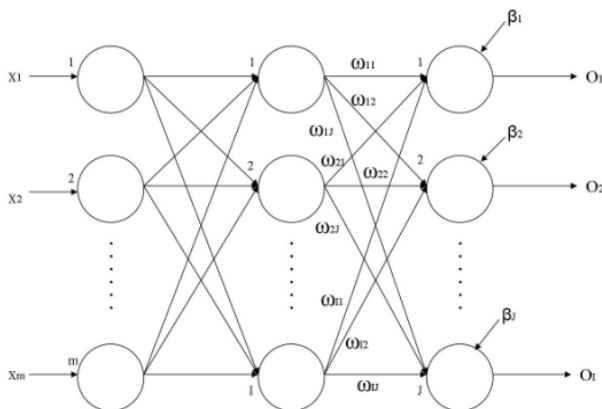


FIGURE 1 The structure of radial basis function network

The input nodes pass the input values $x=(x_1, x_2, \dots, x_m)$ to the internal nodes that construct the hidden layer. Each unit of hidden layer implements a specific activation function called radial basis function. The nonlinear responses of hidden nodes are weighted in order to calculate the final outputs of network in the output layer. The input layer of this network has m units for m dimensional input vectors. The input units are fully connected to I hidden layer units, which are in turn fully connected to the J output layer units, where J is the number of output layer. Each neuron of the hidden layer has a parameter mean vector called centre. Figure 1 shows the detailed structure of an RBF network. Each input data x with m dimensions, $x=(x_1, x_2, \dots, x_m)$, are located in the input layer, which broadcast to hidden layer. The hidden layer has I neurons and each neuron compute the distance between the centres and the inputs. Each activation function of the neuron in hidden layer is

chosen to be Gaussians and is characterized by their mean vectors c_i and its spread parameter α_i ($i=1,2, \dots, I$). That is, the activation function $\phi(x)$ of the i^{th} hidden unit for an input vector x is given by:

$$\phi_i(x) = \exp[-\alpha_i \cdot \|x - c_i\|^2]. \quad (2)$$

The ϕ_i affects the smoothness of the mapping, thus, the output value of the neuron j of output layer \bar{y}_j for training sample x , are given by $o(x)$ in Eq. (3).

$$o(x_i) = (\bar{o}_1, \bar{o}_2, \dots, \bar{o}_j) \quad (3)$$

$$\bar{o}_j = \sum_{h=1}^I w_{hj} \phi_h(x_i) + \beta_j$$

The weights, w_{ij} ($i=1,2, \dots, I, j=1,2, \dots, J$), is the i^{th} node of output of hidden layer that transmitted to j^{th} node of the output layer, and β_j is the bias parameter of the j^{th} node of output layer determined by the RBF network training procedure. In practice, the training procedure of RBF is to find the adequate parameters w_{ij} , α_i , β_i and c_i such that the error metrics such as the mean square error (MSE) is minimum.

$$MSE(w, \alpha, \beta, c) = \frac{1}{N} \sum_{k=1}^N \|d(x_k) - o(x_k)\|^2, \quad (4)$$

where the $d(x_i)$ and $o(x_i)$ are the desired output vector and actual output vector for training sample x_i . In (4), the N is the number of the training samples. The traditional implementation of RBF network uses the gradient descent algorithm to construct the structure of RBF network. In this paper, the gradient descent algorithm is called the Gradient-RBF neural network.

3.2 FIREFLY ALGORITHM

Firefly algorithm (FA) was developed by Xin-She Yang at Cambridge University in 2008. In the firefly algorithm, there are three idealized rules: (1) all fireflies are unisex so that one firefly will be attracted to other fireflies regardless of their sex; (2) Attractiveness is proportional to their brightness, thus for any two flashing fireflies, the less brighter one will move towards the brighter one. If there is no brighter one than a particular firefly, it will move randomly. As firefly attractiveness one should select any monotonically decreasing function of the distance $r_{i,j} = d(x_j, x_i)$ to the chosen j^{th} firefly, e.g. the exponential function.

$$r_{i,j} = \|x_i - x_j\|, \quad (5)$$

$$\beta \leftarrow \beta_0 e^{-\gamma r_{i,j}}, \quad (6)$$

where the β_0 is the attractiveness at $r_{i,j} = 0$ and γ is the light absorption coefficient at the source. The movement

of a firefly i is attracted to another more attractive firefly j is determined by

$$x_{i,k} \leftarrow (1 - \beta)x_{i,k} + \beta x_{j,k} + u_{i,k}, \tag{7}$$

$$u_{i,k} = \sigma(\text{rand1} - \frac{1}{2}). \tag{8}$$

The particular firefly x_i with maximum fitness will move randomly according to the following equation.

$$x_{i^{\max},k} \leftarrow x_{i^{\max},k} + u_{i^{\max},k}, \text{ for } k=1,2,\dots,c, \tag{9}$$

$$u_{i^{\max},k} = \sigma(\text{rand2} - \frac{1}{2}). \tag{10}$$

When rand1 , rand2 are random vector whose each element obtained from the uniform distribution range from 0 to 1; (3). The brightness of a firefly is affected or determined by the landscape of the fitness function. For maximization problem, the brightness I of a firefly at a particular location x can be chosen as $I(x)$ that is proportional to the value of the fitness function.

3.3 FIREFLY RBF NEURAL NETWORK

In the proposed algorithm each individual of the fireflies is composed of the parameters of weights (w), spread parameters (α), centre vector (c) and the bias parameters (β) of network structure of Fig. 1. The mean vector c_i of the i -th neuron of hidden layers is defined by $c_i = (c_{i1}, c_{i2}, \dots, c_{im})$, therefore, the parametric vector of position y_{f_i} of each glow-worm f_i with $IJ + I + mI + J$ parameters is expressed as:

$$y_{f_i} = (w_{11}^i, w_{12}^i, \dots, w_{IJ}^i, \alpha_1^i, \alpha_2^i, \dots, \alpha_I^i, c_{11}^i, c_{12}^i, \dots, c_{1m}^i, \dots, c_{I1}^i, c_{I2}^i, \dots, c_{Im}^i, \beta_1^i, \beta_2^i, \dots, \beta_m^i, \dots, \beta_J^i)$$

In fact, each of fireflies can represent into a parameter vector that can construct a specific RBF network for classification. In our proposed Firefly RBF neural network, the applied fitness function is given in Eq. (11), that is to say, the algorithm of Firefly RBF network is to select the optimal vectors f_{best} of firefly of specific trained Firefly RBF network can maximize the fitness function defined in the Eq. (11).

$$J(y_{f_i}) = \frac{1}{1 + MSE} = \frac{1}{1 + \frac{1}{N} \sum_{k=1}^N \|d(x_k) - o(x_k)\|^2}, \tag{11}$$

where $d(x_k)$ and $o(x_k)$ are denoted to the desired output vector and actual output vector for training sample x_k of RBF network designed by parametric vector y_{f_i} . The N is the number of the training samples.

The steps of the proposed algorithm are described as following in detail.

Step 1. (Generate the initial solutions and given parameters)

In this step, the initial population of m solutions are generating with dimension $IJ + I + mI + J$, denoted by the matrix D .

$$D = [f_1, f_2, \dots, f_m], \tag{12}$$

$$y_{f_i} = (w_{11}^i, w_{12}^i, \dots, w_{IJ}^i, \alpha_1^i, \alpha_2^i, \dots, \alpha_I^i, c_{11}^i, c_{12}^i, \dots, c_{1m}^i, \dots, c_{I1}^i, c_{I2}^i, \dots, c_{Im}^i, \beta_1^i, \beta_2^i, \dots, \beta_m^i, \dots, \beta_J^i) \tag{13}$$

where the values of weights (w) and centres (c) are assigned between -1 and 1, and the values of the spread and bias parameters α and β range from 0 to 1. Furthermore, the step will assign the parameters of firefly algorithm, that are σ , β_0 , the maximum cycle number (MCL) and γ . Let number of cycle l to be 0.

Step 2. Firefly movement

In step 2, each solution (firefly) f_i computes its fitness value $J(y_{f_i})$ as the corresponding the brightness of firefly. For each solution f_i , this step randomly selects another one solution f_j with the more bright and then moves toward to f_j based on the following equations.

$$r_{i,j} = \|y_{f_i} - y_{f_j}\| = \sqrt{\sum_{k=1}^{IJ+I+mI+J} (y_{f_{i,k}} - y_{f_{j,k}})^2}, \tag{14}$$

where the k is an index of the component of the parametric vector form 1 to

$$\beta = \beta_0 e^{-\gamma r_{i,j}}, \tag{15}$$

$$y_{f_{i,k}} = (1 - \beta)y_{f_{i,k}} + \beta y_{f_{j,k}} + u_{j,k}, \tag{16}$$

$$k = 1, 2, \dots, IJ + I + mI + J$$

where $u_{j,k} \sim U(0,1)$ is a randomly number ranged form 0 to 1 and the $y_{f_{i,k}}$ is the k^{th} element of the solution y_{f_i} .

Step 3. (Select the current best solution)

The step 3 selects the best one from the all solutions and defines as $y_{f_{best}}$, that is,

$$i^{\max} = \arg \max_i y_{f_i}; \tag{17}$$

$$y_{f_{best}} = y_{f_{i^{\max}}};$$

Step 4. (Check the termination criterion)

If the cycle number l is equal to the MCL then the algorithm is finished and output the best solution $y_{f_{best}}$. Otherwise, l increases by one and randomly walks the best solution $y_{f_{best}}$ then go to Step 2. The best solution $y_{f_{best}}$ will randomly walk its position based the following equation.

$$y_{f_{best}} = y_{f_{best}} + u_{i^{\max},k}, \quad k = 1, 2, \dots, IJ + I + mI + J, \tag{18}$$

where $u_{i^{\max},k} \sim U(0,1)$ is a random number.

4 Experimental results and discussion

4.1 IMAGE ACQUISITION AND SYSTEM EQUIPMENT

All the ultrasonic images used were recorded from 2004 to 2007, and the ages of patients ranged from 30 to 65 years. In all, 120 shoulders in 120 patients with shoulder pain who had undergone preoperative and subsequent arthroscopy were identified. The arthroscopy diagnosis was a thickness tear in 30, a tendon inflammation in 30, a calcific tendon in 30 and the normal in 30. A longitudinal view of an ultrasound image of each shoulder was acquired using an HDI Ultramark 5000 Ultrasound system (ATL Ultrasound, CA, USA) fitted with a 5.0 MHz dynamic focusing transducer (C5-40 5.0 MHz Curved Linear Array, ATL Ultrasound, CA) from National Cheng Kung University Hospital in Taiwan based on current the clinical setting for ultrasound examination. The captured images were digitized into 256×256 pixels with 256 grey levels via a frame grabber and the stored on a disk. Figure 2 (a)-(d) provides sample images of normal, tendon inflammation, calcific tendonitis and rotator cuff tears, respectively. Among the 120 acquired images, 40 supraspinatus images equally divided into the four classes were selected as the training data to search for powerful features and then to establish the Firefly RBF neural network with 4 hidden nodes for classification. The remaining 80 supraspinatus images were used as the test images for subsequent classifications.

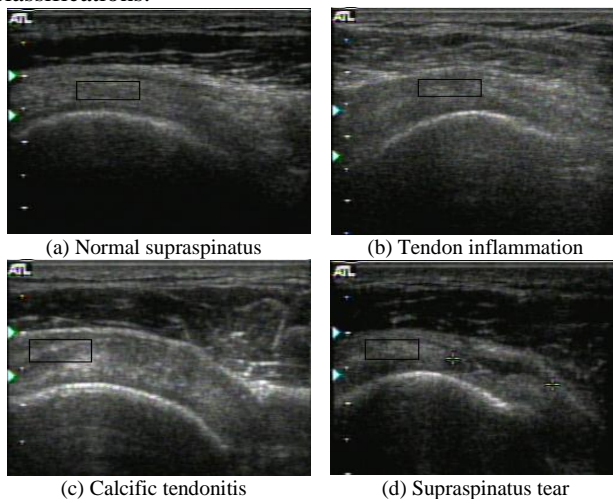


FIGURE 2 Samples of the ultrasonic images of supraspinatus. The image (a) is a normal case, image (b) is a sample of the tendon inflammation. The images (c) and (d) are calcific tendonitis and tear, respectively

The ultrasonic system settings were standardized for all of the participants and kept constant during the image acquisition. We used a depth setting of 3.0 cm. The depth-gain compensation was built into the ultrasound machine. The acoustic signal received by the ultrasonic transducer was digitized by 8 bit intensity values making the ultrasonic image. Each image consisted of pixels that were $0.1172\text{mm} \times 0.1172\text{mm} = 0.0137\text{mm}^2$. For each

image, a region of interest (ROI) with 30×60 pixels at a depth of approximately 1.5 to 2.5 cm from the body surface was manually selected by an ultrasound musculoskeletal radiologist with five years of experience in shoulder examination to extract texture features for subsequent classification. In general, radiologist avoided including the areas of ruptured tendons in the ROI selection process. All programs were implemented in Visual C++ associated with Neural Toolbox of Matlab software under on a personal computer with a 2.4GHz CPU and 1G RAM using the Window XP operating system. The parameters of the proposed Firefly RBF neural network are listed in Table 1.

TABLE 1 The used parameters of Firefly RBF neural network

Parameters	Value
Attractiveness β_0	1.0
Light absorption coefficient γ	1.0
The size of initial firefly	50
Iteration number	100
σ	0.1
Attractiveness β_0	1.0

In the experiment, each training image can be extracted to a total of 80 features that are generated from the above-mentioned four texture analysis methods. All extracted features are first required to be normalized to zero mean and unit standard deviation, which ensure the larger value input features so as not to overwhelm smaller value inputs and to reduce errors before the feature selection and classification. Table 2 lists the selected features by independent usage of the F-score criterion, that are Sum Variance (GLCM), Sum Average (GLCM), Mean Convergence (TFCM), Code Variance (TFCM) and Contrast (GLCM). The results of the texture feature selection were that all of the selected features were generated from the GLCM and TFCM. The average execution time is 0.2919 seconds for classifying an ultrasonic supraspinatus image by using Firefly RBF neural network. These results may reveal that the discriminative capability of the selected features generated from the GLCM and TFCM methods are superior to those of other methods.

TABLE 2 Texture feature selection by using the four texture analysis methods

Methods	Feature selected (d: displacement)
Four texture methods + F-scoring ranking method	Sum variance (GLCM, $d = 2$),
	Sum average (GLCM, $d = 2$),
	Mean convergence (TFCM, $d = 2$),
	Code variance (TFCM, $d = 2$),
	Contrast (GLCM, $d = 3$)

4.2 PERFORMANCE EVALUATION

In order to compare the classification results of the proposed glowworm swarm optimization trained radial

basis function network with the results of the diagnosis of clinical musculoskeletal radiologist, the supraspinatus images were evaluated and further classified with an clinical radiologist with 5 years of experiences in shoulder ultrasound examination. The correct classification rate of the 120 ultrasonic supraspinatus images by this clinical radiologist is about 95.8%. Table 3 shows the classification results by using the proposed Firefly RBF neural network. Its correct classification rate is 93.75%. In order to further investigate the performance of classification using the Firefly RBF neural network classifiers, the performance indices, such as the sensitivity, the specificity and accuracy rate are computed to compare with the results of the ML classifier, the error correcting output code (ECOC), the fuzzy SVM and the gradient descent RBF (Gradient-RBF) network algorithms. The three indices are defined in the four-class supraspinatus image classification as follows.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FN+FP}, \tag{19}$$

$$\text{Sensitivity} = \frac{TP}{TP+FN}, \tag{20}$$

$$\text{Specificity} = \frac{TN}{TN+FP}. \tag{21}$$

TABLE 3 The classification results based on the features of Table 1 using Firefly RBF network classifier

Predicted Results	Actual Results			
	Normal	Inflammation	Calcific	Tear
Normal	20	1	0	0
Inflammation	0	18	1	0
Calcific	0	1	17	0
Tear	0	0	2	20

The definition of TP, TN, FP and FN are specified as follows.

TP (true-positive): the number of correctly diagnosed diseased cases (including tendon inflammation, calcific tendonitis and tear).

TN (true-negative): the number of correctly diagnosed normal cases.

FP (false-positive): the number of misclassifications where patients are considered as being acute disease than the actual diagnosis.

FN (false-negative): the number of misclassifications where patients are classified with less severe diseases than actual diagnosis.

In addition, an effective classification method should decrease the possibility of misclassification, especially for the false-negative rate. A high false-negative rate represents the danger to underestimate the disease severity in a patient while the clinical doctor uses the classification system; therefore, the false-negative rate may be considered as an index for evaluating the performances of the RBF networks of the two different selected feature sets. Table 4 shows the four performance

indices by using different classification algorithms. Form this table, we find that the results of Firefly RBF network is superior to other methods. The accuracy of Firefly RBF network grows 1.5% with comparison to the original gradient-descent RBF network. Furthermore, the false-negative rate by using the Firefly-RBF neural network is only 0.0375. It reveals that Firefly RBF network is promising to develop into a powerful diagnosis tool in the clinical diagnosis application of the ultrasonic supraspinatus images.

TABLE 4 The four performance indices, which are accuracy, sensitivity, specificity and false-negative rate by using the five different algorithms

Performance measures	Classification methods				
	ML	FSVM	ECOC	Gradient RBF	Firefly RBF
Accuracy	84.2%	89.1%	90.8%	92.5%	93.75%
Sensitivity	0.840	0.900	0.922	0.932	0.948
Specificity	0.833	0.860	0.900	0.904	0.909
False negative rate	0.116	0.075	0.058	0.05	0.0375

4.3 RECEIVER OPERATING CHARACTERISTIC ANALYSIS

The receiver operating characteristics (ROC) analysis is based on statistical decision theory and has been applied extensively to the evaluation of classification methods [14, 15]. The ROC curve can manifest the relationship between the true-positive fraction (TPF) and false-positive fraction (FPF) with the variations in decision threshold. In general, the area under the ROC curve (AUC), A_z , is a powerful index for assessing the classification performance of the classifier. In general, a large value of AUC is desirable as AUC values greater than 0.9 suggest that the corresponding diagnosis system is very effective. The area under curve (AUC), A_z , of ROC curves of the five different algorithms are listed in Table 5. Obviously, the proposed Firefly RBF neural network is 0.951 that is superior to other four classification algorithms.

TABLE 5 The A_z of ROC curve of the five different classification algorithms

Method	Classification methods				
	ML	FSVM	ECOC	Gradient RBF	Firefly RBF
A_z value	0.876	0.923	0.927	0.941	0.951

5 Conclusion

Traditionally, the widespread method used by radiologists to diagnose the rotator cuff injury is to examine the micro/macro changes of the supraspinatus in the ultrasonic images; however, manual observations emerge several problems such as inter-observer and intra-observer variability. In this paper, the radial basis function neural network trained by the firefly algorithm was developed as the classifier for the classification of

ultrasound supraspinatus images. Based on the results of the present experiments of the classification of ultrasonic supraspinatus images, the followings can be emphasized:

1. The correct classification rate by using the Firefly RBF neural network [16] is superior to other four classification algorithms, particularly the accuracy by using Firefly RBF network improves near 1.25% compared to gradient-descend RBF network. It reveals that the firefly algorithm is effective in the training the radial basis function neural network. The results drive a probable study to develop new training algorithm for other neural networks.
2. The lower false-negative rate and the high sensitivity and specificity using the Firefly RBF network in the classification of ultrasound supraspinatus images appear that the proposed

algorithm of this paper is a reliable algorithm, and further it has potential to develop into a practical tool for clinical diagnosis. In the future study we will examine the other texture analysis method to obtain other powerful texture features for classifications.

Acknowledgment

The authors would like to thank the National Science council, ROC, under Grant No. NSC NSC 101-2221-E-251-008 for support of the work.

References

- [1] Neer C S 1972 Anterior acromioplasty for the chronic impingement syndrome in the shoulder: A preliminary report *J. Bone Joint Surg.* **54** 41-50
- [2] Iagnocco A, Coari G, Leone A, Valesini G 2003 Sonographic study of painful shoulder *Clinical Experimental Rheumatology* **21** 355-8
- [3] Arslan G, Apaydin A, Kabaalioglu A, Sindel T, Luleci E 1999 Sonographically detected subacromial/subdelroid bursal effusion and biceps tendon sheath fluid: Reliable signs of the rotator cuff tear? *J. Clin. Ultrasound* **27** 335-9
- [4] Chiou H J, Chou Y H, Wu J J, Hsu C C, Tiu C M, Chang C Y Chang 1999 High-resolution ultrasonography of the musculoskeletal system: Analysis of 369 cases *J. Med. Ultrasound* **7** 212-8
- [5] Marcello H N-B, Jose B V, Jorge E J, Gerson M 2002 Diagnostic imaging of shoulder rotator cuff lesions *Acta. Ortop. Bras.* **10** 31-9
- [6] Al-Shawi A, Badge R, Bunker T 2008 The detection of full thickness rotator cuff tears using ultrasound *Journal of Bone and Joint Surgery-British* **90-B** 889-92
- [7] Nielsen P K, Jensen B R, Darvann T, Jorgensen K, Bakke M 2006 Quantitative ultrasound tissue characterization in shoulder and thigh muscles – a new approach *BMC Musculoskeletal Disorders* **7** 1-11
- [8] Horng M H 2009 Multi-class support vector machine for classification of the ultrasonic images of supraspinatus *Expert Systems with Applications* **36** 8124-33
- [9] Horng M H, Chen S M 2009 Multi-class classification of ultrasonic supraspinatus images based on radial basis functional neural network *Journal of Medical and Biological Engineering* **29**(5) 242-50
- [10] Yang X S 2008 *Nature-inspired Metaheuristic Algorithms* Frome; Luniver Press
- [11] Lukaski S, Zak S 2009 Firefly Algorithms for Continuous Constrained Optimization Tasks, Computational Collective Intelligence. Semantic Web, Social Networks and Multiagent Systems *Lecture Notes in Computer Science* **5796** 97-106
- [12] Yang X S 2010 Firefly Algorithm, Stochastic Test Functions and Design Optimization *Int. J. Bio-inspired Computation* **2**(2) 78-84
- [13] Liu H, Motoda H 1998 *Feature selection for knowledge discovery and data mining* Norwell, MA: Kluwer Academic
- [14] Fawcett T 2006 An introduction to ROC analysis *Pattern Recognition Letters* **27** 861-74
- [15] Gonen M 2007 *Analysing Receiver Operating Characteristic Curves Using SAS* SAS Press
- [16] Horng H M, Lee Y X, Lee M C, Liou R J 2012 Firefly Metaheuristic Algorithm for Training the Radial Basis function Network for Data Classification and disease Diagnosis, Theory and New Applications of Swarm Intelligence *InTech* 115-32

Authors



Chih-Feng Chao, born on December 10, 1973, Kaohsiung, Taiwan

Current position, grades: Doctor of Computer Science and Information Engineering, assistant professor in National Pingtung Institute of Commerce.

University studies: Computer Science and Information Engineering in Tamkang University.

Scientific interest: Biologically inspired computing, context-aware computing, and wireless networking.

Publications: 6 Journal Papers.

Experience: Chih-Feng Chao received his Bachelor's degree in electronic engineering from Chinese Culture University in 1998 and a Ph.D. degree in computer science and information engineering from Tamkang University in 2005. Dr. Chao is currently an assistant professor with the Department of Computer Science and Information Engineering, and serves as the chief of Information Network Section, Division of Computer Center in National Pingtung Institute of Commerce, Pingtung, Taiwan. He is also a member of Phi Tau Phi Honor Scholastic Society. His main research interests include biologically inspired computing, context-aware computing, and wireless networking.



Ming-Huwi, Horng

Current position, grades: Doctor of Computer Science and Information Engineering, Professor and doctoral supervisor in computer Science and Information Engineering of National Pingtung Institute of Commerce

University studies: Department of Computer Science and Information Engineering in National ChengKung University.

Scientific interest: Bio-inspiring computing, Image Processing and Pattern Recognition

Publications: 63 referred Journal Papers

Experience: Ming-Huwi Horng received the B.S. degree in 1990, the M.S. in 1992 and the Phd degree in 1997 from the National Cheng Kung University, Taiwan. He is currently a Professor in the National Pingtung Institute of Commerce from 2010. He was working on the bio-inspired computing, image processing and pattern recognition and has published more than 60 referred international journal papers. His current research interests are in the medical image processing and bio-inspired computing. He is a member of IEEE, the Chinese Association of Image Processing and Pattern Recognition and the Chinese Association of Biomedical Engineering.

Material simulation based on Phong illumination model

Shiqing Yang^{1,*}, Lanfang Miao²

^{1,2} College of Mathematics/Physics and Information Engineering at Zhejiang Normal University, Jinhua - Zhejiang – China

Received 1 March, 2014, www.tsi.lv

Abstract

Realistic material simulation is one of the major works in rendering realistic graphics. In this paper, we have studied and discussed the effects of rendering realistic graphics by simulating some materials under different illumination conditions and material attributes and emission based on the OpenGL graphics technology and Phong illumination model. Experimental results show effects of emission on specular and non-specular materials and demonstrate the emission effects to some degree.

Keywords: Phong illumination model, Realistic graphics, Material simulation, Emission

1 Introduction

Realistic picture rendering is one of the most important parts of computer graphics. The major work in rendering realistic graphics is simulating illumination and material attributes of object surface. That means, if we deal with them correctly we can create realistic graphics. In general, we use some illumination model to calculate the brightness values under imaginary lightings, textures and material attributes, when we render realistic pictures. Phong illumination model is good at simulating shading and lighting effects. Moreover, its computational complexity is very low [1].

OpenGL can be used to simulate the real world based on given illumination conditions, it also can be used to control the relation between lighting and objects, and generate many different varieties of visual effects. In order to render realistic pictures, we need to simulate not only the lighting effect but also the effects of absorbing and reflecting light of objects surface. Besides, the attributes of material such as emission and reflection have great impact on rendering realistic pictures. In this paper, we will study and discuss the effects of rendering realistic graphics under different material attributes and emission values based on Phong illumination model.

2 Illumination Model

In the process of rendering realistic 3D object, we need to calculate the brightness of each pixel on visible object surface according to illumination model [2]. Illumination model is the base of rendering realistic pictures in computer graphics [3]. According to the relevant law of optical physics, illumination model calculates the light brightness and colour of each point on the surface of 3D objects projected on the observer's eyes. It defines the characteristics of light source, geometric distribution of light intensity on object irradiation surface and reflection

characteristics of surface to light, and it contains many other factors such as the type of objects, lighting attributes in the scene.

Phong illumination model is a kind of local illumination model. It has been often used in 3D computer graphics. All light sources in Phong illumination model are simulated as point light source, which only consider the effects of light direct exposure to the surface and then directly reflect to the viewpoint. The core of Phong illumination model is how light reflects from the surface and Phong illumination uses the cosine angle between reflection vector and the view vector to simulate specular reflection light.

2.1 CALCULATION OF PHONG ILLUMINATION MODEL

Phong illumination model [4] contains three different lighting types: ambient light, diffuse light and specular reflection light [5], which can be expressed as follow:

$$I = I_{amb} + I_{diff} + I_{spec}, \quad (1)$$

where ambient light I_{amb} is also called background light in Phong illumination model, which is formed by multi-reflecting of light from light source. It is the most basic illumination model in 3D scenes [6]. It enters to object surface uniformly from surrounding and reflects evenly in all directions. It is a uniform distribution of light and can be described as follow:

$$I_{amb} = k_a I_a, \quad (2)$$

where k_a is reflection coefficient of ambient light, I_a is ambient intensity.

Diffuse reflected light is a kind of uniformly reflecting light into space. It is not related with the

*Corresponding author e-mail: ysqzjnu@163.com

observer position and it has same brightness regardless of the position of the observer. In calculating diffuse reflected light, the position of light source should be considered, which is different from ambient light. Diffuse illumination model follows Lambert law [7, 8, 9], which can be described by two vectors: Vector L and Vector N.

As illustrated in fig. 1, vector L describes the position of light source, and vector N describes the vertex on object surface. The less the angle between vector L and vector N is, the greater the diffuse reflected light is. When L and N are in the same direction, diffuse reflected light is greatest. Diffuse reflected light can be expressed as follow:

$$I_{diff} = kaI_l \cos \theta, \tag{3}$$

where ka is percentage of diffuse light in the incoming light, the values range from 0 to 1 [10], I_l is lighting intensity, and $\cos \theta = N \bullet L$.

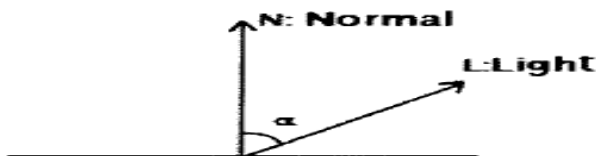


FIGURE 1 Diffuse reflected light

Specular reflection light comes from a specific direction and goes to a specific direction. In Phone illumination model, the colour of specular reflection component is independent of any material attributes [11]. It is often used to simulate the shiny surface of object, such as metal and glass bottle. In the specular reflection model, two vectors are used to calculate the specular reflection component, which are observation point vector V and reflection vector R illustrated in fig.2. The former describes the relative position of observation point (or the camera), the latter describes the reflection direction of light.

As illustrated in fig. 2, the less the angle β between V and R, the brighter the reflection light. Besides, there is an exponent n to represent gloss properties. Specular reflection light can be expressed as follow:

$$I_{spec} = ksI_l \cos^n \beta, \tag{4}$$

where ks is reflection coefficient of specular light, I_l is light intensity, $\cos \beta = V \bullet R$.

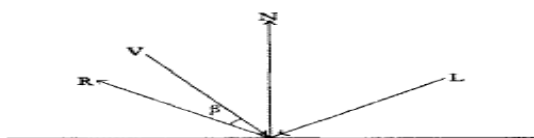


FIGURE 2 Specular reflection

From above description, we can see that the calculation formula of three components (ambient light, diffuse light and specular reflection light) in Phone illumination model can be better to reflect corresponding lighting information in the realistic lighting environments. Ambient light, diffuse light and specular

reflection light are combined to construct Phone illumination model. In RGB 3-color system, Phone illumination model can be expressed as follow:

$$\begin{bmatrix} r \\ g \\ b \end{bmatrix} = ka \begin{bmatrix} r_{pa} \\ g_{pa} \\ b_{pa} \end{bmatrix} + \sum \left[\begin{bmatrix} r_{pd} \\ g_{pd} \\ b_{pd} \end{bmatrix} \cos \theta + k_s \begin{bmatrix} r_{ps} \\ g_{ps} \\ b_{ps} \end{bmatrix} \cos^n \beta \right], \tag{5}$$

where $[r_{pa}, g_{pa}, b_{pa}]$, $[r_{pd}, g_{pd}, b_{pd}]$, $[r_{ps}, g_{ps}, b_{ps}]$ are the corresponding colour of brightness I_{pa} , I_{pd} , I_{ps} respectively. Users can directly specify the colour of ambient light, diffuse reflection light and specular reflection light, and then they can generate the pictures with different colour. By changing the parameters on above formula, different materials effects can be displayed.

2.2 MATERIAL

Material is what the object looks like. In the rendering process, material is the combination of each visual attribute. These visual attributes include colours, textures, smoothness and reflection of material surface. Only if we assign proper materials to objects, the rendering results can be realistic. If there is no illumination model, there will be no realistic material.

The colours we see are decided by the illumination in the scene and reflection of material. The wavelength of reflected light decides the colour of material. If the light falls on a surface with smoothness and brightness, the material will reflect most of the incident light. When light falls on an object, it will be grey after it equally absorb all visible light. Like light characters, material has own ambient light, diffuse light and specular reflection light, but material ambient light is affected by ambient light of panoramic view and all light sources. Specular reflection light of materials are usually decided by the light source. Besides, material has glossiness and emission. Glossiness presents roughness of the material. It affects the reflection of the object. Emission is irrelevant to any kind of illumination and the higher the value, the brighter the material gets.

3 Algorithm Description

This experiment is carried out under the win7 operating system, using VS2005 as development platform and programming under OpenGL1.2. The experiment algorithm is described as follow:

- (1) Create a 3D teapot model, define normal vector of each vertex.
- (2) Create lighting locations and types. In this paper, we use Phong illumination model.
- (3) Create a point source and enable it.
- (4) Define material attributes. In OpenGL, glMaterialfv is used to describe the material. GL-EMISSION is used to simulate the luminous intensity.

4 Experiment Results and Analysis

Illumination in OpenGL is an approximation of real illumination. In OpenGL, we use four parameters build material model: ambient, diffuse, specular and emission, the first three parameters individually present three illumination reflect by the objects, the last one present the emission. In addition, the first three parameters are respectively composed of four components R, G, B, A, which represents red, green, blue, and alpha respectively. The values range from zero to one. We adjust the parameters to achieve some realistic material. Table 1 is some values of the material attributes used in our experiment according to the above model.

TABLE 1 Parameters of the material

Material	Ambient R, G, B, A	Diffuse R, G, B, A	Specular R, G, B, A	Emission
Plastic (a)	0.31, 0.31,0.31,1.0	0.12, 0.1,0.55,1.0	0.2, 0.2, 0.2, 1.0	0.25
Plastic (b)	0.31,0.31,0.31,1.0	0.12,0.1,0.55,1.0	0.2, 0.2, 0.2, 1.0	0.5
Plastic (c)	0.31,0.31,0.31,1.0	0.12,0.1,0.55,1.0	0.2, 0.2, 0.2, 1.0	1.0
Aluminum (a)	0.4, 0.4, 0.4, 1.0	0.4,0.4, 0.4,1.0	0.7, 0.7, 0.8, 1.0	0.25
Aluminum (b)	0.4, 0.4, 0.4, 1.0	0.4,0.4,0.4,1.0	0.7, 0.7, 0.8, 1.0	0.5
Aluminum (c)	0.4, 0.4, 0.4, 1.0	0.4,0.4, 0.4,1.0	0.7, 0.7, 0.8, 1.0	1.0

In this paper, we simulate two types of materials: non-specular material such as plastic, specular material such as aluminium.

Fig.3 shows non-specular material under Phong illumination model, the background is black. From left to right, the value of emission becomes larger, the teapot becomes brighter. At last, purple teapot turns shiny white. Under this condition, emission mainly affects the colour.

Fig. 4 shows specular material under Phong illumination model, the background is also black. From left to right, the value of emission becomes larger, the teapot becomes brighter. Eventually, the teapot turns white and there is no specular on the surface.

In this experiment, the teapot is opaque; the reflected light decides the colour of the teapot. When emission value grows larger, object surface becomes brighter, and vice versa. Under non-specular material condition, the object colour becomes more and brighter with the self-luminous value become large and until the whole object become white. Under specular material, the teapot colour also becomes more and brighter with the self-luminous

value become large and until the object becomes white. Moreover, at this time the colour will not be changed with self-luminous value changing. Because specular reflected light is white, teapot eventually performances white and there is no specular reflection. As we can see from the fig.3 and fig. 4, teapot looks like a light source when emission values is 1.

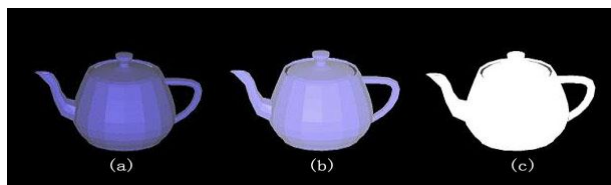


FIGURE 3 Using material parameters simulate non-specular material (plastic)

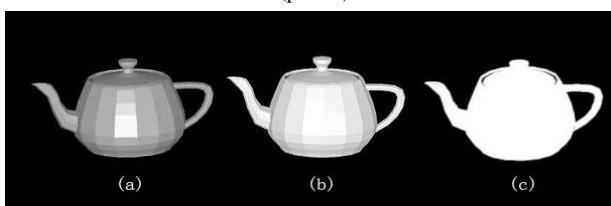


FIGURE 4 Using material parameters simulate specular material (aluminium)

5 Conclusions

Phong illumination model describes the surface characters of the luminous object, and it only shows the most obvious difference between reflective light and non-reflective light. However, it cannot show the viewer's feelings of luminance. In this paper, we have simulated realistic materials in the case of different emission values in Phong illumination model, and mainly discussed the effects of emission on non-specular material and specular material and demonstrated the emission effects material to some degree. In the future, we will focus on how to performance the true colour specular material effects.

Acknowledgment

This work is supported by National Natural Science Foundation of China (Grant No. 61170315), Open Fund of Computer Software and Theory Provincial Priority Disciplines in Zhejiang Normal University (No. ZSDZZZXK07).



References

[1] Phong B T 1975 Illumination for Computer Generated Pictures *Communications of the ACM* 18(6) 311-7 (In Chinese)
 [2] Qunsheng Peng, Hujun Bao, Xiaogang Jin 1999 *Procedural Elements for Computer Graphics* science Press, Beijing 27-34 (In Chinese)
 [3] Jia Ren, Jing Gan 2010 Illumination Simulation for Photo-Realistic Rendering of Diverse Materials' Color Texture *Microcomputer Applications* 31(2) 38-43 (In Chinese)
 [4] Ruchuan Wang, Bei Zou 2002 *Computer Graphics* Beijing Posts & Telecom Press 240-1 (In Chinese)
 [5] Xiaogang Chen, Lin Lu, Xiangyang Liu 2009 Accurate estimation of face image illuminating direction based on classical lighting model *Computer Engineering and Applications* 45 203-5

[6] Hui Wang 2009 Basic research and program realization of illumination model in three-dimensional scene *China Science and Technology Review* 17 101-2 (In Chinese)
 [7] Haipeng Chen, Xuanjing Shen, Yinda Lv 2011 Blind Identification for Image Authenticity Based on Lambert Illumination Model *Journal of Computer Research and Development* 48(7) 1237-45 (In Chinese)
 [8] Chuang Zhang, Donghui Lv 2011 Research and Application of Single Image Light Direction Estimation Algorithm *Television Technology* 35(15) 41-5 (In Chinese)
 [9] Wei Dong Zhong, Zhu Jun Jie, Li Xian Wei, et al. 2012 A new algorithm to exposing image forgeries by detection ambient illumination consistency *Advanced Materials Research* 433-440 5463 - 8

[10] Hongan Jiang, Cunhua Li 2005 The Realistic 3D Terrain Simulation Based on Phong Model *The simulation technology academic conference proceedings* 3 359-360 (In Chinese)

[11] Foley J D 2002 Computer Graphics: Principles and Practice Second Edition in C *Pearson Education* 429-34

Authors	
	<p>Shiqing Yang, born on August 5, 1987, Jiangsu, China</p> <p>Current position, grades: postgraduate of College of Mathematics/Physics and Information Engineering at Zhejiang Normal University University studies: Biological chemical and environmental engineering in Nanjing Xiao Zhuang College Scientific interest: Computer graphics and Virtual reality Publications: 1 Paper Experience: I started computer graphics since 2010 and now have a firm knowledge of it. And in the beginning of 2011 I joined the virtual reality lab, in which I started the research on virtual reality and also achieved certain amounts of results.</p>
	<p>Lanfang Miao, Born on October 15, 1963, Zhejiang, China</p> <p>Current position, grades: Professor of College of Mathematics/Physics and Information Engineering at Zhejiang Normal University University studies: Computer graphics in Zhejiang University Scientific interest: Computer graphics and Virtual reality Publications: 20 Papers Experience: Mainly engaged in computer graphics and computer 3d animation. In recent years has published more than 20 articles. In 2008, served as chairman of the 7th International Conference on Web-based Learning (ICWL).</p>

Parametric identification for GHM and application of viscoelastic damper

L J Tan^{*}, B Fang, M M Li, Y Tang, W H Huang

School of Astronautics, Harbin Institute of Technology, West Dazhi Str. 92, Nan Gang District, Harbin 150001, P R China

Received 1 March 2014, www.tsi.lv

Abstract

The GHM (Golla-Hughes-McTavish) model is extensively utilized by structural designers for studying complex structures with viscoelastic damping treatments in engineering. A kind of shear-type viscoelastic damper is investigated, and the damper is modelled with GHM model. The parameters of GHM model are identified by curve fitting and a detailed experiment in complex frequency domain. The comparison results show that the method proposed in the present paper to determine the parameters of GHM model is correct. A whole-spacecraft vibration isolation experiment is practically performed, and the results show that using the method to design the WSVI (Whole-spacecraft Vibration Isolator) is effective for isolating structure vibrations.

Keywords: viscoelastic damper, GHM model, vibration experiment, WSVI

1 Introduction

Among the passive control systems for attenuation of vibrations in structures, high damping properties of the VEM (Viscoelastic Materials) frequently are utilized for vibration isolation [1, 2]. Although the measuring method to dynamic characteristics of VEM has been standardized [3, 4], but the measuring methods of dynamic characteristics of complex structure with the VEM are all not uniform. The theory of sinusoidal sweep and GHM model are combined to measure the dynamic characteristics of the viscoelastic damper in the paper.

The GHM model is presented by Golla and Hughes [5], and improved by McTavish [6]. The GHM model uses a series of small perturbing term to represent the modulus function of the VEM. The complex approach is able to account for damping effects over a range of frequencies and complex mode behaviour. The procedure of GHM starts with plots of experimentally obtained transmissibility in the form of frequency dependent complex moduli that are then curve fit to a rational polynomial over a frequency range of interest. The rational polynomial, with coefficients reflecting the material properties of the test specimen, is used to represent the Laplace transform of the stress-strain relationship. The validation and assessment are addressed with different viscoelastic damping area under consideration. Incorporation of VEM with metal plate can provide an effective means of vibration isolation in the design of whole-spacecraft vibration isolator [7]. The parameters identified provide a conduct for designing the viscoelastic damping vibration isolator.

2 Mathematical model

The constitutive behaviour of VEM might be depended upon the frequency, temperature, amplitude and type of excitation [8]. A mathematical model considering all these effects simultaneously is very difficult to conceive. Therefore, for practical reasons, isothermal conditions are usually considered in the simulation conditions. The frequency dependent constitutive behaviour is taken into account upon the constitutive mathematical model [9].

Considering an isotropic VEM under isothermal and one-dimensional stress conditions, the Boltzmann integral constitutive equation [10] of viscoelastic material can be given as:

$$\sigma_v(t) = \int_0^t \dot{g}_v(t-\tau)\varepsilon_v(\tau)d\tau + \varepsilon_v(t)g_v(0), \quad (1)$$

where $\sigma_v(t)$ is the stress of VEM; $\varepsilon_v(t)$ is the corresponding strain; $g_v(t)$ is the relaxation function. Since the fading memory characteristics of VEM, the relaxation function is monotone decreasing function. Considering nil initial conditions, the Laplace transform of the above equation yields:

$$\sigma^*(s) = s\hat{G}(s)\varepsilon(s) = G^*(s)\varepsilon(s), \quad (2)$$

where $\sigma^*(s)$, $\varepsilon(s)$ and $\hat{G}(s)$ are the Laplace transforms of $\sigma(t)$, $\varepsilon(t)$ and $g(t)$ respectively. s is the Laplace complex variable. The GHM model uses a series of mini-oscillator terms to represent the material modulus function in the Laplace domain such that:

^{*} *Corresponding author* e-mail: ysjuuplj@163.com

$$s\hat{G}(s) = G_\infty \left[1 + \sum_{r=1}^N \alpha_r \frac{s^2 + 2\hat{\zeta}_r \hat{\omega}_r s}{s^2 + 2\hat{\zeta}_r \hat{\omega}_r s + \hat{\omega}_r^2} \right], \quad (3)$$

where the constant G_∞ corresponds to the stable value of the modulus called the final value of the relaxation function. Three positive constants $\{\alpha_r, \hat{\zeta}_r, \hat{\omega}_r\}$ are material parameters determined by curve fitting of the experimental master curves of the VEM, which govern the shape of the modulus function in the Laplace domain.

Consider a single-term GHM, the representation of the complex modulus in the frequency domain is:

$$G^*(\omega) = G_\infty \left[1 + \alpha \frac{\omega^4 + (4\hat{\zeta}^2 - 1)\hat{\omega}^2 \omega^2 + 2\hat{\zeta}\hat{\omega}^3 \omega j}{\hat{\omega}^4 + 2(2\hat{\zeta}^2 - 1)\hat{\omega}^2 \omega^2 + \omega^4} \right]. \quad (4)$$

There are four constants to be identified to describe the complex mode behaviour of VEM. The stress and strain caused by shear of the VEM can be written as:

$$\tau = \frac{F_s}{A_s}, \quad \gamma = \frac{X}{H}, \quad (5)$$

where F_s is external force; A_s is damping area in shear; X is the deformation in shear and H is the thickness of VEM. With the above equations, the complex stiffness is:

$$K_v^* = \frac{F}{X} = \frac{G^* A_s}{H} = \frac{G_\infty A_s}{H} \left[1 + \alpha \frac{\omega^4 + (4\hat{\zeta}^2 - 1)\hat{\omega}^2 \omega^2 + 2\hat{\zeta}\hat{\omega}^3 \omega j}{\hat{\omega}^4 + 2(2\hat{\zeta}^2 - 1)\hat{\omega}^2 \omega^2 + \omega^4} \right]. \quad (6)$$

3 The parametric identification of GHM model

The VEM is scarcely used as engineering structure alone due to its low modulus but as sandwich plate with constrained damp layer. Viscoelastic damping material is suitable for sandwich plate with the excellent characteristic of high loss factor, excellent bonding strength with metal.

A kind of shear-type viscoelastic damper is designed within the requirement temperature range considering isothermal conditions. Viscoelastic damping material is pasted with the upper and lower two pieces of metal plate. The dynamic characteristics of viscoelastic damper are analogous with that of the VEM used.

The experiment system, to identify the complex modulus parameters of GHM and the laws of transmissibility versus the frequency, is illustrated in Figure 1.

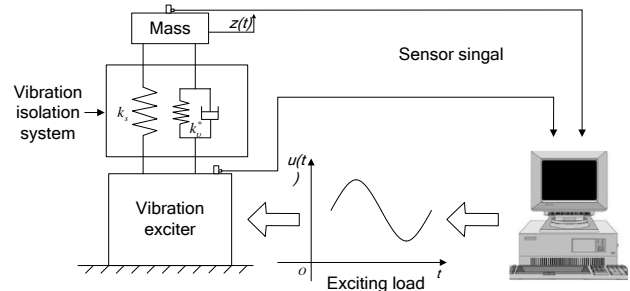


FIGURE 1 Schematic of the experimental system

The system dynamic equation is:

$$m\ddot{z} + k_s(z-u) + k_v^*(z-u) = 0, \quad (7)$$

where m , k_s and k_v^* are the mass, stiffness of thin lamina spring and complex stiffness. Excitation displacement and response of isolated structure are:

$$u = Ue^{j\omega t}, \quad z = Ze^{j(\omega t + \varphi)}. \quad (8)$$

The transmissibility from the bottom of flange to the top of mass is:

$$|T| = \left| \frac{Z}{U} \right| = \left| \frac{k_s + \frac{G_\infty A_s}{H} \left[1 + \alpha \frac{\omega^4 + (4\hat{\zeta}^2 - 1)\hat{\omega}^2 \omega^2 + 2\hat{\zeta}\hat{\omega}^3 \omega j}{\hat{\omega}^4 + 2(2\hat{\zeta}^2 - 1)\hat{\omega}^2 \omega^2 + \omega^4} \right]}{k_s - m\omega^2 + \frac{G_\infty A_s}{H} \left[1 + \alpha \frac{\omega^4 + (4\hat{\zeta}^2 - 1)\hat{\omega}^2 \omega^2 + 2\hat{\zeta}\hat{\omega}^3 \omega j}{\hat{\omega}^4 + 2(2\hat{\zeta}^2 - 1)\hat{\omega}^2 \omega^2 + \omega^4} \right]} \right|. \quad (9)$$

If the transmissibility data can be obtained, the parameters of GHM model will be identified by curve fitting. The test rig is shown in Figure 2.

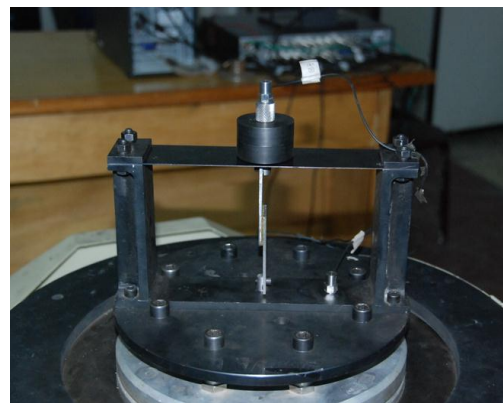


FIGURE 2 The experimental set-up for identifying parameters

The upper thin lamina spring provides the restraint of the specimen in pure shear deformation. Electrodynamics' vibration generator provides the dynamic excitation. The acceleration response of the moving mass is measured using a piezoelectric acceleration transducer. One piezoelectric acceleration transducer is located on the flange face as exciting signal

and the other is located on the cylindrical core rod which connecting mass blocks and upper plate of the viscoelastic damper. The dynamic frequency responses are all measured in the frequency domain.

The transmissibility from the excitation-point to isolated structure is acquired with sinusoidal excitation. The data are acquired as transmissibility, and complex modulus are extracted from the moduli of the transmissibility. The parameters to identify the GHM model are given: $k_s=1.32e^5N/m$, $H=e^{-3}m$, $m=6.02e^{-1}kg$ and $A_s=4e^{-4}m^2$. The comparing results of curve fitting for the transmissibility and the data obtained by practical measurement are illustrated in figure 3. The fitting curve can well represent the experimental measurement results.

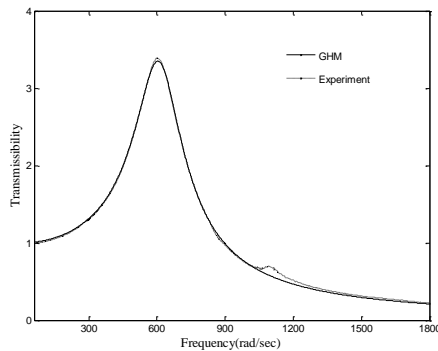


FIGURE 3 Curve fitting of GHM and experimental data ($A_s=4e^{-4}$)

The parameters obtained by above method are shown in table 1. For this material, the parameters were found to represent well the frequency range 10-300Hz.

TABLE 1 The parameters of GHM model

Parameters	G_∞	α	$\hat{\zeta}$	$\hat{\omega}$
Identification	2.132e+05	5.358	13.17	9.989e+04

To verify the parameters, G_∞ , α , $\hat{\zeta}$ and $\hat{\omega}$, the experiment with the damping area, $A_s=e^{-3}m^2$, is performed after plotting curve with the identified parameters in table 1.

In figure 4, the experimental measurement results well reflect the results of GHM model. From figure 3 and figure 4, a conclusion can be drawn that the modulus of viscoelastic damper can be well-fitted using GHM model over a frequency range of interest.

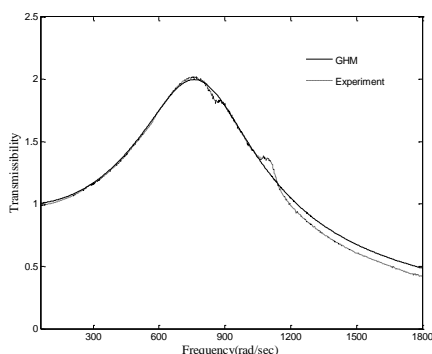


FIGURE 4 Curve fitting of GHM and experimental data ($A_s=e^{-3}$)

4 Application in WSVI

Vibration isolation is an efficient solution for the control of vibrations in structures subjected to viscoelastic damping material. The WSVI system can be modelled as two degrees of freedom, as shown in figure 5.

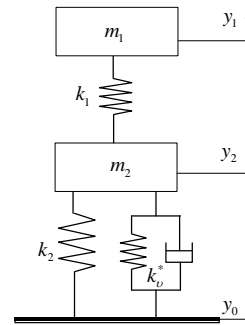


FIGURE 5 Model of WSVI

The dynamic equation of WSVI is:

$$\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{Bmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{Bmatrix} + \begin{bmatrix} k_1 & -k_1 \\ -k_1 & k_1 + k_2 + k_v^* \end{bmatrix} \begin{Bmatrix} y_1 \\ y_2 \end{Bmatrix} = \begin{bmatrix} 0 \\ k_2 + k_v^* \end{bmatrix} \begin{Bmatrix} 0 \\ y_0 \end{Bmatrix}, \quad (10)$$

where m_1 and k_1 are the mass and stiffness of whole-spacecraft, while m_2 and k_2 are the mass and stiffness of vibration isolator.

Excitation displacement, responses of whole-spacecraft and vibration isolator are:

$$y_0 = Y_0 e^{j\omega t}, \quad y_1 = Y_1 e^{j(\omega t + \psi_1)}, \quad y_2 = Y_2 e^{j(\omega t + \psi_2)}. \quad (11)$$

The Laplace form of the dynamic equation is:

$$\begin{bmatrix} -m_1\omega^2 + k_1 & -k_1 \\ -k_1 & -m_2\omega^2 + k_1 + k_2 + k_v^* \end{bmatrix} \begin{Bmatrix} Y_1 \\ Y_2 \end{Bmatrix} = \begin{bmatrix} 0 \\ k_2 + k_v^* \end{bmatrix} Y_0. \quad (12)$$

The transmissibility from the bottom of vibration isolator to the mass centre of the whole-spacecraft is:

$$T_1 = \frac{Y_2}{Y_0} = \frac{(k_2 + k_v^*)/k_1}{(1 - m_1\omega^2/k_1)(1 + (k_2 + k_v^*)/k_1 - m_2\omega^2/k_1) - 1}. \quad (13)$$

A ratio of transmissibility decline, $\delta > 30\%$, is the aim of vibration isolation for the whole-spacecraft. Aiming at δ and using the parameters identified of GHM model, the complex stiffness of viscoelastic dampers can be determined. Then, the damping area and the quantity of the viscoelastic dampers can be well computed. That is the design process of WSVI. The WSVI with the viscoelastic dampers is design to decrease the vibration transmissibility from the bottom of vibration isolator to the mass centre of whole-spacecraft.



FIGURE 6 Experiment set-up of WSVI

The verified experiment is carried out, and the experimental specimen is shown in figure 6. Six viscoelastic dampers, each damping area dimension of which is $0.02 \times 0.05 \text{ mm}^2$, are installed in the WSVI. The thickness of VEM is 0.001m. The mass of whole-spacecraft and the vibration isolator are 40.25kg and 19.35kg. The measurement experiment is driven by sinusoidal signal over a frequency rang of 10-100 Hz.

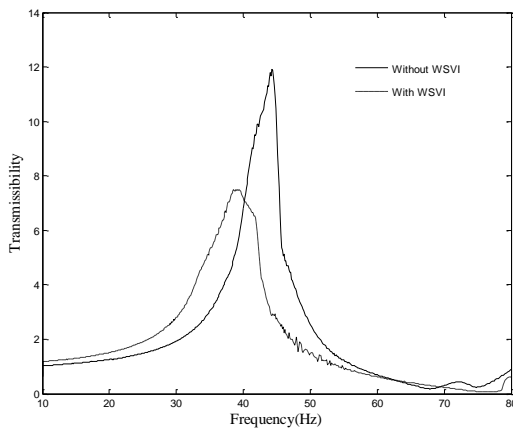


FIGURE 7 Experiment curves of before and after isolation by WSVI

The vibration isolator is connected to the cone-shell adapter. Both the vibration transmissibility from the

References

- [1] Arikoglu A, Ozkol I 2012 *AIAA Journal* **50**(3) 620-30
- [2] Meftah S A, Daya E M, Tounsi A 2012 *Thin-Walled Structures* **51** 174-85
- [3] Lili C, Baokun H, Tianchang W 2012 *Technical Acoustics* **31**(4) 402-4
- [4] Jrad H, Dion J L, Renaud F, Tawfiq I, Haddar M *European Journal of Mechanics-A/Solid* **42** 176-87
- [5] Golla D F, Hughes P C 1985 *Journal of Applied Mechanics* **52** 897-907
- [6] McTavish D J, Hughes P C, Soucy Y, Graham W B 1992 *AIAA Journal* **30**(5) 1392-9
- [7] Shaochong Z, Jinghua L, Jian C, Guoan T 2009 *Journal of Astronautics* **30**(4) 1346-50
- [8] Chunhong L, Hongbai B 2007 *Polymer Materials Science and Engineering* **23**(6) 28-31
- [9] Vasques C M A, Moreira R A S, Rodrigues J D 2010 *Mechanical Engineering* **1**(2) 96-110
- [10] Park S W 2001 *International Journal of Solids and Structures* **38**(44) 8065-92

bottom of platform without vibration isolator to the mass centre of the spacecraft and the vibration transmissibility from the bottom of WSVI to the mass centre of the spacecraft can be acquired. The experiment results are shown in figure 7.

The transmissibility presents great decline at the first-order modal peak after the viscoelastic dampers applied in the vibration system. The transmissibility decreases from 11.87 to 7.449. The ratio of transmissibility decline $\delta = 37.25\%$ exceeds 30%, which can meet the needs of whole-spacecraft vibration isolation.

5 Conclusions

Viscoelastic damping technology is an effective approach to reduce vibration transmitted to whole-spacecraft. An experimental procedure to identify the complex modulus was presented and the obtained data was used to fit the GHM model. The other experimental data was compared in order to validate the experimental procedure and parameters identified for the damping model.

Combing curve fitting and transmissibility experimental data to identification the parameters of GHM model is correct and simple to employ GHM model to perform dynamic analysis. The vibration transmissibility can be decreased signally using the parameters identification method in whole-spacecraft vibration isolation system.

Acknowledgments

The authors gratefully acknowledge the support of the Natural Science Foundation of Heilongjiang Province and the Commission of Science Technology and Industry for National Defense Project, China for this work.

Authors



Lijun Tan, born on February 15, 1984, Heilongjiang, China

Current position, grades: Ph.D. Candidate in Harbin Institute of Technology.

University studies: Master degree in Mechanical Design and Theory from Harbin Institute of Technology.

Scientific interest: Structure design and Vibration control.

Publications: 4 papers, 3 Patents.

Experience: He has participated six scientific research projects. His current research activities concern vibration control for whole-spacecraft vibration isolation, experiment research for viscoelastic damping structure and parameter identification for nonlinear systems.

	<p>Bo Fang, born on August 25, 1964, Liaoning, China</p> <p>Current position, grades: Professor and doctoral supervisor in Harbin Institute of Technology, Director of Chinese Society for Vibration Engineering.</p> <p>University studies: Ph.D. degree in General Mechanics from Harbin Institute of Technology.</p> <p>Scientific interest: Spacecraft dynamics and Vibration control.</p> <p>Publications: over 60 papers, 5 Patents.</p> <p>Experience: He is visiting scholar at IOWA State University from 2000 until 2004. He is the director of Chinese Society for Vibration Engineering. His current research activities concern active and passive vibration control for whole-spacecraft vibration isolation.</p>
	<p>Mingming Li, born on October 27, 1982, Heilongjiang, China</p> <p>Current position, grades: Ph.D. in Harbin Institute of Technology.</p> <p>University studies: Master degree in General and Fundamental Mechanics from Harbin Institute of Technology.</p> <p>Scientific interest: Active vibration control and Constrained layer damping structure.</p> <p>Publications: 5 papers.</p> <p>Experience: He has participated three scientific research projects. His current research activities concern piezoelectric constrained layer damping treatment.</p>
	<p>Ye Tang, born on March 10, 1984, Liaoning, China</p> <p>Current position, grades: Ph.D. Candidate in Harbin Institute of Technology.</p> <p>University studies: Master degree in Mechanical and Electronic Engineering from Harbin Engineering University.</p> <p>Scientific interest: Fluid-solid coupling dynamics.</p> <p>Publications: 4 papers.</p> <p>Experience: He has participated three scientific research projects. His current research activities concern dynamics of fluid-solid coupling structure.</p>
	<p>Wenhui Huang, born on July 22, 1926, Zhejiang, China</p> <p>Current position, grades: Professor and doctoral supervisor in Harbin Institute of Technology, Academician of Chinese Academy of Engineering, Honorary President of Chinese Society for Vibration Engineering.</p> <p>University studies: B.S degree in Mechanical Dynamics from Zhejiang University, diploma of Seminar in Harbin Institute of Technology.</p> <p>Scientific interest: Failure diagnosis and Spacecraft vibration control.</p> <p>Publications: over 150 papers, 6 monographs</p> <p>Experience: He is an expert in mechanics and vibration. He obtained significant results in aerospace engineering, failure diagnosis and kinetic control. He is academician of the Chinese Academy of Engineering and member of the National Natural Science Foundation of China. His current research activities concern failure diagnosis and nonlinear vibration control.</p>

Study on the sealing properties of the sealing structure for the rotating chamber of a certain cased telescoped ammunition gun

Longmiao Chen*, Qiang Fu, Gui Lin

College of Mechanical Engineer, Nanjing University of science and technology, Nanjing 210094, Jiangsu, PR China

Received 1 March 2014, www.tsi.lv

Abstract

To solve the spherical transient high pressure gas seal problem for the rotating chamber of a certain medium calibre cased telescoped ammunition gun, a self-impacted combined sealing system was newly designed which can be placed at semi combustible cartridge of the cased telescoped ammunition. The sealing mechanism of the structure was analysed and simulation studies on the comprehensive properties of the sealing structure were carried out via the FEM dynamic response. In addition, the simulation and verification tests were conducted to test the sealing performance of the sealing structure. The results of the simulation analysis and the experiments demonstrate that the designed sealing structure has a good sealing performance and can solve the spherical transient high pressure gas seal problem for the rotating chamber of the medium calibre cased telescoped ammunition gun, and it is expected to offer a reference value to solve related problems in engineering.

Keywords: Sealing for the Rotating Chamber, FEM, Test Verification, Cased Telescoped Ammunition

1 Introduction

Cased telescoped ammunition (CTA) gun is a kind of artillery, which use an integrative ammunition with the projectile embedded in simple cylinder cartridge case. The most important feature of CTA is its regular ammunition shape and short length. CTA gun commonly use new latching principle of rotating chamber, which will compact the loader structure, and reduce the overall size of the artillery systems [1, 2]. Due to the separate design of CTA chamber and the gun tube, and the frequent rotation of opening and closing, this special structure is inevitable to make there a gap between the chamber and the spherical interface of barrel ends. If there is no reliable sealing structure, the condition of pressure up to 400 MPa and temperatures up to 3000K in the bore and rotating chamber will lead to gas leak and serious effect on gun performance which contains great security risks. Thus, the seal problem for the rotating chamber of a CTA gun is a key technology, which must be resolved [3].

Rotating chamber high temperature and high-pressure gas make a great impact on the seal member in a very short period, which is a typical transient high-pressure seal problem at the gap of the spherical interface. In this paper, a self-impacted combined sealing system was newly designed which can be placed at semi combustible cartridge of the CTA. The sealing mechanism of the structure was analysed and simulation studies on the comprehensive properties of the sealing structure were carried out via the FEM dynamic response. In addition,

the simulation and verification tests were conducted to test the sealing performance of the sealing structure.

2 How to use the template

The self-impacted combined sealing structure is placed at a non-combustible structure of both ends of semi combustible cartridge of the CTA. The structure is shown in Figure 1. Taking into account the similarities of the sealing structure and the mechanism between both ends of the cartridge, only the front sealing structure of the cartridge is analysed. There are three major leakage channels A, B, C in the sealing structure as shown in Figure 1. The sealing mechanism is as follows [4]:

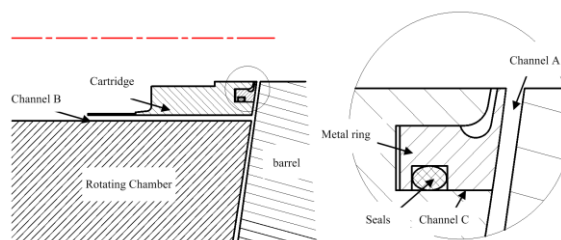


FIGURE 1 Schematic Diagram of Sealing System

Channel A: the special design of spoon-shaped metal ring makes generated gas quickly spread to the spoon-shaped area before the gas into the channel A. In addition, relative to the "spoon" area, channel A is just a small gap, so the right force along the barrel axial of metal rings is much larger than the left. At the beginning period of gunfire, the gas can quickly push the metal ring moving right along the barrel axis to meet the barrel end,

* Corresponding author e-mail: 13913870745@139.com

so that the metal ring can seal the channel A. With the continual rising of chamber pressure, the force to the metal ring is gradually increasing. The more closely is the metal ring wedge to the right, the more reliable the sealing performance is.

Channel B: according to the elastic-plastic theory of the thin-walled cylinder and the thick-walled cylinder, when the chamber pressure is small, the thin-walled segment and the thick-walled segment of the cartridge come into being radial expansion and elastic deformation, and at this time, a small amount of gunpowder gas leak in channel B. When the chamber pressure continues to raise, the cartridge thin-walled segment become in contact with the rotating chamber, after that, channel B is sealed. Moreover, the contact area gradually expanded over time, channel B is further sealed. When the chamber pressure reaches a certain value, the thin-walled segment and thick-walled segment are fully fit with the rotating chamber, than there are expansions, plastic deformation. In the process of chamber pressure down to zero by the peak, the plastic deformation of the cartridge cannot be restored to the initial state, so there may be residual bonded in a local area of the cartridge and the rotating chamber.

Channel C: in the assembly, since under the preload from the metal ring of the cartridge, the seals always fit closely with the cartridge at a low pressure, which can ensure the sealing in the low-pressure. In addition, as the pressure increases, the metal ring radial expansion force progressively larger. The greater pressure metal ring passed to seals is, the better the seal performance is.

The above-mentioned sealing mechanism indicates that, channel A, C can be achieved quick seal at low pressure, and at high pressure to achieve a reliable seal. As to channel B, when to seal or whether to seal in the end of the combustible cartridge is largely dependent on the design of the thin-walled segment structure, and it is the key to the successful seal in the entire sealing system. Therefore, this paper focuses on analysing the sealing performance of channel B.

3 Dynamic Simulation Analysis of Sealing Performance of the Sealing Structure

3.1 SIMULATION MODEL

As to the semi-combustible cartridge geometry, it is a typical axisymmetric problem that boundary conditions are symmetrical to the rotary axis, and displacement under load, stress and strain are also symmetrical to this axis [5, 6, 7].

In order to analyse the impact of the metal ring and rubber seals on the contact condition of cartridge and rotating chamber, a finite element model with and without metal rings and seals of cartridge were established. In addition, the error is only 1.25% after calculating and contrasting the result, so the metal ring

and rubber seals are omitted during models simplification.

Taking into account the accuracy and computation time, the model is sliced in the appropriate region, mesh refinement in the contact area, the stress concentration and other key part. On one hand, the non-contact area of the chamber, on the other hand, to increase the mesh density appropriately and reduce the element number of the non-focus part, which can reduce the calculation time and improve the efficiency of the finite element calculation. A four-node axis of symmetry hourglass control reduced integration elements are used in calculation. The mesh density in the key areas is 0.5mm, other parts is adjusted according to density [8, 9].

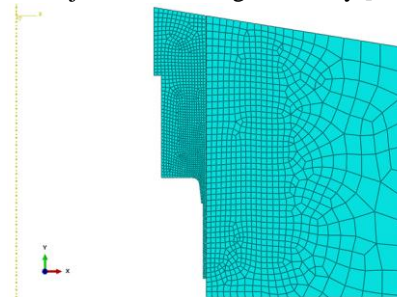


FIGURE 2 Finite Element Model of Semi-combustible Cartridge

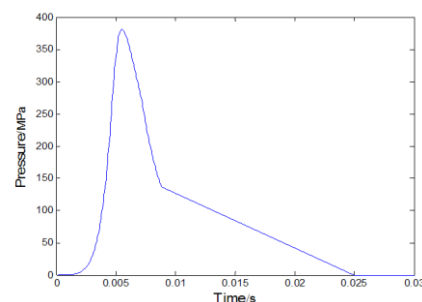


FIGURE 3 New Artillery Bore Pressure Curve

The time of the chamber pressure load curve is 0.03s. The constraints in the end surface of the model limit the axis movement of the rotating chamber and the semi-combustible cartridge. Then loading chamber pressure is exerted on each face of the cartridge. Moreover, the chamber pressure curve is shown in Figure 3.

ABAQUS/Explicit analysis is adopted to analyse nonlinear problems of the instantaneous dynamic seal. Penalty friction formula is used to define the contact parameters. ABAQUS / Explicit general contact algorithm employs the contact algorithm, and specify the master surface and the slave surface. The sampling interval, n of every n interval is set to 3000, the sampling period is 0.01ms, which not only ensure the appropriate scale data, but also achieve the purpose of smooth curves.

3.2 RESULTS OF SIMULATION

Structural strength analysis. Figure 4 and Figure 5 indicate a stress contours of the sealing structure. The yield and the deformation of various parts can be seen. The results show that sharp corners, the transition portion has stress concentration, while the other regions stress

fluctuation is small. No severe deformation has occurred in the sealing structure and the structural strength can meet the requirements.

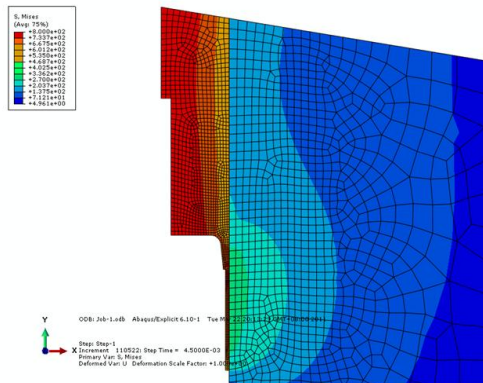


FIGURE 4 Stress Contours at 4.5ms

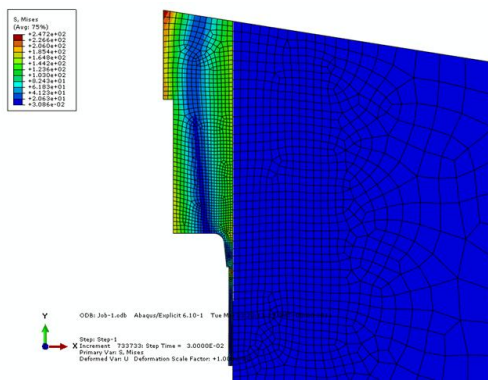


FIGURE 5 Stress Contours at 30ms

3.3 ANALYSIS OF SEALING STRUCTURE OBTURATION PERFORMANCE

In order to estimate the obturation of the sealing structure, namely the fit condition of contact area at each moment, Node7, Node143, Node164 and Node178 are selected as research subjects with a distance of 0mm, 23.0689mm, 74.5016mm and 102.550mm respectively from the top of the contacting area on the top. The contact stress vary with time is analysed:

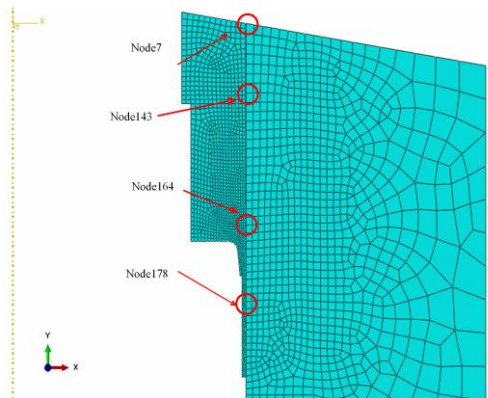


FIGURE 6 Node Coordinate in the Sealing Structure

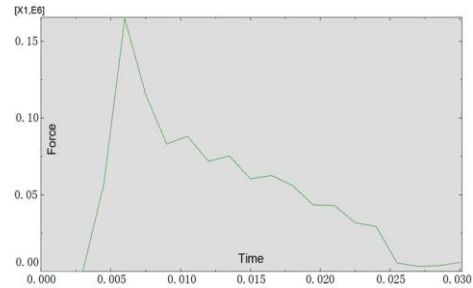


FIGURE 7 Contact Stress on Node7

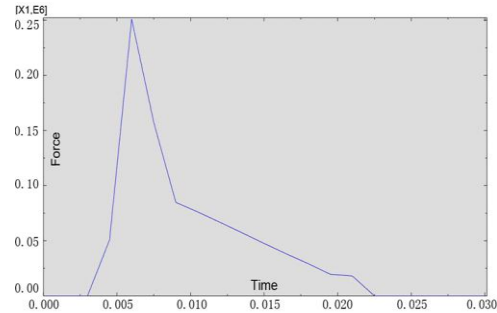


FIGURE 8 Contact Stress on Node143

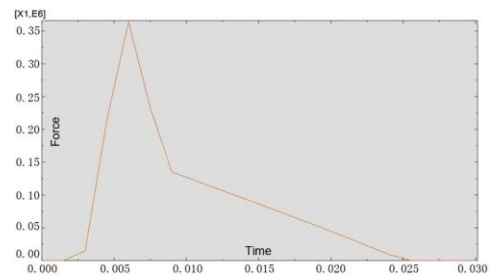


FIGURE 9 Contact Stress on Node164

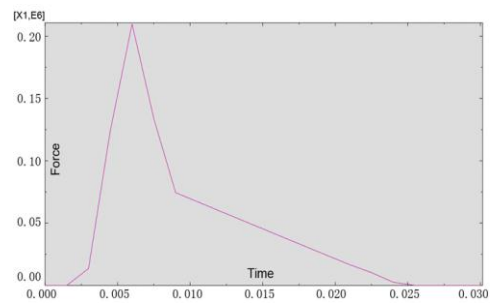


FIGURE 10 Contact Stress on Node178

It can be seen in figures 8 to 9 that Nodes 143 and 164 are in contact with chamber first when the sealing structure are exposed to the chamber pressure of 4Mpa about 2ms. With the chamber pressure rising, the bonded area gradually extends from the middle of the sealing structure to both ends as shown in figures 7 and 10. As is show in figures 7 and 10, Node 7 and Node 178 are also in contact with the chamber about 2.8ms. Therefore, the entire sealing segment has completely contacted with the chamber. Compared with the bore pressure curve, contact stress at each node reaches the maximum in 5.5ms, which is 0.3ms after the chamber pressure reaches the maximum. Since then, chamber pressure gradually decreases, and Node 143 is separated from chamber when the gas reaction time is about 23ms. Moreover, Node 164 and Node 178 are separated from chamber in 2ms later.

When the chamber pressure drops to zero, there is still residual contact stress with Node 7, which is small and exerts little effect on shell performance.

It can be seen from the above analysis that, the sealing structure seals quickly in the initial sealing and the contact pressure of the sealing structure increases while the chamber pressure increases. The performance of sealing is valid throughout the work period.

4 Sealing performance test

In order to verify the sealing performance of the self-impacted sealing structure, a seal test device is designed to test the seal effect. In the test, the gunpowder in confined device is ignited, to produce analogue bore environment of high temperature and high pressure gas, test pressure curve and the main part of the obdurate force. The test device is shown in Figure 11, and the sealing device test bench real photos are shown in Figure 12.

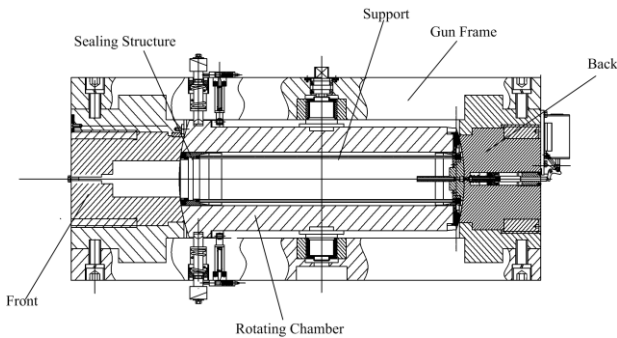


FIGURE 11 Schematic Diagram of Test Device



FIGURE 12 Experimental Device

Sealing performance is tested under the pressure in the bore of about 200MPa, and transient pressure curve and the strain of the three positions in rotating chamber. The specific location is shown in the Figure 13. The bore pressure curve and specific location of the stress curve is shown in Figure 14 and it can be seen that after bore crimp reaches to the peak of 221.4MPa , under the action of the sealing structure, the pressure drop almost to zero within 30 milliseconds, the pressure curve maintains the level to meet the requirements of the bore seal time of 20 milliseconds. Therefore, the sealing device and good sealing effect can meet the needs of practical engineering.

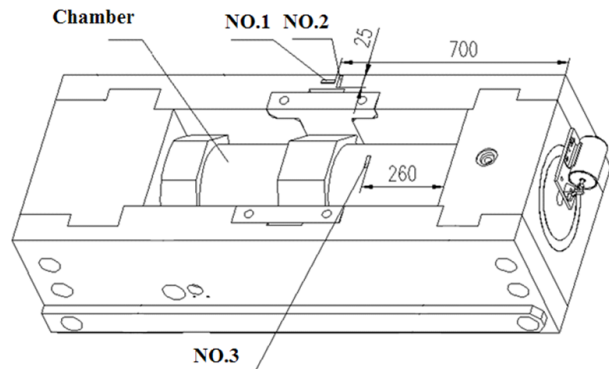


FIGURE 13 Location of Test Point

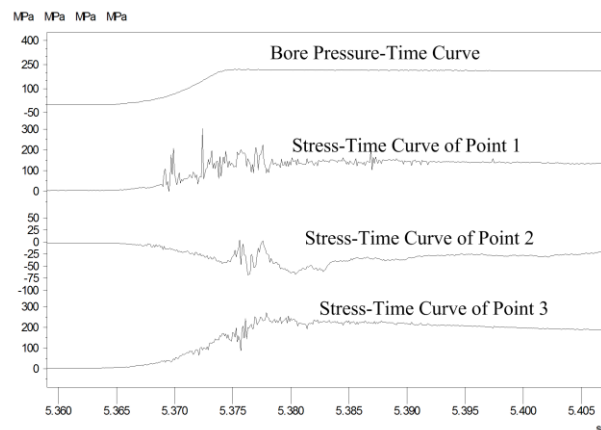


FIGURE 14 Test Results

5 Conclusions

In this paper, a self-impacted combined sealing system was newly designed which can be placed at semi combustible cartridge of the CTA. The sealing mechanism of the structure was analysed and simulation studies on the comprehensive properties of the sealing structure were carried out via the FEM dynamic response. In addition, the simulation and verification tests were conducted to test the sealing performance of the sealing structure. The results of the simulation analysis and the experiments demonstrate that the designed sealing structure has a good sealing performance and can solve the spherical transient high-pressure gas seal problem for the rotating chamber of the medium calibre CTA gun, and it is expected to offer a reference value to solve related problems in engineering.

Acknowledgments

The authors would like to thank for financial support the National Natural Science Foundation of China. (51205207)

References

- [1] Kelly W J, Marcher W D 2002 *Cased Telescoped Ammunition Smart Seal Development* Army TACOM-ARDEC
- [2] Nusca M J, Horst A W, Newill J F 2004 *Simulations of Notional Telescoped Ammunition Propelling Charge* ARL-TR-3306
- [3] Nusca M J, Horst A W 2005 *Progress in Modelling Ignition in a Solid Propellant Charge for Telescoped Ammunition* ARL-TR-3673
- [4] Xu Hao Seal 1990 *Metallurgical Industry Press* 20-28 Beijing (In Chinese)
- [5] Zhang Hao, Zhou Yan-huang 2006 A seal design under transient high pressure condition *Lubrication Engineering* (5) 132-4 (In Chinese)
- [6] Zhang Xun, Zhang Xiang-yan 2009 Simulation Research on New Combined Barrel Seal Structure *Journal of Ballistics* 21(1) 63-6 (In Chinese)
- [7] Jun Yan, Strenkowski J S 2006 A Finite Element Analysis of Orthogonal Rubber Cutting *Journal of Materials Processing Technology* 174 102-8
- [8] Nikas G K, Sayles R 2004 Nonlinear Elasticity of Rectangular Elastomeric Seals and Its Effect on Elastohydrodynamic Numerical Analysis *Tribol. Int.* 37 651-60
- [9] Yu Weiwei, Cai Lixun, Ye Yuming, et al. 2006 Spring back Properties of Inconel718 Alloy O-Ring *Engineering Mechanics* 23(6) 142-7 (In Chinese)

Authors

	<p>Long MiaoChen, born on May 6, 1979, Jiangsu, China</p> <p>Current position, grades: Assistant Professor of Mechanical Engineering School, Nanjing University of Science and Technology 200 Xiaolingwei St., Nanjing 210094, China.</p> <p>University studies: Ph.D Mechanical Design and Manufacturing in Nanjing University of Science and Technology</p> <p>Scientific interest: Computer Modelling, Mechanical vibration and control, Advanced materials ,and system integration technique in Mechanical Engineering</p> <p>Publications: 18 Patents, 25 Papers</p> <p>Experience: B.E. Nanjing University of Science and Technology, China, 2001; M.S.and Ph.D.Nanjing University of Science and Technology, China, 2005; February, 2006 –present: Nanjing University of Science and Technology, China ,Assistant Professor, Adjunct Professor, School of Mechanical Engineering; First prize, Science and Technology Progress Award of Jiangsu province; Excellent worker in Science and Technology of Nanjing University of Science and Technology; Member of Chinese Society of Vibration Engineering</p>
	<p>Qiang Fu, born on October 14, 1985, Jiangsu, China</p> <p>Current position, grades: Postgraduate of Mechanical Engineering School, Nanjing University of Science and Technology 200 Xiaolingwei St., Nanjing 210094, China.</p> <p>University studies: M.S. Mechanical Design and Manufacturing in Nanjing University of Science and Technology</p> <p>Scientific interest: Computer Modelling, Structure Optimization</p> <p>Publications: 1 Paper</p> <p>Experience: B.E. Nanjing University of Science and Technology, China, 2007; M.S. Nanjing University of Science and Technology, China, 2014</p>
	<p>Gui Lin, born on February 28, 1990, Jiangsu, China</p> <p>Current position, grades: Postgraduate of Mechanical Engineering School, Nanjing University of Science and Technology 200 Xiaolingwei St., Nanjing 210094, China.</p> <p>University studies: M.S. Mechanical Design and Manufacturing in Nanjing University of Science and Technology</p> <p>Scientific interest: Computer Modelling, Structure Optimization</p> <p>Publications: 1 Paper</p> <p>Experience: B.E. HuaQiao University, China, 2011; M.S. Nanjing University of Science and Technology, China, 2014</p>

A self-adaptive selective method of remote sensing image classification algorithms

Xin Pan, Hongbin Sun*

School of Computer Project & Technology, Changchun Institute of Technology, Changchun City, Jilin Province, China, 130012

Received 28 January 2014, www.tsi.lv

Abstract

Remote sensing image classification algorithms, which can obtain information of land use/cover quickly and inexpensively have been widely used in the field of GIS. The quality of classification results is not only affected by the quality of remote sensing data, but also affected by the character of classification algorithm. At present, despite a lot of algorithms have been proposed, but users usually meet difficulties in algorithm selection due to single classification algorithm can not applicable to all classification cases. This study proposes a self-adaptive selective method for remote sensing image classification algorithms based on data complexity evaluation, through data complexity evaluation, our method can distinguish remote sensing data's character even from same satellite sensor and give user recommendation of algorithm selection. Experiments indicate that the algorithms selected by this method can achieve higher classification accuracy, which provides the recommendation for the selection of appropriate classification models to users.

Keywords: Remote Sensing image, Classification, Algorithms Evaluate, Data Complexity

1 Introduction

Remote sensing image classification algorithms classify the entire images by using a few training samples, which can obtain information of land use quickly and inexpensively and have been widely used in the field of GIS. At present, many different algorithms (including Naïve Bayes, ID3, CART, KNN, Neural Networks, SVM etc.) have been applied to remote sensing image classification [1]. The quality of classification results is not only affected by the quality of remote sensing data, but also affected by the classification algorithm [2]. Just as like No Free Lunch theorem represent: "If algorithm A outperforms algorithm B on some cost functions, then loosely speaking there must exist exactly as many other functions, where B outperforms A" [3], data set's characteristics varying greatly and there are no single classification algorithm can applicable to all the cases [4], so select appropriate classification algorithm for remote sensing image classification is very important.

Many scholars have made research in classification algorithms selection field: Brodley proposed a knowledge based method to search an algorithm [5]; Gama proposed a linear regression method to predict algorithm's accuracy [6], Brazdil further presented a meta-learning method to select candidate classification algorithms [7]. Song gave an automatic recommendation framework for classification model selection [4]. However, the studies above are based on data characteristic of feature structure, type, range features; remote sensing image data will have same data structure information, which obtained from

same satellite sensor, and have similar data structure and statistical information, which from different satellite sensor, this will lead a difficulty to distinguish the data using the above methods. Therefore, it is necessary to introduce new evaluation method to describe remote sensing data's character.

Data complexity is a method, which can characterize data measures on the training data instead of experimenting with train data [8]. A lot of data complexity measures, concerning statistical, geometrical and information theoretic descriptions have been proposed in past few years [9]. It can give a relation between classifier performance and training data character and we can further give recommendation for select appropriate classifiers by the help of data complexity evaluation [10-12]. Therefore, data complexity can be a more effective method to describe the characteristics of the data.

This study proposes a new method named Self-adaptive Selective Method of Remote Sensing Image Classification Algorithms based on Data Complexity Evaluation (SSMRICADCE), which describes the characteristics of remote sensing images with data complexity, obtains the relationship between the characteristics of data and classification accuracy of algorithms through a large number of remote sensing datasets, and further proposes suggestions for the selection of classification algorithms on that basis. Experiments indicate that the algorithms selected by this method can achieve higher classification accuracy, which provides the knowledge for the selection of appropriate

*Corresponding author e-mail: 101103991@qq.com

classification models in the study of land use.

The remainder of this paper is organized as follows: Section 2 provides remote sensing data characterization method based on data complexity evaluation, section 3 provides the algorithm of proposed method, section 4 gives experiments result and section 5 draws conclusions.

2 Remote sensing data characterization based on data complexity

There are many formulas in data complexity evaluation field; it is difficult for a single formula to describe remote sensing data thoroughly. Therefore, we use a data complexity evaluation vector, which has three indexes to describe the character of a remote data set.

1) *Index of Fisher’s discriminant ratio*

Fisher’s discriminant ratio can describe how separated classes according to features:

$$i1 = \frac{\sum_{i=1}^c n_i \times \delta(m, m_i)}{\sum_{i=1}^c \sum_{j=1}^{m_i} \delta(x_j^i, m_i)}, \tag{1}$$

where n_i denotes the number of samples in class i , δ is a metric, m is the overall mean, m_i is the mean of class i , and x_j^i represent the sample j belonging to class i [11].

2) *Index of volume of overlap region*

The volume of the overlap region for two classes can be represented by the product of normalized lengths of overlapping ranges for all features [11]:

$$i2 = \prod_k \frac{\min \max k - \max \min k}{\max \max k - \min \min k}, \tag{2}$$

where $k=1,2,3..n$ and $\min \max k = \min\{\max(f_k, c_1), \max(f_k, c_2)\}$, $\max \min k = \max\{\min(f_k, c_1), \min(f_k, c_2)\}$, $\max \max k = \max\{\max(f_k, c_1), \max(f_k, c_2)\}$; $\min \min k = \min\{\min(f_k, c_1), \min(f_k, c_2)\}$.

3) *Index of pooled Mahalanobis distance*

Pooled Mahalanobis distance can describe the distance between classes i and j :

$$D_{pooled}^2(i, j) = (\mu_i - \mu_j)' \left(\frac{\Sigma_i + \Sigma_j}{2} \right)^{-1} (\mu_i - \mu_j), \tag{3}$$

where m is the mean vector of reflectance values, and Σ is the variance–covariance matrix. The index of pooled Mahalanobis distance can be describe as:

$$i3 = \sum_i^c \sum_j^c D_{pooled}^2(i, j). \tag{4}$$

Through formulas above, we can describe remote sensing data’s character by a vector from discriminant, overlap and classes distance level:

$$dcV=(i1,i2,i3). \tag{5}$$

Data sets which have similar data complexity evaluation vector would have similar requirement in classification, the distance of two data complexity evaluation vector in a data set d can have represented by the following formula:

$$dDCV(vi, vj, d) = \sum_k^3 \frac{|v_{ik} - v_{jk}|}{2\delta(d1)}. \tag{6}$$

The following meta-information in a remote sensing image can distinguish between remote sensing images also plays an important role:

- $m1$ =number of bands;
- $m2$ =cellsize;
- $m3$ =Source Type;
- $m4$ =pixel Type;
- $m5$ =pixe Depth;

A meta-information vector can be represent as follows:

$$metaV=(m1,m2,m3,m4,m5); \tag{7}$$

the distance of two meta-information vector can have represented by the following formula:

$$dMI(mi1,mi2)= (number\ of\ difference\ items)/5; \tag{8}$$

A remote sensing data character can be represented by meta-information vector and data complexity evaluation vector:

$$RSC(data)={miv; dcv}=(meta\text{-information vector; data complexity evaluation vector); \tag{9}$$

the RSC can describe a remote sensing date d and the differences between $R1$ and $R2$ can be calculate from following formulas:

$$\begin{aligned} distance(R1, R2, d) = & \alpha 1 \times dM1(R1, R2) \\ & + \alpha 2 \times dDCV(R1, R2, d). \end{aligned} \tag{10}$$

The process of remote sensing data characterization can be represented as the following algorithm:

Algorithm: Remote Sensing Image Data Characterization ($RSIDC$)

Input: Remote Sensing image RS , training samples’ positions and catalogues SPC

Output: remote sensing data character *RSC*

- 1) *dataSet*=construct multi feature training data set from *RS* and *SPC*;
- 2) *mateInfo*= gather number of bands, cellsize, Source Type, pixel Type, pixe Depth information from *RS*;
- 3) *metaV*=construct vector by formula (7) from *mateInfo*;
- 4) *dcV*=construct vector by formula (5) from dataset;
- 5) *RSC*={ *metaV* ; *dcV* }
- 6) return *RSC*

End

From *RSIDC* Algorithm, we can characterize a remote Sensing Image Data and its training samples.

3 The self-adaptive selective method of remote sensing image classification algorithms

With the help of remote sensing data characterization from above section, we can realize the method of sensing image classification algorithms selection; the method can be described as follows:

Method: Self-adaptive selective method of remote sensing image classification algorithms (*SSMRSICADCE*)

Stage 1: Training Stage

Input: A large number of remote sensing image and their training samples

Ouput: Relationship data set *RDS*

RDS=Combine remote sensing image, data character, classification algorithm, classification accuracy together.

End

Stage 2:

Input: A remote sensing image *RS* and training samples *TS*

Output: recommend algorithms *RAS*

RAS= Self-adaptive recommend user to select classification algorithms with the help of *RDS*

End

Through *SSMRSICADCE* we can self-adaptive recommend user to select classification algorithms; it has two stages, as depicted in Fig 1, stage 1 aims at get the relationships from remote sensing image, data character, classification algorithm, classification accuracy.

From Figure 1, the detail of Stage 1 can be represented as follows:

- 1) Construct a training database from Remote sensing images and training samples. In the database, each group of remote sensing image and corresponding samples can construct a train training sample set, and each set can obtain remote sensing data character through *RSIDC* algorithm;
- 2) Lots of classical classification algorithms were gathered and construct a Classification algorithms database.
- 3) Classification algorithms database provide

algorithm and classification each training sample set and obtain their classification accuracy.

- 4) Each training sample set and each classification algorithms together with their data character and corresponding classification accuracy were combine into relationship, we can obtain the relation describe as follows:

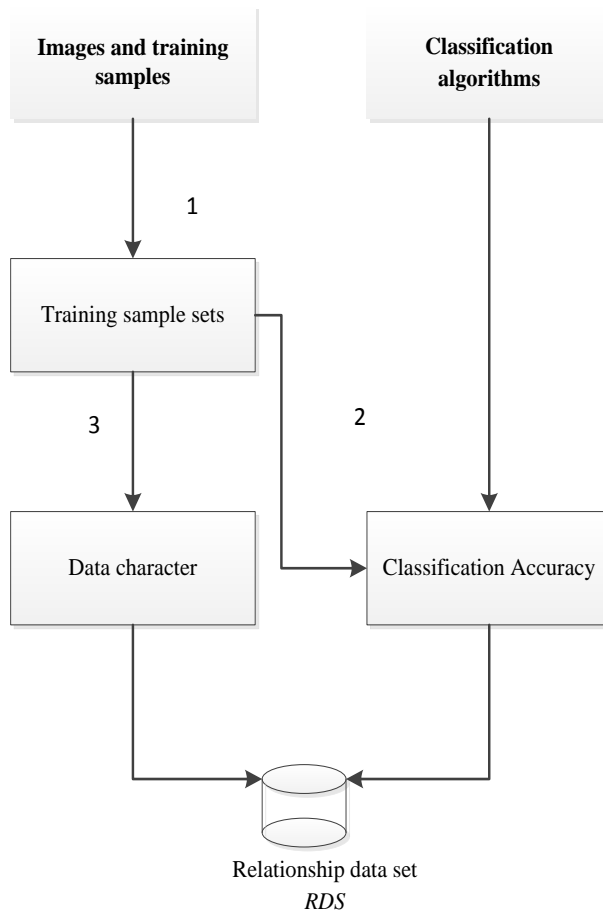


FIGURE 1 Get the relationships from remote sensing image, data character, classification algorithm, classification accuracy

Further save all the data mentioned above into Relationship data set *RDS*. Through stage 1, we can obtain a lot of relationships which can be regarded as knowledge to recommend classification algorithm selection. In stage 2 this knowledge are firstly used to select similar data character (see Figure 2).

Algorithm: Find similar data character (*FSDC*)

Input: Relationship data set *RDS*, A remote sensing image *RS* and training samples *TS*

Output: founded data characters *FDC*

- 1) *T_RSC*= get data character from *RS* and *TS* by algorithm *RSIDC*;
- 2) *FDC*= select *RSC* from *RDS* where *RDS.DC.miv=T_RSC*;
- 3) if *FDC*<>NULL Then return *FDC*; return data characters with exactly same data structure end if;

- 4) $FDC_distance$ = select all the DC from RDS , and calculate formula (10) with $\alpha1=1$ and $\alpha2=0$;
- 5) FDC = select the items from RDS with $FDC_distance < a$ threshold;
- 6) return FDC ; return data character with similar data structure.

End

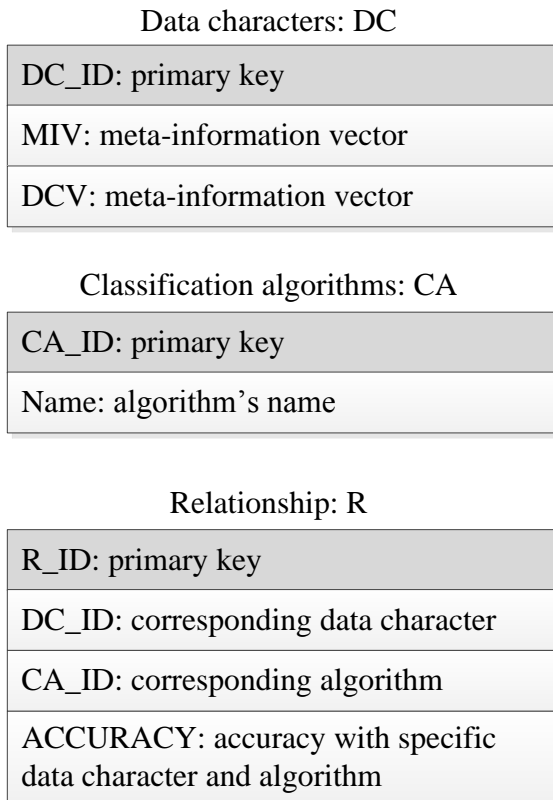


FIGURE 2 The table structure of relationship

The algorithm $FSDC$ we can find similar data characters, in the next step stage 2 need select most similar data character and select algorithms with may be archive higher classification accuracy (see Figure 3).

Algorithm: get recommend algorithms (GRA)

Input: founded data characters FDC , relationship data set RDS , data character T_RSC , number threshold NT ;

Output: recommend algorithms RAS

- 1) $DC_distance = T_RSC$ get all the distance from FDC , calculate formula (10) with $\alpha1=0.5$ and $\alpha2=0.5$;
- 2) FDC = select top NT data characters from FDC with $DC_distance$ ascend order;
- 3) $Relations$ =select all the relationships $RDS.R$ where $RDS.R.DC_ID$ in (FDC);
- 4) $OrderedGroup$ = $Relations$ split into groups by $RDS.R.CA_ID$ and each group's corresponding accuracy grade=average(accuracy in this group)+ (current group member number)/(FDC member), and all the group arrange in descending order by accuracy grade;

- 5) RAS = select classification algorithms from $RDS.CA$ where $RDS.CA.CA_ID$ in ($OrderedGroup$);
- 6) return RAS .

End

The flowchart of classification algorithm recommendation and obtain a classification result can be seen as follows:

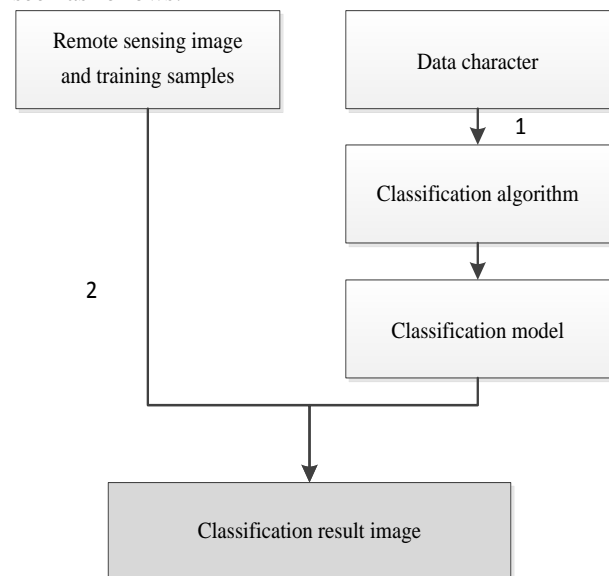


FIGURE 3 Classification algorithm recommendation and classification result

As can be seen from Figure 3:

- 1) remote sensing image and its training samples are characterized by $RSIDC$ algorithm and then find similar characters from Relationship data set by $FSDC$ algorithm, through GRA algorithm get Recommend algorithms with prediction accuracy descending order;
- 2) select corresponding algorithms and train it by raining samples to obtain a Classification model, this classification model can be used to classify the whole remote sensing image.

4 Result of experiments

Self-adaptive method of remote sensing image classification algorithms selection has two stages. In stage 1, the method need to construct relationship data set RDS ; in this study RDS will be generated through the following data set: There are two different types of the remote sensing images in the data set list as follows:

TABLE 1 The details of data which to construct RDS

Sensor type	Content	Number of sub-images
Landset TM	forest, grass and water	12
	building and farmland	20
SPOT5	wetland, grass and farmland	31
	building and road	15

Landset TM and SPOT5 are different remote sensing sensor which have different resolution, and images are further cut into sub-images with 200×200 pixel size from whole scene image, and each sub-image will be designated the training samples in manual interpretation way. Each group of sub-image and training samples are characterized by *RSIDC* Algorithm. The Classification algorithms database contains following 5 classical algorithms: Naïve Bayes, ID3 Tree, CART Tree, SVM and ANN. The entire classification algorithm classify all the sum-images (features maybe discretized for some algorithms) and obtain the corresponding classification accuracy, all the data collected and stored into database to construct *RDS*.

In order to verify the correctness of method proposed, this study has introduced a remote sensing images and cut into 10 sub-images as test data set, use 5 classification algorithm to classify and obtain the accuracy as Table 2:

TABLE 2 Test data sets classification accuracy

Data set	Classification Algorithms (%)				
	Naïve Bayes	ID3 Tree	CART Tree	SVM	ANN
1	70	<u>87</u>	84	86	86
2	83	80	80	<u>86</u>	83
3	82	85	87	<u>89</u>	87
4	75	77	76	<u>85</u>	83
5	90	<u>93</u>	<u>93</u>	<u>93</u>	<u>93</u>
6	<u>99</u>	98	98	<u>99</u>	<u>99</u>
7	85	87	88	<u>90</u>	89
8	78	<u>89</u>	81	81	82
9	60	85	<u>92</u>	90	90
10	89	90	90	<u>94</u>	90

The one marked with underlined in the table is the highest classification accuracy. Further, the self-adaptive method of remote sensing image classification algorithms selection is utilized to select algorithm, which may have highest classification accuracy.

As can be seen from Table 3, there are 6 times in 10 selection process find the best algorithm, and 3 times select the 2nd algorithm, only 1 time select the 3rd algorithm. This proved that the proposed method has good algorithm selection ability.

References

- [1] Shao Y, Lunetta R S 2012 *ISPRS Journal of Photogrammetry and Remote Sensing* **70** 78-87
- [2] Pan X, Zhang S, Zhang H, Na X, Li X 2010 *Computers & Geoscience* **36**(12) 1466-73
- [3] Wolpert D H, Macready W G 1997 *IEEE Transactions on Evolutionary Computation* **1** 67-82
- [4] Song Q, Wang G, Wang C 2012 *Pattern Recognition* **45** 2672-89
- [5] Brodley C E 1993 Addressing the selective superiority problem: automatic algorithm/model class selection *Proceedings of the 10th International Conference on Machine Learning* **1** 17-24
- [6] Gama J, Brazdil P B 1995 *Regress in Artificial Intelligence* **1** 189-200
- [7] Brazdil P B, Soares C, Da Costa J P 2003 *Machine Learning* **50** 251-77
- [8] Ho T K, Baird H S 1998 *Computer Vision and Image Understanding* **70** 101-10
- [9] Ho T K, Basu M 2002 *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(3) 289-300
- [10] Mollineda R A, Sánchez J S, Sotoca J M 2005 *Lecture Notes in Computer Science* **3523** 27-34
- [11] Baumgartner R, Somorjai R L 2006 Data complexity assessment in undersampled classification *Pattern Recognition Letters* **27** 1383-89
- [12] Luengo J, Herrero F 2010 *Fuzzy Sets and Systems* **16** 13-9

TABLE 3 Selected algorithm and its rank

Data set	Selected algorithm	Highest accuracy in Table 2	Rank
1	ID3 Tree	ID3 Tree	1
2	SVM	SVM	1
3	SVM	SVM	1
4	ANN	SVM	2
5	ID3 Tree	ID3 Tree	1
6	Naïve Bayes	Naïve Bayes	1
7	ANN	SVM	2
8	CART Tree	ID3 Tree	3
9	SVM	CART Tree	2
10	SVM	SVM	1

5 Conclusions

To select appropriate classification algorithm for a remote sensing image classification is very important, but the similar data structure of remote sensing image data and statistical information hinders the traditional algorithm selection method, this research introduced data complexity evaluation into remote sensing image classification algorithm selection field, and proposed a self-adaptive selective method of remote sensing image classification algorithms (SSMRSICADCE). The method has two stages: in stage 1, users can input a large number of remote sensing image and their training samples, method can combine remote sensing image, data character, classification algorithm, classification accuracy together and save these information into Relationship data set RDS, the RDS is the knowledge of algorithm selection; in stage 2, A self-adaptive selective algorithm mechanism was proposed with the help of RDS. Experiments indicate that the algorithms selected by this method can achieve higher classification accuracy, which provides the recommendation for the selection of appropriate classification models to users.

Acknowledgments

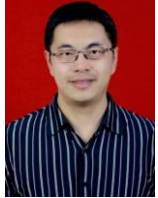
This research was supported by the National Natural Science Foundation Youth Fund of China (41101384); Natural Science Foundation of Jilin Provincial Science & Technology Department (No. 20140101178JC).

Authors



Xin Pan, born on January, 1978, Changchun City, Jilin Province, P.R. China

Current position, grades: an associate professor in the Changchun Institute of Technology, DSc
University studies: BSc degree in computer science from Changchun University of Technology (2000),
 MSc degree from Changchun University of Technology (2005),
 DSc degree in remote sensing and geographic information sciences from Chinese Academy of Sciences (2009)
Scientific interest: His research interest fields include Data mining, GIS, Remote sensing image analysis
Publications: more than 10 papers published in various journals
Experience: He has teaching experience of 8 years, has completed two scientific research projects



Hongbin Sun, born on May, 1969, Changchun City, Jilin Province, P.R. China

University studies: BSc and MSc in Electrical Engineering from Huabei Electric University, China (1991, 1997),
 Ph.D. in Electrical Engineering from Donghua University Shanghai, China (2007)
Scientific interest: His research interest fields include include complex network systems, Data mining
Publications: more than 15 papers published in various journals.
Experience: He has teaching experience of 10 years, has completed five scientific research projects.

A study and implementation on the data reduction based on the curvature of point clouds

Hong Jiang^{1,3}, Wenlei Sun^{1*}, Yongfang Shi², Yanhua Huang¹

¹ The School of Mechanical Engineering, Xinjiang University, Urumqi, China, 830047

² The School of Medical Engineering, Xinjiang Medical University, Urumqi, China, 830011

³ The School of Software, Xinjiang University, Urumqi, China, 830008

Received 6 February 2014, www.tsi.lv

Abstract

In the process of Reverse Engineering (RE), higher density of measured data from all kinds of parts with complex curved surface will not only lead to lower efficiency in computing, storing and data processing, but also affect the fairness of reconstructed surface. According to the advantages and disadvantages of common algorithms, an algorithm for data reduction is proposed in this paper, in which the neighbourhood search method based on the point cloud's curvature is used. With the utilization of proposed algorithm, high precision and the desired effect can be ensured. Finally, a roller bit's data cloud, as an example, is reduced efficiently and validly by the algorithm in this paper.

Keywords: Reverse Engineering, Data Reduction, Point Cloud, Curvature

1 Introduction

As an advanced manufacturing method, RE employs the digital technology for measurement and Three-Dimensional (3D) model reconstruction. Based on the 3D reconstruction model, the products with defects and deficiencies can be redesigned and remanufactured according to such analyses as mechanics, dynamics and so on. In the RE process, a crucial work is to obtain and process the data of objects surface.

The digital data obtained through the 3D laser or camera scanner are regarded as the point cloud. When the measurement is required by a complex curved surface, especially for some quadratic surface and cubic surface, lots of sufficient data are needed to be measured. However, due to too much sampled points on the reconstructing surface, the computer will be made too slow to implement the storage and calculation and the surface will be made less smooth. So it is necessary to reduce the point cloud data according to the characteristics and requirements of reconstruction parts [1-4, 14].

2 The common algorithm of data reduction

Chen *et al* [5] put forward a method through which the measured data can reduce data point through reducing triangular mesh generated directly by itself in 1999. Lee *et al* presented a data reduction method for laser scanning and measuring, including the minimum distance, angle deviation method and uniform grid method. The Preceding methods have the common shortcoming, that

is, difficult to determine the boundary point cloud data. Shang *et al.* [6] put forward the adaptive minimum distance method based on the above methods. By means of this method and based on the principle that the smaller minimum distance is chosen in mutation and transitional regions and the larger minimum distance is chosen in flat regions, the minimum distance is firstly chosen for data points in each zone according to the accuracy and curvature change. Thereafter, the method simplifies the data for different regions separately, which is better to keep the detailed feature of original data. However, it can be found that the effect of data processing using this method is more preferable for the smaller amount of point cloud data than for the larger one, and the efficiency and accuracy are unable to meet the design requirements. Shi in Xi'an Jiaotong University [7] presented a method of data reduction based on the remaining features, which can not only simplify point cloud data effectively but remain the features of original data. This method chooses one representative point as the original class and then classifies other data into the original class. Afterwards it traverses each class as well as replaces the class using the obtained local modal points so as to accomplish data reduction. The method is able to remain the geometrical shape of the original model for the surfaces better with larger curvature change and more additional features.

The common algorithms of data reduction include bounding box, random sampling and uniform grid method.

Bounding box algorithm starts at one point of cloud data. A certain size of rectangular bounding box is firstly established and then divided into some small cubes of

*Corresponding author e-mail: onlyxjhh@126.com

uniform size by given ruler. In each cube, the centre point of this cube is selected to replace all points in this cube. As a result, all points of this cloud point are divided and replaced.

However, the edge length of the bounding box in this method is given by user arbitrarily and all cubes are of the same size, therefore, it is impossible to ensure the reduction accuracy. As for the density data, a centre point replacing all the points in the cube will lose more important and key points with detailed characters. Therefore this method will fail when the complicated surface curvature change obviously.

Random sampling algorithm is relatively simple. It is easy to be realized as following. A random function that can represent all cloud data points needs to be found first. Meanwhile this function can generate random numbers at the range of point cloud data. Then, the number of points that the point cloud will be left is assumed. Certain random numbers produced by the random function which represent some points are deleted next. Repeat this step until the number of the left points is reduced to a given value. This algorithm is simple and easy to be implemented. The efficiency of the algorithm is higher. However, the shortcoming of this algorithm is more obvious. When massive amount of data are reduced, the reduction results are close to a uniform simplification because of its high randomness. The precision of reduction results is uncontrollable and the uniformity of the reduced data will be reflected. In the meantime, it is difficult to reconstruct a 3D mode due to the possibility of presented hole.

For uniform grid algorithm, it assigns all data points into the corresponding grids based on “median filter” principle. Then a median point to replace all points in this grid was selected. This is an improved bounding box algorithm in fact.

The uniform grid method can overcome some shortcomings of the spline curve, the same grid size and the grid that is divided too small, easily result in producing some empty grids, and a lack of the flexibility for capturing the shape will lead to a waste of time and space.

Considering the large quantity of dense points and complex surfaces, the above traditional algorithms are not good at processing the point cloud data since the features of some key positions could be lost. The subsequent surface modelling could also be affected and the precision of model could not be guaranteed. The main reason is that different degrees of data reduction cannot be determined according to the different surface features.

2.1 CURVATURE-BASED ALGORITHM OF DATA REDUCTION

Because of the insufficiencies of traditional algorithms, the algorithm of curvature has been drawn a lot of attention by researchers and a number of research achievements have been made including neighbourhood

establishment, curvature estimation and principles for reducing data [8, 9]. There are still deficiencies in these research achievements in some degree. For example, in the principles of data reduction, some mistakes appear constantly by using the methods of minimum distance and angular deviation. Therefore, this paper proposes a curvature algorithm based on paraboloid fitting, which is suitable neighbourhood search of even or uneven point cloud. In the described algorithm, the mean value of the curvature is regarded as a criterion for data reducing and data subdivision in which the data are uneven in high curvature region. The algorithm is good at dealing with the point cloud with complex features and high curvature.

2.2 CURVED SURFACE AND CURVATURE

Curvature is a basis for measuring the uneven degree of geometry and reflects important features of the surface. The normal curvature at the main direction of one point on the surface is regarded as the principal curvature of the point.

Suppose a surface $S:r=r(u,v)$, make $v=v_0$ and fix it, let u change, then a curve, named u -curve, will be drawn by $r=r(u,v_0)$; Likewise, make $u=u_0$ and fix it, let v change, then a curve, named v -curve, will be drawn by $r=r(u_0,v)$. Curve u and Curve v form a coordinate network of curved lines. The mesh consisting of the line of curvature is a coordinate network of curved lines. Setting the principal curvatures along the “line” as k_1 , the principal curvatures along the “line v ” as k_2 and the intersection angle between the random direction of surface, $d=du:dv$, and the curve as θ , then k_n , the normal curvature along (d) , satisfies the Euler’s formula: $k_n = k_1 \cos^2 \theta + k_2 \sin^2 \theta$. Setting $(d)= du:dv$ as the main direction of curved surface $s: \vec{r} = \vec{r}(u,v)$ at point P, the principal curvatures along the main direction as k_N , then the computational formula of k_N is:

$$\begin{vmatrix} L - k_N E & M - k_N F \\ M - k_N F & N - k_N G \end{vmatrix} = 0. \tag{1}$$

That is:

$$(EG - F^2)k_N^2 - (LG - 2MF + NE)k_N + (LN - M^2) = 0, \tag{2}$$

where E, F, G are first type elements of surface and L, M, N are second type elements of surface.

Suppose k_1, k_2 are two principal curvatures of a point on the surface, the $k_1 k_2$ is called Gaussian curvature of the point and recorded as K , that is $k = k_1 k_2$; the average of the curvatures is called mean curvature of this point on the surface and recorded as H , that is:

$$H = (k_1 + k_2) / 2, \tag{3}$$

According to the formula of principal curvature and Wada's theorem, the formulas of Gaussian curvature are shown as:

$$K = \frac{LN - M^2}{EG - F^2} \text{ (Gaussian curvature),} \tag{4}$$

and

$$H = \frac{LG - 2MF + NE}{2(EG - F^2)} \text{ (Mean curvature).} \tag{5}$$

2.3 NEIGHBOURHOOD SEARCH

There are several method in solving neighbourhood, including line-by-line searching method, neighbourhood ball searching method [11], 3D grid searching method [12] and the octree searching method [13] etc. Among them, line-by-line searching method is only suitable for the point cloud with obvious scanning line characteristics, such as point cloud by laser scanning.

For neighbourhood ball searching method, a good deal of time and internal storage are occupied during searching neighbourhood ball of every point according to the radiuses. In addition, this method has the following problems while dealing with non-uniform scattering point cloud. Since some neighbourhoods probably have no point or fewer than three points in small density point area, the subsequent curvature estimation could not be accomplished. However, in large density point area, some neighbourhoods may have too many points, which will slow the computational speed.

$$M = \left\lceil \frac{x_{box}}{cube_size} \right\rceil, N = \left\lceil \frac{y_{box}}{cube_size} \right\rceil, L = \left\lceil \frac{z_{box}}{cube_size} \right\rceil, \tag{8}$$

and its edge length "cube_size" is:

$$cube_size = \sqrt{2 \times (x_{box} \times y_{box} + x_{box} \times z_{box} + y_{box} \times z_{box}) \times \frac{NUM}{N}}, \tag{9}$$

where N is the number of points of this cloud point, the NUM is the expected number of cubic grid contained points.

In practice, the measured points are non-uniformly distributed. Therefore, the number of cubic grid is not

$$cube_size = \beta \sqrt{2 \times (x_{box} \times y_{box} + x_{box} \times z_{box} + y_{box} \times z_{box}) \times \frac{NUM}{N}}. \tag{10}$$

Next, a linked list array $List[i][j][k]$ is established. According to the index numbers of the data points in the X, Y, Z axial directions, the points are inserted into this

For 3D grid searching method, firstly, the measured data must be read-in one-dimensional array. Then the maximum value and minimum value of the data on the coordinate axis of X, Y, Z could be found. The minimum space bounding box and the three edges length are constructed as following:

$$\begin{cases} x_{box} = x_{MAX} - x_{MIN} \\ y_{box} = y_{MAX} - y_{MIN} \\ z_{box} = z_{MAX} - z_{MIN} \end{cases}, \tag{6}$$

where x_{max} is max and x_{min} is min in X direction; y_{max} is max and y_{min} is min in Y direction; z_{max} is max and z_{min} is min in Z direction. x_{box}, y_{box} and z_{box} are edges length in X, Y, Z direction respectively.

In order to ensure that all data points are included in the minimum space bounding box, the (6) is revised as:

$$\begin{cases} x_{box} = (x_{MAX} - x_{MIN}) \times 1.1 \\ y_{box} = (y_{MAX} - y_{MIN}) \times 1.1 \\ z_{box} = (z_{MAX} - z_{MIN}) \times 1.1 \end{cases}. \tag{7}$$

Then the minimum bounding box is divided into $M \times N \times L$ small cube grids and the length of edge is "cube_size". The numbers of points included in each cube grid are the same. The number of cubic grids in $X, Y,$ and Z axis direction are:

necessarily the expected number NUM . Hence for calculating the edge length of cubic grid, an adjusting coefficient β is added. The edge length is adjusted dynamically:

linked list. For example, the point $P_i=(x_i, y_i, z_i)$, and its index number are:

$$i = \left\lfloor \frac{x_i - x_{MIN}}{cub_size} \right\rfloor; \quad j = \left\lfloor \frac{y_i - y_{MIN}}{cub_size} \right\rfloor; \quad k = \left\lfloor \frac{z_i - z_{MIN}}{cub_size} \right\rfloor, \quad (11)$$

and the point P_i is inserted into $List[i][j][k]$.

After the list of cube grid is established, searching for the nearest neighbour points of each sample point would continue and the number of nearest neighbour points is k . While establishing a nearest neighbour points linked list for every sample point, the distance from the neighbour point to base point is calculated, the point into the nearest neighbour points list is added based on the distance value. The method for searching a nearest neighbour points is, firstly its index number is calculated. Secondly, its nearest neighbour point is searched in the cube grid including itself. But this cubic grid can't be considered alone since no matter how big this cube grid is, the neighbour points of sample point are likely to exist in the grid adjacent to this grid. For example, when a point is located on the faces of the cube, then its neighbour points may be found in the other cube which has the same face with this cube. Therefore, when searching the k neighbour points of a sample points, its twenty six adjacent cubes must be searched.

In the searching process, an adjacent ball is used to describe the current k nearest neighbour points. The adjacent ball is updated if the nearest neighbour is found. Thus, this adjacent ball is always the smallest one that contains the searched k neighbouring points. The process

starts with the searching of the sample points in the cubic grids. Then it searches the points in the nearest neighbour grid. When a grid is identified, the cubic grid disjointing with the current adjacent ball will be removed from the grid list. In the meantime, the identified grid will also be excluded from the grid list after searching. Likewise, the searching will be continued until the grid list becomes empty. Such a method guarantees the minimum number of checked cubic grids and the improved computational efficiency.

Therefore, the path of an octree for seeking space point can be obtained by interpreting the position codes in the process of subspace decomposition and transforming decomposition range each time. Due to the direct relationship of position codes and the position of cubes, it can be specified as the position code of a child node on X axis plus 1 relative to the adjacent nodes on the left side, the one on Y axis plus 2 relatives to the adjacent nodes on the bottom, and the one on Z axis plus 4 relatives to the adjacent nodes on the rear side.

According to the encoding characteristics of the octree, each node of the tree is solely corresponding to a 8-hexadecimal number coded in binary form based on the coordinates of nodes in the divided space. These coordinates are calculated as follows:

$$\begin{cases} x = a_{n-1}2^{n-1} + a_{n-2}2^{n-2} + \dots + a_k2^k + \dots + a_12^1 + a_02^0 \\ y = b_{n-1}2^{n-1} + b_{n-2}2^{n-2} + \dots + b_k2^k + \dots + b_12^1 + b_02^0 \\ z = c_{n-1}2^{n-1} + c_{n-2}2^{n-2} + \dots + c_k2^k + \dots + c_12^1 + c_02^0 \end{cases}, \quad (12)$$

where $a_k, b_k, c_k \in \{0,1\}$, and $k \in \{0,1,\dots,n-1\}$.

If the number of a cube uniting the octree were known, the coordinates can be rewritten as:

$$\begin{cases} x = \sum_{i=1}^{n-1} (q_i \bmod 2) \times 2^i \\ y = \sum_{i=1}^{n-1} ((q_i / 2) \bmod 2) \times 2^i \\ z = \sum_{i=1}^{n-1} ((q_i / 4) \bmod 2) \times 2^i \end{cases}, \quad (13)$$

where n is the depth of the current node.

The following equation (14) can be obtained based on the numbers of cube units in the octree space:

$$q_i = c_i2^2 + b_i2^1 + a_i2^0. \quad (14)$$

Likewise, if one sub-cube unit number can be known, its coordinate can be calculated inversely using equation

(13) and (14). Assume the relative coordinate of the node in the final bounding box is (x,y,z) , 26 relative coordinates of the adjacent minimum bounding box can be represented by the following equation considering the division characteristics of a bounding box space:

$$\begin{cases} x' = x \pm \delta \\ y' = y \pm \delta \\ z' = z \pm \delta \end{cases}, \quad (15)$$

where $\delta \in [0,1]$; x', y', z' and x, y, z cannot be equal simultaneously.

Moreover, for such isolated and sparse noisy points, Due to the far distance from other points, their bounding boxes for the isolated and sparse noisy points are in relatively isolated positions when the divided bounding box space or the bounding box contains fewer points. Given a threshold value T , any points outside the centre with a value less than T are identified. The points contained in the bounding box are noisy points and thus

can be removed to avoid their impact on the following work.

Further, several issues associated with the partition of the space with bounding box are noticed: (a) the positions of some bounding boxes are relatively isolated; (b) the number of points in the bounding box is less because of sparsity, and (c) isolated noisy data points, which are far away from other data collection. In order to solve these issues, a threshold value “T” is used. If the identified points are less than the appointed value after the outward expansion of “T” times, the points in the bounding box are treated as noisy points that can be deleted.

Based on aforementioned information, this paper proposes an approach that is applicable to search in the neighbourhood of uniform and non-uniform distribution point clouds.

Firstly, meshing the point clouds and searching in the grids that potentially will lead to avoid calculating the distance between P_i and other points. Secondly, the maximum and minimum neighbour points are denoted as P_i . In the grid that includes the point P_i , once the number of points is less than the minimum, one is added to the length of a side of the grid in order to expand the grid until at least “min” points are searched. If the number of points in the grid including the point P_i is greater than the max, then the value is kept at the closest to the max. Thus, the extreme conditions with either too many or very limited number of neighbour points can be avoided. The searching process is shown as Figure 1.

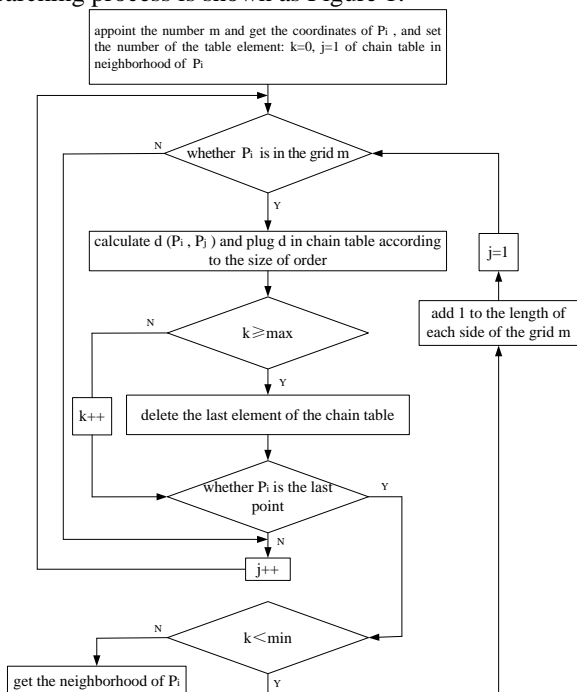


FIGURE 1 T the process diagram of searching neighbourhood

2.4 CURVATURE ESTIMATION

Curvature estimation, an important algorithm, has become the thematic topic in areas such as computer vision, computer graphics, geometric modelling and

bioengineering. The frequently used methods of curvature estimation are: paraboloid fitting, circle fitting, Gauss-Bonnet, Watallable & Belyaev, Taubin, etc. Among these methods, Gauss-Bonnet can gain optimal Gauss curvature; paraboloid fitting can obtain suboptimum Gauss curvature as well as optimal average curvature. Particularly, paraboloid fitting has optimal stability in the neighbourhood. Thus, paraboloid fitting is adopted in this paper.

Set the paraboloid equation as:

$$z = ax^2 + bxy + cy^2, \tag{16}$$

where x, y and z are three coordinates of points on parabolic, and a, b and c are the coefficients of the equation.

There are k points in neighbourhood of P_i .

The problem to obtain the coefficients a, b and c by the least square paraboloid fitting for P_i and its neighbourhood essentially is a linear least square problem. Householder converter technique can be used to solve the linear equations. The process to estimate the curvature of arbitrary falls mostly into four steps:

Step 1: Plug P_i and the points in its neighbourhood into (16). Then the equation set is obtained as follows:

$$AX=Z, \tag{17}$$

where A is the transformation matrix composed of P_i and K points' x and y coordinates in the neighbourhood of P_i :

$$A = \begin{pmatrix} x_1^2 & x_1y_1 & y_1^2 \\ x_2^2 & x_2y_2 & y_2^2 \\ \vdots & \vdots & \vdots \\ x_{k+1}^2 & x_{k+1}y_{k+1} & y_{k+1}^2 \end{pmatrix}_{(k+1) \times 3}. \tag{18}$$

X is the transposed matrix of the parabola coefficients:

$$X = \{a, b, c\}^T. \tag{19}$$

Z is the transposed matrix consists of P_i and K points' z coordinate in the neighbourhood of P_i :

$$Z = [z_1, z_2, \dots, z_{k+1}]_{k+1}^T. \tag{20}$$

Step 2: Solve (18) and (17) using Householder converter technique and obtain the coefficients a, b , and c of the paraboloid equation.

According to the properties of parabola, we can obtain:

$$K = 4ac - b^2 \quad (\text{Gaussian curvature}), \tag{21}$$

and

$$H = a + c \quad (\text{Mean curvature}). \tag{22}$$

Step 3: Repeat steps 1 and 2 until the Gaussian curvatures and the average curvatures of all points are obtained.

2.5 IMPLEMENTATION OF THE ALGORITHM

It is common that the point clouds are reduced according to the curvature when reconstructing their surface. The basic principle is to keep a small number of points in a small curvature zone and sufficient points in large curvature zone for reserving the detailed feature of the surface. The method cannot only reduce the number of the points effectively but also keep the features of the surface accurately. In this paper, a module for reducing point clouds based on the curvature has been developed, and the procedure of the algorithm is described below.

Firstly, divide the value of the curvature into many intervals and set the deviation ϵ based on the curvature. T special condition that the value of the curvature approaches zero (that is, the curve is similar to the line) must be considered.

Secondly, setting the curvature deviation as ϵ within one interval. If the point P_i meets the criteria deviation: $|H_j - H_i| \leq \epsilon$ (H_j and H_i are the average curvature of P_j and P_i , respectively) then delete the point P_j . Otherwise, keep the point P_j and set it as the datum point. Repeat this process. This principle cannot only reduce the quantity of the point clouds but also retain the geometrical features of the point clouds better.

The specific realization process is to: (1) search the neighbourhood of each point in point clouds; (2) estimate the curvature in neighbourhood while searching; and (3) reduce the quantity of the points according to the reduced principle in the condition of ensuring the accuracy. The process is illustrated in Figure 2.

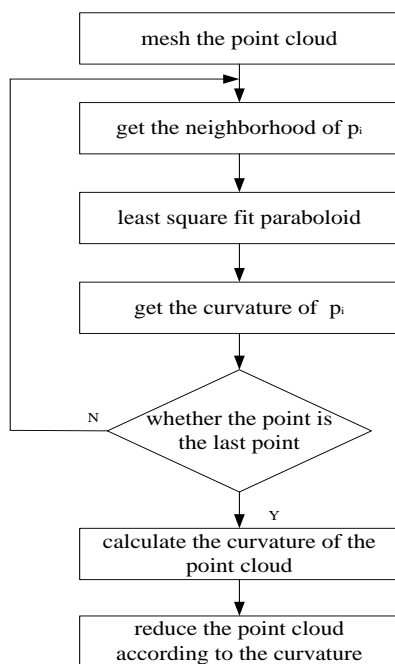


FIGURE 2 The process diagram of reducing point clouds

3 Reducing the point cloud of roller bit

In this paper, the method mentioned above is used to reduce the point clouds of roller bit, and the result is shown as Figures 3 and 4.



FIGURE 3 The point clouds of roller bit before reducing

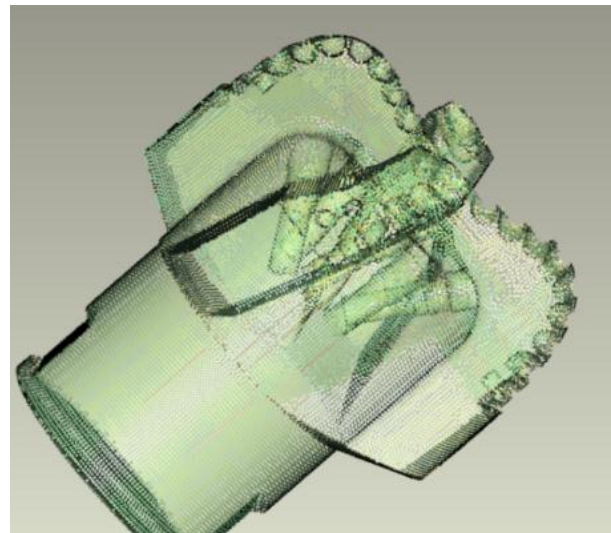


FIGURE 4 The point clouds of roller bit after reducing

4 Conclusions

For parts with complex surface in RE process, if the density of metrical data is high, the efficiency of computer is relative low in terms of running, storage and operation. In addition, the smoothness of the reconstructed surface is affected.

This paper focuses on the pros and cons of common algorithms so as to reduce point clouds. It also proposes a new reducing algorithm based on the curvature change.

The case study illustrates that the proposed method performs well in reducing the point clouds with complex surface. Detailed features of the original data are better maintained on the premise of guaranteeing the accuracy and the quantity of the point clouds. These features are very important for further reconstructing surface.

Acknowledgements

This study is funded by the joint project of school and university of Xinjiang (Grant No. 201204361319) and the Science and Technology Supporting Project of Xinjiang

Uyghur Autonomous Region (Grant No.: 201304071393) and Science and Technology Major Special Project of Xinjiang Uyghur Autonomous Region (Grant No. 201204071134).

References

- [1] Filip D, Magedson R, Markot R 1986 *Computer Aided Geometric Design* 3(4) 295-311
- [2] Barbero B R 2009 *Computers and Industrial Engineering* 56(4) 1265-75
- [3] Liu Tao, Xu Zheng, Sha Chengmei and Zhao Juntian 2009 Curvature estimation of scattered-point cloud data based on bounding box method *Science Technology and Engineering* 9(12) 3333-6 (in Chinese)
- [4] Wan Jun, Ju Lu-yue 2004 Data Reduction for Reverse Engineering *Journal of shanghai university (natural science)* 10(1) 63-69 (in Chinese)
- [5] Chen Y H, NG CT, Wang Y Z 1999 Data reduction in integrated reverse engineering and rapid prototyping *International Journal of Computer Integrated Manufacturing* 12(2) 97-103 (in Chinese)
- [6] Shangguan Jian-lin, Guo San-ci 2011 A study on data reduction in Reverse Engineering *Mechanical Management and Development* 122(4) 16-18&20 (in Chinese)
- [7] Shi Baoquan, Liang Jin 2010 Research on point cloud simplification with preserved features *Journal of Xi'an Jiaotong University* 44(11) 37-40 (in Chinese)
- [8] Ma Juan, Fang Yuanmin, Zhao Wenliang, Feng Yujin 2011 Algorithm for finding k-nearest neighbors based on spatial sub-cubes and dynamic sphere *Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomatics and Information Science of Wuhan University* 36(3) 358-62 (in Chinese)
- [9] Vanco M, Brunnett G, Schreiber Th 1999 A hashing strategy for efficient k-nearest neighbors computation *Proceedings of the International Conference on Computer Graphics, Canmore, Alta., Canada (7-11 June 1999)* 120-8
- [10] Goodsell G 2000 *Computer Aided Geometric Design* 17(4) 387-92
- [11] Fu Jing, Joshi S B, Simpson T W 2008 *Computer-Aided Design* 40(3) 311-23
- [12] Zhu Maomao, Feng Zujun, Guo Jianhua, Ju Luyue 2005 Data point reduction using a 3-Dimensional cube algorithm *Journal of Shanghai University(Natural Science Edition)* 11(3) 242-6 (in Chinese)
- [13] Liu Chun-ming, Fang Yi 2009 An algorithm of searching topological structure from 3D unorganized points *Journal of QingDao University* 18(3) 20-4 (in Chinese)
- [14] Wei Chen, Jiayi Shen, Jiabin Xu 2013 Study on a robot 3D laser radar information collection system *Journal of Digital Information Management* 11(1) 69-75(in Chinese)

Authors

	<p>Hong Jiang, born on December, 1976, Urumqi, Xinjiang, P.R. China</p> <p>Current position, grades: the Associate Professor of School of Mechanical Engineering, Xinjiang University, China. University studies: B.Sc in Mechanical Engineering from Xinjiang Institute in China, MSc from Xinjiang University in China Scientific interest: digitized manufacture, RE Publications: more than 30 papers Experience: He has teaching experience of 15 years, has completed ten scientific research projects.</p>
	<p>Wenlei Sun, born on October, 1962, Urumqi, Xinjiang, P.R. China</p> <p>Current position, grades: the Professor of School of Mechanical Engineering, Xinjiang University, China. University studies: BSc in Mechanical Engineering from Xinjiang Institute in China, MSc from Huazhong University of Science and Technology in China. Scientific interest: His research interest fields include CAD/CAM, RE Publications: more than 250 papers Experience: He has teaching experience of 30 years, has completed eighty scientific research projects.</p>
	<p>Yongfang Shi, born on October, 1978, Urumqi, Xinjiang, P.R. China</p> <p>Current position, grades: the Associate Professor of School of Mechanical Engineering, Xinjiang Medical University, China. University studies: BSc in Mechanical Engineering from Xinjiang Institute in China, MSc from Xinjiang University in China Scientific interest: Her research interest fields include biomanufacturing engineering, RE Publications: more than 16 papers Experience: She has teaching experience of 15 years, has completed eight scientific research projects.</p>
	<p>Yanhua Huang, born on February, 1978, Urumqi, Xinjiang, P.R. China</p> <p>Current position, grades: the Associate Professor of School of Mechanical Engineering, Xinjiang University, China. University studies: BSc in Mechanical Engineering from Xinjiang Institute in China, MSc from Huazhong University of Science and Technology in China Scientific interest: Her research interest fields include digital design, RE Publications: more than 40 papers Experience: She has teaching experience of 15 years, has completed ten scientific research projects.</p>

Multi-objective improved algorithm for flow allocations in hazardous chemicals logistics preference paths

Xin Zhang*

Jiaxing University, Jiaxing City, Zhejiang Province, China, 314200

Received 6 October 2013, www.tsi.lv

Abstract

The flow allocation of paths was a key stage of the transportation network's efficiency, particularly in the hazardous chemicals logistics network where many weights were stochastic. Over the years, a variety of methods (or heuristics) have been proposed to solve this complex optimization problem, with good results in some cases just with limitations in the special fields. In this work, we develop an algorithm for model multi-objective that combines ideas from stochastic weight. Our method performs well even when the order of magnitude and/or the range of the parameters were unknown. The method refines iteratively a sequence of parameter distributions through preference combined with partial exempling from a historical prior defined over the support of all previous iterations. We exemplify our method with multi-objective improved models using both simulated and real experimental data and estimate the weight efficiently even in the absence of a priori knowledge about the weight.

Keywords: Hazardous chemicals transportation, Flow allocation, Multi-objective optimization, Path preference, Control

1 Introduction

The efficiency and benefit of logistics network is determined directly by network topology and logistics nodes of the network and flow allocation of the transportation lines. But once the topology of logistics network is identified, it will not be changed in a fairly long period of time. However, the flow allocation in the logistics network will be influenced by many factors such as supply, demand and network capacity [1]. In the process of the flow allocation of logistics network, we often not only needs to consider cost of logistics, also the delivery time, the delivery distance, the reliability of the network and other factors, and also the preference of decision-makers (this issue is an important reason that causes the transportation problem as NP-hard problem [2, 3]). Therefore, how to realize the scientific and reasonable flow allocation in the hazardous chemicals logistics network of a fixed topology structure of is the overall optimal scheme of flow allocation of logistics network with the prerequisites of logistics demand such as goods transportation, storage and distribution from supply to demand of the goods. The goals of this problem are often coupled together, but also in competition with each other, contradictory, and each goal has its distinct significances and dimensions, and their competition complexity makes the optimization of the problem very difficult. Therefore, multi-objective optimization of logistics transportation is always a hotspot in the field of investigation [4, 5].

Considering the existing limitations of multi-objective algorithm in the flow allocation of logistics network, this

study considers the preference of decision-makers and the comprehensive effect of the target value on the objective weight based on weighting method, and puts forwards the indefinite traffic assignment method. The flow assignment method of logistics network needs not to find all the optimal solution, but by adjusting preference coefficient, we can obtain the traffic assignment results in accordance with the actual needs of the flow allocation decision.

The flow allocation in logistics network is that the desired goods are distributed to the logistics nodes and the transportation lines of logistics network in order to achieve goods delivery from supply to demand. This study will take delivery task as a path connecting the supply and demand nodes of multi-strip, only need the allocation flow of each path to make the target value achieve the optimal, and guarantee the flow balance in all the transportation lines and the logistics nodes. The process is divided into three stages, i.e. the calculation of the path state value based on multiple optimization targets, the selection of optimal path and the flow allocation. The path is composed of a series of logistics nodes and transportation lines in order form. Path state value (weight) is determined by the logistics nodes and the state values of transportation line. According to the contingency theory, each logistics nodes and transportation lines of the path should not be absolute equality but fixed which are relative to the weight of the whole path. Therefore, the logistics nodes and the state value of transportation lines and weight becomes the critical for path selection; similarly, logistics nodes and the state value calculation of transportation lines should

*Corresponding author e-mail: zcxin9@163.com

be based on the variable weight synthesis of multiple attribute of optimization objectives definition, not the constant weight sum method. According to this principle, we can avoid the NP-hard in the logistics activities to some extent.

2 Model and algorithm

Logistics nodes, transport routes and paths are collectively referred to as logistics network factors, in order to solve the problem easily; its expression form is [6]: $\Phi_l = \{u_\alpha, v_\alpha(\Phi_l), w_\alpha(\Phi_l)\}, \alpha, l \in N$.

Among them, Φ_l expresses No. l factor in the transportation network, u_α expresses No. property of Φ_l factor, $v_\alpha(\Phi_l)$ expresses No. α state of u_α property in Φ_l factor, $w_\alpha(\Phi_l)$ expresses No. α weight values (preference) of u_α property in Φ_l factor, and $\sum_\alpha w_\alpha(\Phi_l) = 1$.

The essence of flow allocation is to select the optimal path, i.e. optimal path of state value. A mathematical model:

$$\begin{aligned} \text{Max} \quad & F(x) = (f_1, f_2, \dots) \\ \text{s.t.} \quad & G(x) \leq 0 \end{aligned} \tag{1}$$

Among them, f_1, f_2, \dots as optimization object, X as decision variables.

From the expression of path factor, we can select the path according to the state value and distribute the flow to the optimal path of state value. Calculation steps of state values for attributes:

- Step 1:** determine the path attribute and its initial weights;
- Step 2:** calculate the attribute state value and its preference of logistics nodes and the transportation lines;
- Step 3:** determine the state value of logistics nodes and the transportation lines taking optimization target as attributes;
- Step 4:** calculate the attribute state value of sub path and its preference;
- Step 5:** determine the state value of the sub path taking logistics nodes and the transportation line as attributes;
- Step 6:** calculate the path attribute values and its preference;
- Step 7:** determine the state values of the path taking sub path as attributes.

Because the attributes of all the logistics network factors are based on type (1) multi optimization objectives of problem, but the optimization target values

have different dimensions. So, firstly, the optimization target values should be transformed into dimensionless quantity in the [1] interval and then conduct standardized processing. By using linear membership function, we can achieve the dimensionless transformation of optimal target values [7]. Membership function of optimization object represents the closeness between function of optimization target and the possible ideal values.

v^+ is defined as the upper bound of the state value, v^- as the lower bound of the state value, there is: $v_\alpha^+ = \max(v_\alpha(\Phi_l)), v_\alpha^- = \min(v_\alpha(\Phi_l)), \alpha, l \in N$.

According to the linear membership functions, the dimensionless and standardization process of the properties are [8-9]: $z_\alpha = \frac{v_\alpha(\Phi_l) - v_\alpha^-}{v_\alpha^+ - v_\alpha^-}$.

The factor attribute of the transportation network after normalization is:

$$v_\alpha(\Phi_l)' = \sum_\alpha (z_\alpha * w_\alpha(\Phi_l)). \tag{2}$$

By formulas (1) and (2), we can obtain the steps of freight traffic allocation algorithm of path preference.

Step 1: initialization. The initial volume of freight X_0 of the lines and nodes of logistics network is 0; the capacity is Y_{\max} .

Step 2: path selection. By formula (1) (2) we can calculate all the paths state value, and select the state optimal path Q^* (n as the network path number):

$$Q^* = \max(v_{\alpha\beta}^1, v_{\alpha\beta}^2, \dots, v_{\alpha\beta}^n). \tag{3}$$

Step 3: the traffic volume allocation. For Q^* , distribute freight volume X^0 , and $X^0 \leq Y_{\max}$.

Step 4: update the path freight volume and capacity, there is:

$$X^1 = X_0 + X^0, Y^1 = Y_{\max} - X^0. \tag{4}$$

Step 5: judge whether the goods are distributed or not, if yes, then ending. Otherwise, change to **Step 2**, until the freight task was all arranged.

3 Algorithm demonstration

As shown in Figure 1, the logistics transportation network is a simple logistics network which is composed of 2 supply nodes, 2 intermediate nodes, 3 demand nodes, and transportation lines among the logistics nodes [8] (each node and route preference coefficient is

$(\beta_1, \beta_2, \beta_3) = (-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})$, i.e. no difference in preference.)

As we know the capacity, the time, the cost attribute parameters of the network, the value after standardized processing is shown in Table 1. V is the rated capacity of unit carrier, C0 is the fixed cost, C1 is the variable cost.

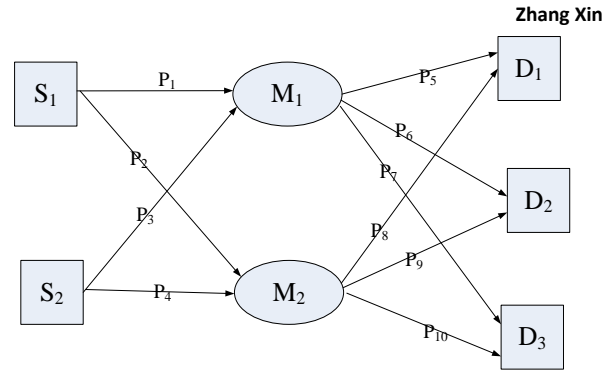


FIGURE 1 Transportation network

TABLE 1 The attribute parameters of transport network

Attribute	Nodes							Paths									
	S1	S2	M1	M2	D1	D2	D3	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
capacity (Y)	60	80	90	150	70	55	15	2	3	4	3	2	1	1	3	2	1
time (t)	-	-	0.5	0.4	-	-	-	89	62	94	77	10	7	14	18	11	6
cost (c)	-	-	1.3	1.5	-	-	-	V=20, C0=1, C1=2									

According to equation (1) and the corresponding algorithm, we can obtain:

$$V_s(y) = (v^{s_1}, v^{s_2});$$

$$V_d(y) = (v^{d_1}, v^{d_2}, v^{d_3});$$

$$V_m(y) = (v_y^{m_1}, v_y^{m_2}, v_y^{m_3});$$

$$V_m(t) = (v_t^{m_1}, v_t^{m_2}, v_t^{m_3});$$

$$V_m(c) = (v_c^{m_1}, v_c^{m_2}, v_c^{m_3}).$$

According to equation (2), we can carry out the dimensionless and standardization. The results are presented in Table 2.

Similarly, the attribute and weights of the network path factor can be obtained (see, Table 3).

By equations (3) and (4), we obtain

$$Q_1^* = (s_2, p_4, m_2, p_8, d_1),$$

$$x_1 = \min(y^{s_2}, y^{p_4}, y^{m_2}, y^{p_8}, y^{d_1}) = 60$$

TABLE 2 Attribute value and weight of network node factors

Path factor	y			t			c		
	v	z	w	v	z	w	v	z	w
s1	60	0.75	0.50	-	-	-	-	-	-
s2	80	1.00	0.50	-	-	-	-	-	-
m1	90	0.60	0.52	0.02	0.80	0.50	0.01	1	0.55
m2	150	1.00	0.48	0.03	1.00	0.50	0.01	0.87	0.45
d1	70	1.00	0.33	-	-	-	-	-	-
d2	55	0.79	0.34	-	-	-	-	-	-
d3	15	0.21	0.33	-	-	-	-	-	-

TABLE 3 The path attribute values and weight

Paths	y			t			c		
	v	z	w	v	z	w	v	z	w
p1	0.75	0.95	0.076	0.66	0.83	0.35	0.79	1.00	0.096
p2	0.75	0.79	0.081	0.91	0.95	0.33	0.95	1.00	0.096
p3	1.00	1.00	0.073	0.80	0.80	0.34	0.79	0.79	0.105
p4	1.00	1.00	0.074	0.81	0.81	0.35	0.95	0.95	0.099
p5	0.72	0.72	0.125	1.00	1.00	0.47	0.75	0.95	0.099
p6	0.68	0.87	0.118	0.79	1.00	0.49	0.75	0.79	0.105
p7	0.51	1.00	0.108	0.21	0.42	0.54	1.00	1.00	0.093
p8	0.67	0.67	0.125	1.00	1.00	0.46	0.66	0.83	0.105
p9	0.69	0.88	0.118	0.77	1.00	0.49	0.91	0.95	0.099
p10	0.73	1.00	0.102	0.21	0.30	0.56	0.80	0.80	0.103

Similarly we have:

$$Q_2^* = (s_1, p_2, m_2, p_9, d_2), x_2 = 40;$$

$$Q_3^* = (s_2, p_3, m_1, p_6, d_2), x_3 = 15;$$

$$Q_4^* = (s_1, p_2, m_2, p_{10}, d_3), x_4 = 15;$$

$$Q_5^* = (s_2, p_3, m_1, p_5, d_1), x_5 = 5;$$

$$Q_6^* = (s_1, p_1, m_1, p_5, d_1), x_6 = 5.$$

When the preference of the transportation network is identical, the final result of freight flow allocation is: $(Q_1^*, Q_2^*, Q_3^*, Q_4^*, Q_5^*, Q_6^*) = (60, 40, 15, 15, 5, 5)$.

4 Conclusions

Based on the method of the preference of decision-makers, this study integrates the objective evaluation values into state values of the hazardous chemicals logistics nodes and the transportation lines in stages, and puts forward the methods of goods flow allocation based on the path preference, which not only considers the influence of the optimization objectives on the decision

results, but also the effect of each hazardous chemicals transportation network factor on the decision results, and synthesizes the preference of subjective decision-makers and influence of objective state value on decision-making, to make allocation of freight task more in line with the actual need to the decision. The method of flow allocation based on path preferences does not need to generate multiple optimal solutions, but through adjusting the preference we can obtain the optimal flow allocation scheme that conforms to decision-makers' expectation. The study can solve the NP problem of flow allocation in the hazardous chemicals logistics network.

References

- [1] Liu Nan, Chen Yuangao, Li Yumin 2007 Comprehensive decision making method for optimal location of logistics hubs *Journal of Southeast University* 23(S1) 71-5 (in Chinese)
- [2] Sun Huijun, Gao Ziyou, Wu Jianjun 2008 A bi-level programming model and solution algorithm for the location of logistics distribution centers *Applied Mathematical Modelling* 32(4) 610-6
- [3] Lee Der-Horng, Dong Meng 2008 A heuristic approach to logistics network design for end-of-lease computer products recovery *Transportation Research Part E: Logistics and Transportation Review* 44(3) 455-74 (in Chinese)
- [4] Melo M T, Nickel S, Saldanha-da-Gama F 2009 *European Journal of Operational Research* 196(2) 401-12
- [5] Moon I, Kang S 1998 Rationing policies for some inventory systems *Journal of the Operational Research Society* 49(5) 509-18
- [6] Li Chunyan, Tang Siyuan 2013 Application of grey relational in analysis work satisfaction of nursing personnel and influencing factors *Applied Mathematics and Statistics* 48(18) 99-107
- [7] Rizvi Ahsan Z, Bhattacharya C 2012 An efficient algorithm and schematic for computation of GC, GC3, and AT3 bias spectra *International journal of artificial intelligence* 9(A12) 19-25
- [8] Jianjun Xu, Shaoxiang Chen, Bing Lin, Rohit Bhatnagar 2010 Optimal production and rationing policies of a make-to-stock production system with batch demand and backordering *Operations Research Letters* 38(3) 231-5
- [9] Saif Benjaafar, Mohsen ElHafsi, Tingliang Huang 2010 Optimal control of a production-inventory system with both backorders and lost sales *Naval Research Logistics* 57(3) 252-65

Author



Xin Zhang, born on November, 1973, Jiaxing County, Zhejiang Province, P.R. China

Current position, grades: the lecturer of Jiaxing University, China.

University studies: received his Ph.D.Sc. in Transportation planning and management from Wuhan University of technology in China. He received his M.Sc. from Wuhan University of technology in China.

Scientific interest: His research interest fields include Transportation, Logistics, et al.

Publications: more than 10 papers published in various journals.

Experience: He has teaching experience of 10 years, has completed three scientific research projects.

Spatially aware in implicit human robot interaction

Wei Wang*, Xiao-dan Huang

School of Information & Electrical Engineering Hebei University of Engineering, Handan, Hebei Province, China, 056038

Received 1 March 2014, www.tsi.lv

Abstract

Implicit interaction pattern between the human user and the robot is important for reducing cognitive burden and enhancing cooperation effect. Given that the spatially aware is a foundation for human-robot cooperation, for the existing researches of robots, in this paper, a reachable space for a serial robot arm with a fixed monocular vision system and five degrees of freedom (5-DoF) was built. Based on the link frame with D-H notation, analysis and simulating experiments were carried on to show the reachable space in three dimensions. In addition, multi ultrasonic sensors are used to detect the space realizing proximity controlled.

Keywords: Human Robot Cooperation, Spatially Aware, Reachable Space, Proximity Controlled, Implicit Pattern

1 Introduction

With the development of theories and technologies of automatic door is opened in advance. In addition, when we mail to somebody, contact with old friends, call a taxi, Smartphone or tablet PC with iOS, Android, or Windows Mobile system can help us to deal with many things too. Furthermore, in special environment, robot as an intelligent device is wildly used. It is good at repeating jobs or helping people doing something heavy, dangerous, and accurate.

Because of increasing number of devices, human computer interaction (HCI) is particularly important. Given that initiative and attention of devices, HCI can be divided into two types: one is explicit HCI, which is mainly concentrated at present, and the other is implicit one, which is gradually drawing more attention [1-3]. Studies have shown that implicit interaction can replace the explicit one effectively in the interactive process, though implicit interaction is not as accurate as the explicit one [4]. In fact, the recent researches suggest that if taking full advantage of user behaviour contexts as the implicit interactive information, we can do better in the interaction than using explicit information [5, 6].

Implicit human computer interaction (IHCI) is the frontier of the HCI [7]. Nicole Kaiyan, coming from Australia Swinburne University of technology, has put forward the concept of IHCI in 1996, but do not study in detail [8]. Since 2005, universities and institutes in the pervasive computing, more and more intelligent devices appear in our life. When we leave home, road condition signs always direct us. When we entry a building, an United States, Germany, China, Austria and so on, have concentrated on the IHCI theories, technologies and applications. Albrecht Schmidt, coming from Karlsruhe University, carried out related researches on IHCI earlier. Given that two elements of the implicit interactions are perception and reasoning, he proposed that context

information was extremely important to the interactive process, and modelled the interaction process with XML language at the same time [9, 10]. Andrew Wilson and Nuria Oliver, coming from Microsoft research in the United States, have developed four systems based on the machine vision to study the implicit interaction technology. Since 2007, Tao, coming from Tsinghua University in China, has developed an adaptive visual system to detect and understand user behaviours in order to apply on the implicit interactions [11]. Meanwhile, Tian, coming from Institute of Software in Chinese Academy of Science, studied the characteristics of the implicit interactions in the view of the WIMP user interface [12]. The development of IHCI at home and abroad was reviewed as shown in the Figure 1.

Table 1 compares IHCI researches from different views in details.

While the implicit interaction pattern is applied in the human robot cooperation, called IHRC, the robot can be smarter to work, active to cooperate, easy to understand partners, and safe to persons and itself. For example, when a robot collaborates with a partner, it does not only work by command, but also predict what the partner wants. So implicit interaction pattern is important to realize human robot cooperating. IHRC includes identity controlled, proximity controlled, profile controlled, and context controlled. Given the proximity controlled aspect, we discussed a method to enhance the cooperating safety in human robot cooperation.

The rest of the paper is organized as follows. In Section 2, the manipulator body we analysing is introduced briefly. Section 3 establishes the rear coordinate system of the manipulator firstly, and then analyses its reachable space. Section 4 discusses proximity controlled method based on multi ultrasonic sensors to enhance cooperating safety, which is an important aspect for human robot cooperation. Section 5 provides the conclusions and the future work.

*Corresponding author e-mail: wangwei8311@163.com

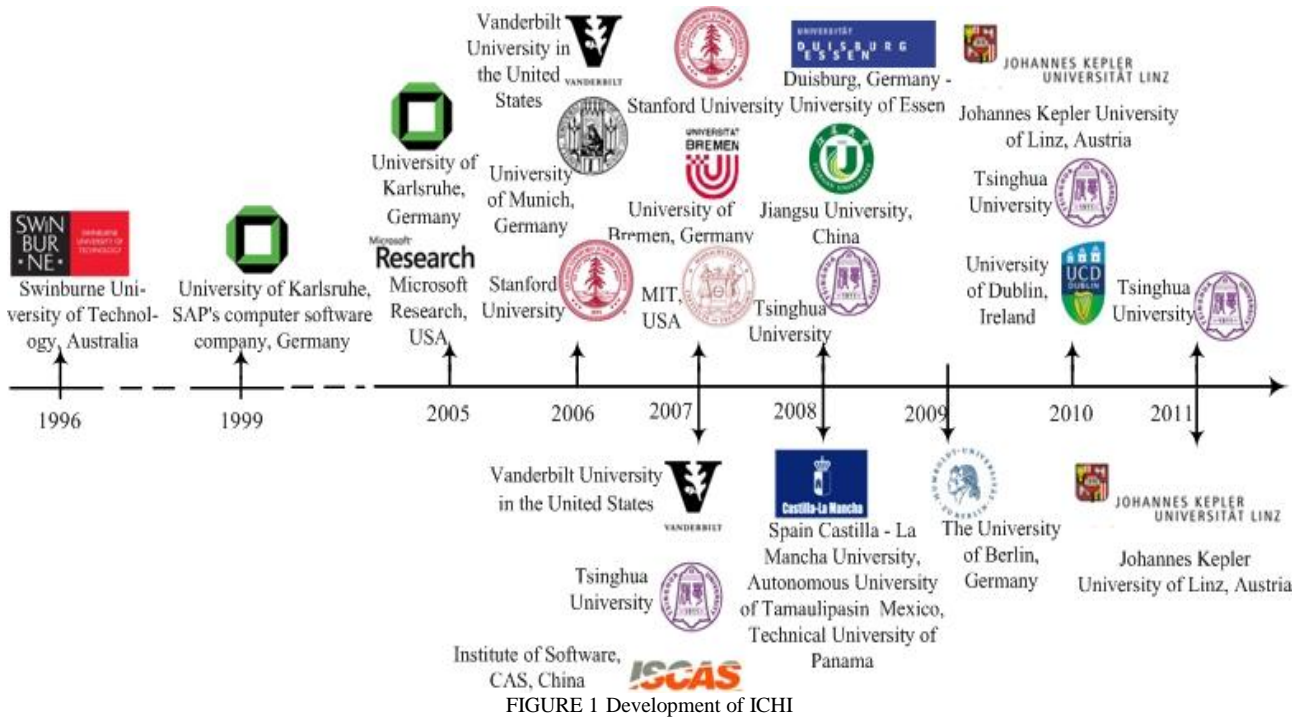


FIGURE 1 Development of ICHI

2 Platform of IHRC

The platform of implicit human robot cooperation is a double-loop system with a joint servo control loop and a

vision control loop based on a fixed monocular vision system. It consists of a camera, many types of sensors, a control board, PC, a wireless module and a serial robot arm with five degrees of freedom, as shown in Figure 2.

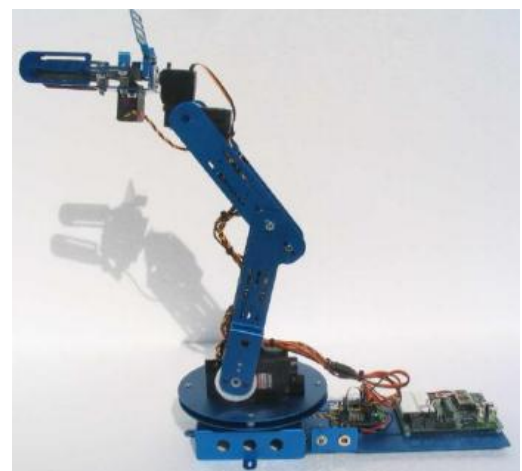
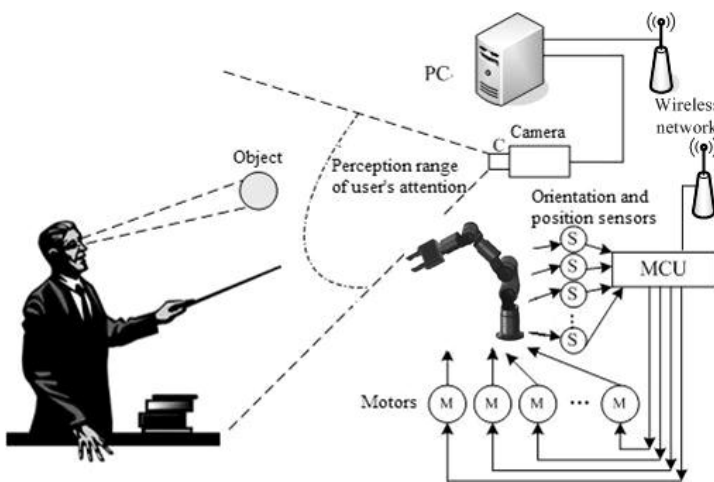


FIGURE 2 Hardware platform of IHRC

3 Reachable space of human robot cooperation

3.1 BODY COORDINATE SYSTEM OF ROBOT ARM

The rear coordinate system of the robot arm is establishes. There are six coordinate systems in total.

They are $O_0X_0Y_0Z_0$, $O_1X_1Y_1Z_1$, $O_2X_2Y_2Z_2$, $O_3X_3Y_3Z_3$, $O_4X_4Y_4Z_4$, $O_5X_5Y_5Z_5$, where $O_0X_0Y_0Z_0$ coincides with the world coordinate system as shown in Figure 3.

TABLE 1 Features of domestic and foreign IHCI

Type	Researchers	Contents	View	Context	Sensors	Continuity
Theory	M.J O'Grady, J Ye ^[13]	Context reasoning, embedded agents, middleware	-	-	-	-
	J Wendy, L Larry ^[14]	Mode, frame, type	-	-	-	-
	N Kaiyan ^[8]	Implicit theories	-	-	-	-
	A Schmidt ^[7]	Definition, identification, modelling, context-aware	-	-	-	-
	H Sebastien, S Ron ^[15]	Bayes Factor model for explicit and implicit interaction	-	-	-	-
	R Pramila ^[16]	Implicit human-computer interaction framework	User	Affection	Physiology	Continuous
	Li Feng, Pei Jun ^[17]	Adaptive user interest model	User	Action	Mouse	Discrete
	Tian Feng, Deng Changzhi ^[12]	Post-WIMP Interface Interactive tasks Generation	User	Action	Pen	Discrete
	Wang Guojian, Tao Linmi ^[11]	Shared services model, distributed vision systems	User	Action	Video	Continuous
Application	W Hendrik ^[18]	Get workers' behaviours	User	Action	Angle	Discrete
	J Wan, M J O'Grady ^[19]	Implicit health monitoring	User	Physiology	Physiology	Continuous
	H Drewes, A Schmidt ^[20]	Interaction based on sight	User	Sight	Video	Discrete
	W Andrew, O Nuria ^[10]	Implicit interaction based on gesture	User	Action	Video	Discrete
	J Wendy, A L Brian ^[21]	Whiteboard	User	Position	IR	Discrete
	A Riener ^[22]	Auxiliary Driving	User	Body gesture	Pressure	Discrete
	E A Arroyo ^[23]	The establishment of a task interference management model	User	Sight	Video	Discrete
	A Richard, W Monika ^[24]	Web usability enhancements based on implicit interaction	User	Action	Mouse	Discrete
	A Schmidt, H W Gellersen ^[25]	Input problem in wearable computing	Environment	Item Specifics	RFID	Discrete
	J Bravo, R Hervás ^[26]	Implicit interaction based on RFID and NFC	Environment	Position	RFID	Discrete
	R Pramila, S Nilanjan ^[16]	Implicit interaction between human and robot	User	Affection	Physiology	Continuous
	P Dai, L M Tao, et al ^[27]	Adaptive vision system for meeting	User	Action, complexion	Video	Discrete
	Ye Xiyong, Tao Linmi ^[28]	Implicit interaction based on the action understanding	User	Action	Video	Discrete

Based on the established D-H coordinate system of the robot arm, D-H parameters and the joint variables of each link are shown in Table 2. According to parameters above and D-H formulas:

$${}^0_1T = \begin{bmatrix} c\theta_1 & 0 & -s\theta_1 & 0 \\ s\theta_1 & 0 & c\theta_1 & 0 \\ 0 & -1 & 0 & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

$${}^1_2T = \begin{bmatrix} c\theta_2 & -s\theta_2 & 0 & a_2 \cdot c\theta_2 \\ s\theta_2 & c\theta_2 & 0 & a_2 \cdot s\theta_2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

$${}^2_3T = \begin{bmatrix} c\theta_3 & -s\theta_3 & 0 & a_3 \cdot c\theta_3 \\ s\theta_3 & c\theta_3 & 0 & a_3 \cdot s\theta_3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (3)$$

$${}^3_4T = \begin{bmatrix} c\theta_4 & 0 & -s\theta_4 & 0 \\ s\theta_4 & 0 & c\theta_4 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (4)$$

$${}^4_5T = \begin{bmatrix} c\theta_5 & -s\theta_5 & 0 & 0 \\ s\theta_5 & c\theta_5 & 0 & 0 \\ 0 & 0 & 1 & d_5 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (5)$$

where s denotes $\sin(\cdot)$, and c denotes $\cos(\cdot)$. Multiply each link matrix to obtain the posture and orientation of the end effector of the robot arm in the world coordinate system.

$${}^0_5T = {}^0_1T \cdot {}^1_2T \cdot {}^2_3T \cdot {}^3_4T \cdot {}^4_5T. \quad (6)$$

Thus, by solving inverse kinematics, joint variables of the robot arm can be calculated with the posture and orientation of the end effector.

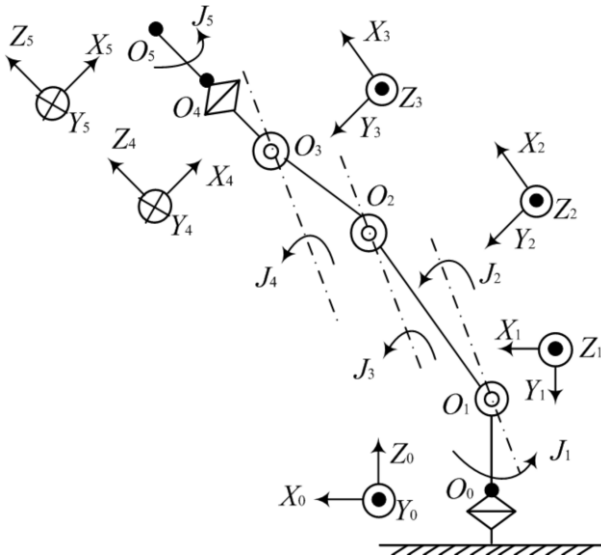


FIGURE 3 D-H coordinate of facial expression robot

TABLE 2 D-H parameter of Arm

Joint i	α_i /degree	a_i /mm	d_i /mm	θ_i /degree
1	-90	0	38	$\theta_1 \in [-90,90]$
2	0	154	0	$\theta_2 \in [-180,0]$
3	0	122	0	$\theta_3 \in [-30,150]$
4	-90	0	0	$\theta_4 \in [-180,0]$
5	0	0	181	$\theta_5 \in [-90,90]$

3.2 ANALYSIS TO THE REACHABLE SPACE

Reachable space of joint attention refers to a maximum range of activities of a reference point in the wrist mechanical interface coordinate system (the end effector coordinate system). Without considering the restrictions for the joint angle, for rotary joints, a fixed reference point P_n is assumed on the robot arm. It rotates around Z_n axis with the end coordinate system $O_n X_n Y_n Z_n$ together. And it forms a joint attention reachable sub-space $W_{n-1}(P_n)$ in the coordinate system $O_{n-1} X_{n-1} Y_{n-1} Z_{n-1}$, which is a circle. Then, the movement of linkage $n-1$ drives $W_{n-1}(P_n)$ rotating around Z_{n-1} axis to form another joint attention reachable sub-space $W_{n-2}(P_n)$ in the coordinate system $O_{n-2} X_{n-2} Y_{n-2} Z_{n-2}$, which is a toroid. After that, the movement of linkage $n-2$ drives $W_{n-2}(P_n)$ rotating around Z_{n-2} axis to form another joint attention reachable sub-space $W_{n-3}(P_n)$ in the coordinate system $O_{n-3} X_{n-3} Y_{n-3} Z_{n-3}$, which is a spinning body. Rotate around the former axis continually, the joint attention reachable sub-space of the fixed reference point P_n is also a spinning body. There is a transformation between neighbouring sub-spaces:

$$W_{n-j-1}(P_n) = Rot(z_{n-j}, \theta_{n-j}) \cdot W_{n-j}(P_n), \quad (7)$$

where, $j=0,1,\dots,n-2$.

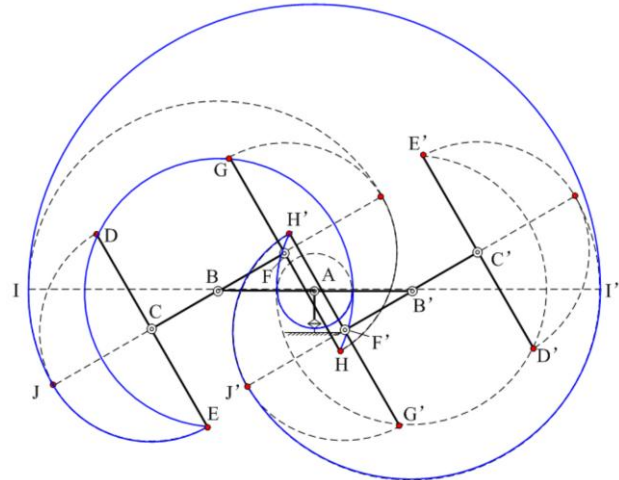


FIGURE 4 Reachable space of joint attention

According to the structure and parameters of the robot arm, $n=5$. As the statement above, we can get the reachable space of joint attention in human robot cooperating process shown in Figure 4. Due to the rotation axis Z_5 of the wrist joint J_5 is consistent with the linkage, the rotation does not have an impact on the reachable space. Therefore this joint can be ignored. For the rest of joints $J_1 \sim J_4$, firstly, fix $J_1 \sim J_3$ in the extreme position ABC. And rotate the linkage in the permissible range of θ_4 to form arc $DJE, \overset{\frown}{}$. Secondly, considering point B as the circle centre, rotate the linkage BC in the permissible range of θ_3 to form a plane encircled by arc $EJH, \overset{\frown}{}$ and arc $EDH, \overset{\frown}{}$, which is the swept area ρ of arc $DJE, \overset{\frown}{}$. Thirdly, considering point A as the circle centre, rotate the linkage AB in the permissible range of θ_2 to form a plane encircled by arc $EIH', \overset{\frown}{}$ and arc $EDH', \overset{\frown}{}$, which is the swept area σ of the plane ρ . Lastly, joint J_1 rotates in the permissible range of θ_1 . Reachable space of joint attention W_1 is obtained.

In the human robot cooperation, the reachable space of joint attention is determined by θ_i . Due to the limitations of the actual structure and driving device, the value of the joint variable is in a certain range, which is:

$$\min(\theta_i) \leq \theta_i \leq \max(\theta_i) \quad (i=1,2,\dots,5). \quad (8)$$

Suppose that the homogeneous coordinates of the end effector centre of the robot arm is $\vec{r}_5^{P_5} = [x_5^{P_5}(\vec{\theta}_i) \quad y_5^{P_5}(\vec{\theta}_i) \quad z_5^{P_5}(\vec{\theta}_i)]$ in the coordinate system $O_5 X_5 Y_5 Z_5$, where $\vec{\theta}_i = [\theta_1 \quad \theta_2 \quad \dots \quad \theta_5]$, and the homogeneous coordinates in the world coordinate system is $\vec{r}_0^{P_5}$, which satisfies:

$$\vec{r}_0^{P_5} = {}^0_5T \cdot \vec{r}_5^{P_5} \quad (9)$$

Therefore, the reachable space of joint attention is the set of $\vec{r}_0^{P_5}$ denoting as:

$$W_1(P_5) = \left\{ \vec{r}_0^{P_5} \mid \vec{r}_0^{P_5} = \begin{bmatrix} x_0^P(\vec{\theta}_1) & y_0^P(\vec{\theta}_1) & z_0^P(\vec{\theta}_1) \end{bmatrix} \right\} \quad (10)$$

3.3 REACHABLE SPACE SIMULATION

Based on Monte Carlo method, the reachable space of the robot arm can be simulated. Considering the number of random values for each joint angle is $N = 5000$, the reachable space is shown in Figure 5.

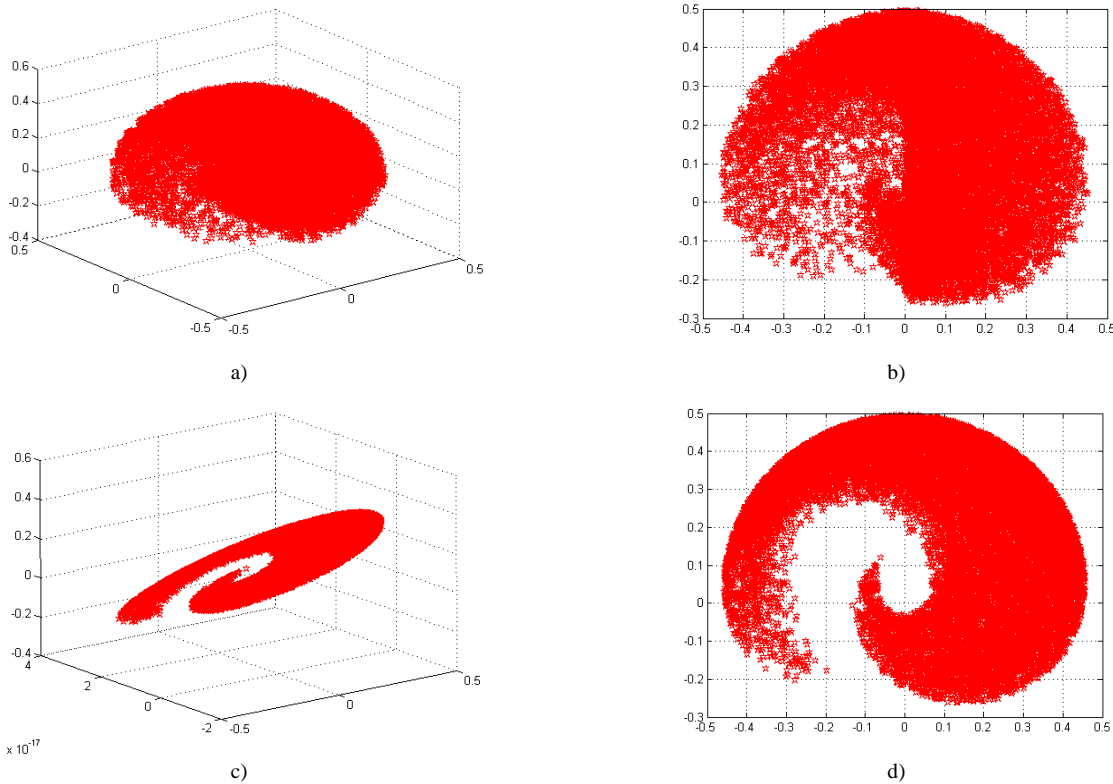


FIGURE 5 The shape of reachable space: a) The view of overall space in three-dimension; b) The view of overall space in XZ plane projection; c) The view of middle section in three-dimension; d) The view of middle section in XZ plane projection

The movement of each joint forms position cloud atlas of the end effector of the robot arm as shown in Figure 5(a) above. Its shape is similar to an ellipsoid. But due to its internal part is not a complete ellipsoid, the cloud atlas has different density. Its projection in the XZ plane is shown in Figure 5(b). In order to reveal the interior of the reachable space $W_1(P_5)$ of the robot arm, considering that the rotation of the joint 1 only makes position cloud atlas of the end effector shown in Figure 5(a) be symmetrical, draw a middle section $Y = 0$ as Figure 5(c). Figure 5(d) is the projection of the cloud

atlas of the middle section to XZ plane, which has the same shape and size as Figure 4.

According to the image coordinates of user's attention focus (u, v) , the corresponding world coordinates (X, Y, Z) can be calculated. Then solve inverse kinematic equations of the robot arm. In this paper, 8 attention focuses in different directions are considered to obtain joint angles. The results are shown in Table 3.

Based on the Robot Toolbox in Matlab, for 8 conditions stated above, kinematics simulation is carried on. The results are shown in Figure 6.

TABLE 3 Joint angles calculating with image coordinates

No.	Image coordinates (pixel)	World coordinates (mm)	Joint angle (degree)	Attention direction
1	(u,v)=(320,51)	(e,f,g)=(442.6,0,151.7)	[0,-14.4,0,0]	Above
2	(u,v)=(320,529)	(e,f,g)=(424.9,0,-130.2)	[0,21.6,0,0]	Below
3	(u,v)=(43,240)	(e,f,g)=(424.9,168.2,38)	[21.6,0,0,0]	Left
4	(u,v)=(596,240)	(e,f,g)=(424.9,-168.2,38)	[-21.6,0,0,0]	Right
5	(u,v)=(596,-131)	(e,f,g)=(384.5,-152.2,232.6)	[-21.6,-25.2,0,0]	Up right
6	(u,v)=(43,610)	(e,f,g)=(384.5,152.2,-156.6)	[-21.6,25.2,0,0]	Bottom right
7	(u,v)=(186,-111)	(e,f,g)=(406.2,77.5,232.6)	[10.8,-25.2,0,0]	Up left
8	(u,v)=(186,590)	(e,f,g)=(406.2,77.5,-156.6)	[10.8, 25.2,0,0]	Bottom left

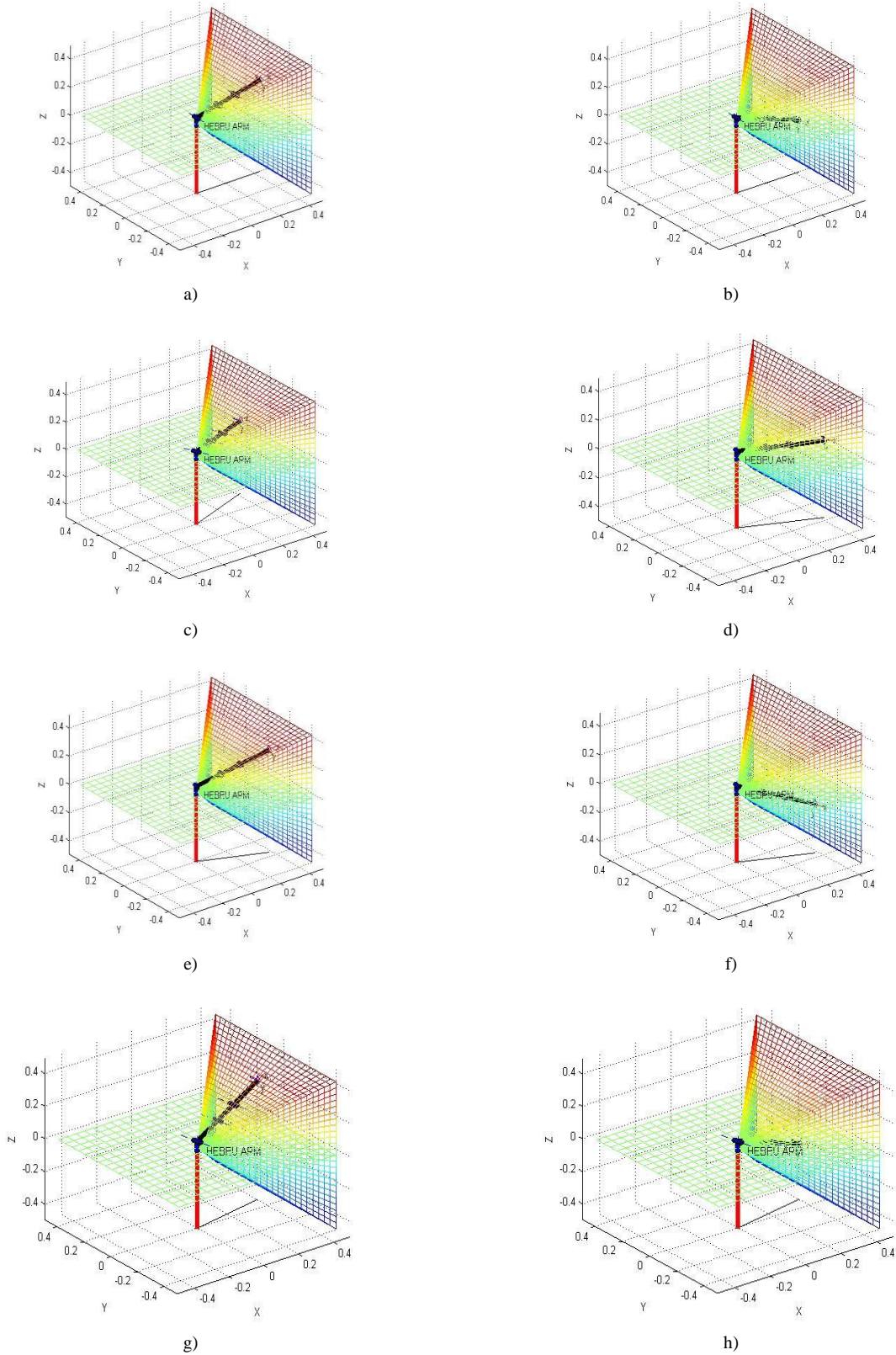


FIGURE 6 Simulation results: a) Above, b) Below, c) Left, d) Right, e) Up-right, f) Bottom-right, g) Up-left, h) Bottom-left

With the method in section 3, we can derive the world coordinates of user's attention from image coordinates. And then solve inverse kinematic equations to get every

joint angle. Figure 14 shows results, and manifests that the robot can track user's attention to realize joint attention.

4 Proximity controlled in IHRC

In order to make the human robot cooperating process be safe and effective, proximity controlled problem is considered. Based on multi ultrasonic sensors, we can detect an area, which is the same as the reachable space shown in figure 4. When the robot arm and its partner do one job together, they must share the same intent, attention, and task. So it is necessary to get the accurate positions of each other. On the one hand, proximity controlled keeps the person, who is in the reachable space of the robot arm, safe. On the other hand, the accurate positions information makes the cooperation effective.

In this paper, we use ultrasonic sensor to cover the reachable space. Directivity in sound pressure level of a single sensor is shown in figure 7. From this figure, we can see that its valid detecting angle is 60 degree.

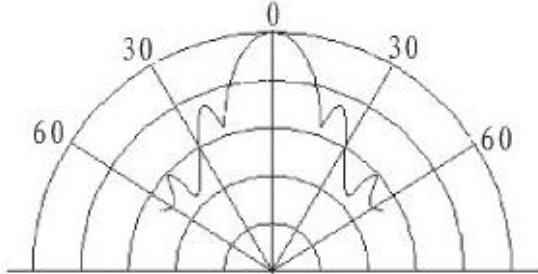


FIGURE 7 Directivity in sound pressure level

We arrange them together one by one as shown in Figure 8. Therefore, according to the tightness of valid detecting areas, degree of overlap (DoO) is defined, which means the maximum number of overlapped detecting areas. In addition, this index can be represented by the centre distance k of detecting areas. Arrangement of ultrasonic sensors falls into three special types: *Case 1*, *Case 2*, and *Case 3*.

Given that the valid detecting angle of ultrasonic sensor is 60 degree as shown in Figure 7, $\angle MOO_1 = 30^\circ$.

Case 1

In the triangle ΔMOO_1 , consider that

$OM = a_2 + a_3 + d_5$, so $OO_1 = OM \times \cos \angle MOO_1$,
and $MO_1 = OM \times \sin \angle MOO_1$. *Case 1*:

In this case, there is no overlap among detecting areas, so we consider DoO is 0. And centre distance is:

$$k = 2 \times MO_1 = 2 \times OM \times \sin \angle MOO_1. \tag{11}$$

However, if arranging sensors as Case 1, the middle area among them does not be sensed, this is a blind zone.

Case 2:

In this case, there is maximum degree of overlap in the middle area. DoO is 4. On this condition, the centre distance is:

$$k = MO_1 = OM \times \sin \angle MOO_1. \tag{12}$$

As we know, if arranging sensors as *Case 2*, the middle area is overlapped, which uses too many sensors. It is not economic.

Case 3:

In this case, the maximum degree of overlap is 2. And the centre distance is:

$$k = 2 \times PO_1. \tag{13}$$

According to geometric relations, $PO_1 = \frac{\sqrt{2}}{2} \times MO_1$. So,

$$k = 2 \times \frac{\sqrt{2}}{2} \times MO_1 = \sqrt{2} \times MO_1.$$

Arrangement as *Case 3* is a good method. It does not only make full use of the valid detecting angle of every ultrasonic sensor, but also have not any blind zone.

Considering the statement above, on the one hand, while $k \in [MO_1, \sqrt{2}MO_1]$, the arrangement of ultrasonic sensors seems too tight, this uses too many sensors and disturbs with each other strongly. On the other hand, while $k \in [\sqrt{2}MO_1, 2MO_1]$, the arrangement of ultrasonic sensors get looser, which make the middle area of sensors be a blind zone. Therefore, $k = \sqrt{2} \times MO_1$, the situation shown in *Case 3* is suitable.

For *Case 3*, the number of ultrasonic sensors could be calculated as follows. Because inner and outer circles are in the same plane, we can obtain that:

$$\begin{aligned} \angle POO_1 &= \arctan\left(\frac{PO_1}{OO_1}\right) = \\ &= \arctan\left(\frac{\frac{\sqrt{2}}{2} \times OM \times \sin \angle MOO_1}{OM \times \cos \angle MOO_1}\right) \\ &= \arctan\left(\frac{\sqrt{2}}{2} \times \tan \angle MOO_1\right) = 22.2^\circ. \end{aligned} \tag{14}$$

Therefore, the number of ultrasonic sensors we needed N is:

$$\begin{aligned} N &= \left\lceil \frac{1}{2} \times \left(\frac{360^\circ}{2\angle POQ} \times \frac{360^\circ}{\angle POQ} - \frac{360^\circ}{\angle POQ} \right) \right\rceil \\ &= \left\lceil \left(\frac{180^\circ}{\angle POQ} \right)^2 - \frac{180^\circ}{\angle POQ} \right\rceil = 12. \end{aligned} \tag{15}$$

Arrange sensors as shown in figure 8(c), and make them work one by one to reduce interference between them. While cooperating, sensors detect and ensure us

safe. With the accurate positions, cooperation process will be effective.

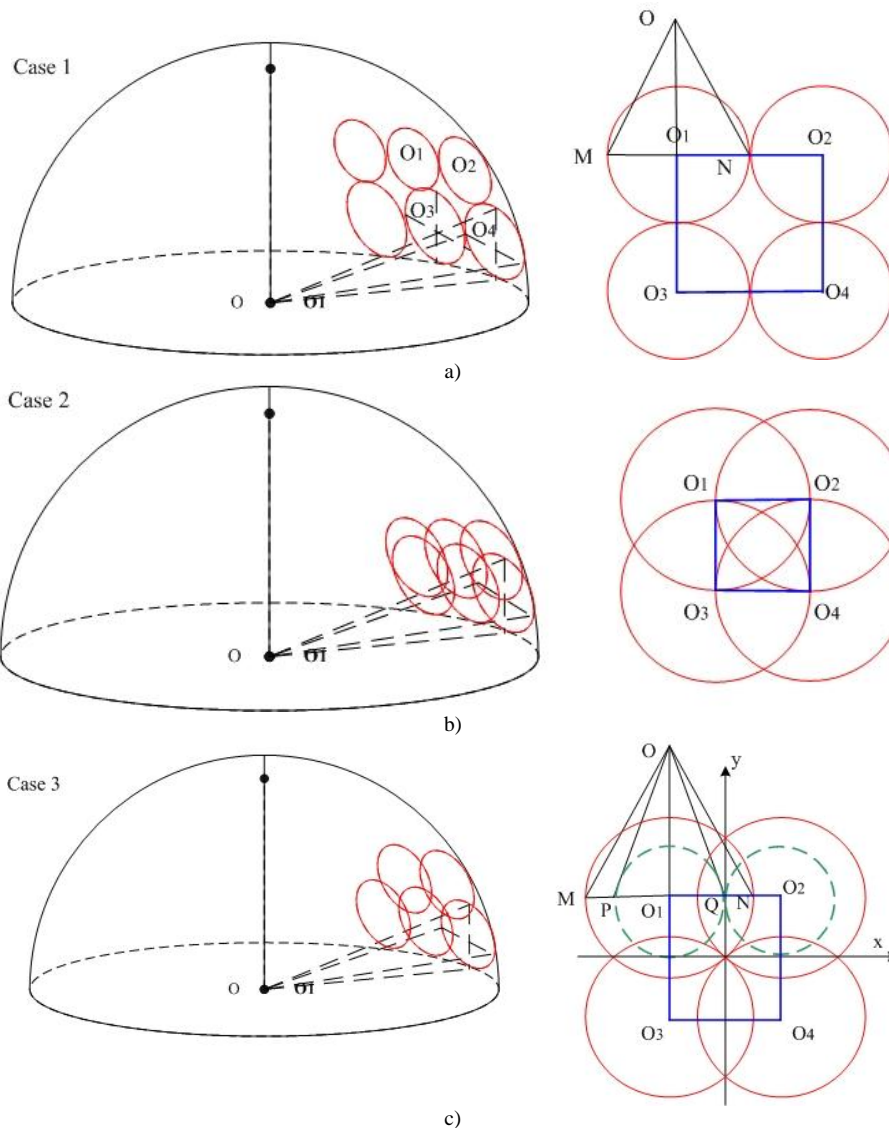


FIGURE 8 Arrangement of ultrasonic sensors: a) Case 1, b) Case 2, c) Case 3

5 Conclusions

- 1) Proximity controlled in implicit human robot interaction is important. It is not only a key part to realize the effective human robot interaction, but also the foundation of safety.
- 2) According to the establishment and transform of the coordinate systems, calculation of three-dimensional coordinate, connection of the robot's joints and linkages, and other theories and technologies, reachable space of a robot arm with five degrees of freedom was researched to prepare for human-machine cooperating subsequently.

- 3) Based on the Robot Toolbox in Matlab, simulation environment for reachable space of the robot arm was built. And analyse the reachable space by simulation.
- 4) With multi ultrasonic sensors, we establish their applying method, and calculate the number of ultrasonic sensors.


Acknowledgments

This work was supported by the Foundation for Young Scholars of Hebei Educational Committee (QN20131152, ZD20131084).

References

- [1] Maurice C 2009 *Invisible Computing* (9) 86-8
- [2] Eric H 1999 Attention-sensitive alerting *Proceedings of UAI '99 Conference on Uncertainty and Artificial Intelligence* San Francisco 305-13
- [3] James F 2005 Examining task engagement in sensor-based statistical models of human interruptibility *Proceedings of the ACM Conference on Human Factors in Computing Systems* 331-40
- [4] Kelly D 2003 Implicit feedback for inferring user preference: A bibliography *SIGIR Forum* 37(2) 18-28
- [5] Teevan J 2005 Personalizing search via automated analysis of interests and activities *Proceedings of the SIGIR* 449-56
- [6] Lv Y H 2007 *A study of personalized information retrieval based on implicit feedback* Beijing: Institute of Software Chinese Academy of Sciences Press China (in Chinese)
- [7] Albrecht S 2010 *IEEE Pervasive Computing* 9(1) 85-8
- [8] Nicole K 1996 Exploratory study of implicit theories in human computer interaction *Proceedings of the 6th Australian Conference on Computer-Human Interaction* 338-9
- [9] Albrecht S 2000 Implicit human computer interaction through context *Personal Technologies* 4(2&3) 191-9
- [10] Andrew W 2005 Multimodal sensing for explicit and implicit interaction *Proceedings of the 11th International Conference on Human-Computer Interaction* 1-10
- [11] Wang G J 2010 *Journal of Image and Graphics* 15(8) 1133-8 (in Chinese)
- [12] Tian F 2007 *Journal of Frontiers of Computer Science and Technology* 1(2) 160-9
- [13] Michael J O, Ye J 2009 A middleware for implicit interaction *Proceedings of Instinctive Computing Workshop Carnegie Mellon* 143-61
- [14] Wendy J, Larry L 2008 *Design Issues: Special Issue on Design Research in Interaction Design* 24(3) 72-84
- [15] Sebastien H, Sun R 2009 Simulating incubation effects using the explicit implicit interaction with bayes factor (EII-BF) model *Proceedings of IJCNN* 1199-205
- [16] Pramila R, Nilanjan S 2007 *Advanced Engineering Informatics* 21(3) 323-34
- [17] Li F, Pei J 2008 *Computer Engineering and Applications* 44(9) 76-79 (in Chinese)
- [18] Hendrik W, Holger K 2006 Towards implicit interaction by using wearable interaction device sensors for more than one task *Proceedings of the 3rd international conference on mobile technology, applications & systems (Mobility '06)* 270-5
- [19] Wan J, Michael J O 2009 Implicit interaction: a modality for ambient exercise monitoring *Proceedings of INTERACT* 900-3
- [20] Drewes H, Schmidt A. 2007 Interacting with the Computer using gaze gestures *Proceedings of INTERACT* 475-488
- [21] Wendy J, Brian A L 2008 Range: exploring implicit interaction through electronic whiteboard design *Proceedings of CSCW* 17-26
- [22] Andreas R 2010 *Sensor-Actuator Supported Implicit Interaction in Driver Assistance Systems*, Springer Press: Heidelberg
- [23] Ernesto A A 2007 *Mediating Disruption in Human-Computer Interaction from Implicit Metrics of Attention*, Massachusetts Institute of Technology Press: Massachusetts
- [24] Richard A, Monika W 2006 Knowing the user's every move - user activity tracking for website usability evaluation and implicit interaction *Proceedings of WWW* 203-12
- [25] Albrecht S, Hans W G 2000 Enabling implicit human computer interaction a wearable RFID-Tag reader *Proceedings of ISWC*, 193-194
- [26] Bravo J, Hervás R 2008 From implicit to touching interaction: RFID and NFC approaches *Proceedings of HSI7* 43-48
- [27] Dai P, Tao L M 2007 An adaptive vision system toward implicit human computer interaction *Proceedings of 4th International Conference on Universal Access in Human-Computer Interaction*, 792-801
- [28] Ye X Y, Tao L M 2011 Implicit interaction based on action understanding *Proceedings of 7th HHME. China: Multimedia Technical Committee in Computer Society* 91-8

Authors

	<p>Wei Wang, born on November, 1983, Handan, Hebei Province, P.R. China</p> <p>Current position, grades: the lecturer of School of Information & Electrical Engineering, Hebei University of Engineering, China. University studies: Ms.Sc. in Control Theory and Control Engineering, Jiangnan University in China; Dr.Sc., University of Science and Technology Beijing in China Scientific interest: Human-robot Cooperation, Implicit Interaction. Publications: more than 30 papers Experience: teaching experience of 3 years, three scientific research projects</p>
	<p>Xiao-dan Huang, born on November, 1983, Handan, Hebei Province, P.R. China</p> <p>Current position, grades: the lecturer of School of Information & Electrical Engineering, Hebei University of Engineering, China University studies: Ms.Sc. in Electronic Information from University of Science and Technology Beijing in China Scientific interest: Implicit Interaction, Robot. Publications: more than 10 papers Experience: teaching experience of 3 years, two scientific research projects</p>

Study on T-S fuzzy sliding mode control based on a new reaching law

Hongqing Hou, Qian Miao, Chuanqiang Yu, Qinhe Gao*

Xi'an Research Inst. of hi-Tech hongqing Town, Xi'an City, Shan'xi Province, China, 710025

Received 1 March 2014, www.tsi.lv

Abstract

Fuzzy Sliding Mode controller based on fuzzy T-S Model is designed for the nonlinear, uncertainties and fast variable time characterizes question of bank-to-turn aircraft control and guidance system model. BTT control model was obtained by using T-S modelling method, asymptotically stable sliding surface was designed, and a new sliding mode reaching law is proposed. Based on the new reaching law, sliding mode stable tracking controller is designed. At last, the rationality and the effectiveness of the designed T-S fuzzy sliding mode stable tracking controller with the new reaching law are verified by the theoretical proof and the simulation experiments.

Keywords: Bank-to-turn Aircraft, Fuzzy Control, Sliding Mode Control, Reaching Law

1 Introduction

BTT(bank-to-turn is called BTT for short.) aircraft control technology is a kind of advanced control technology, It is great theoretical significance and the value of engineering applications for the development of high-speed, high precision and high mobile aircraft system[1]. However, BTT aircraft is a multivariate controlled object of kinematic coupling, inertia coupling, pneumatic coupling and control coupling [2].

Fuzzy control has not depended on accurate mathematical model of the characteristics and robustness and adaptability, it is an effective method to solve complex control problems [3-4]. T-S fuzzy model is an important tool for the modelling of nonlinear uncertain systems; Literature [5] has proved the T-S fuzzy model is a better approximation performance than the Mamdani fuzzy model. Sliding mode control method has strong robustness for parameter perturbation and the external interference, Therefore it is subject to the attention of scholars from various countries [6]. this program of Fuzzy control and sliding mode control combining is maintained the advantages of sliding mode control and overcome the sliding mode control shaking, has strong robustness for model uncertainties and external disturbances, improved reaching the dynamic quality of the system segment. The Lin and Boulkroune [7-9, 13] design controller the application of fuzzy system, many scholars have adopted the fuzzy basis function system. Combining the characteristics of BTT control, reaching law T-S model fuzzy sliding mode control will be applied to the control of the BTT aircraft in order to meet the needs of the BTT Control. Simulation results show

that the system has good dynamic qualities and strong robustness.

2 T-S Fuzzy modelling of nonlinear systems

2.1 PROBLEM DESCRIPTION

Considering a class of second-order multi-input and multi-output nonlinear systems:

$$\ddot{\mathbf{y}} = \mathbf{F}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}. \quad (1)$$

In the formula, $\mathbf{y} = [x_1 \ x_2 \ \dots \ x_n]^T \in \mathbb{R}^n$ is the output vector of system; $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$ is the state vector of system, and State can be measured; $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^n$ and $\mathbf{G}(\mathbf{x}) \in \mathbb{R}^n$ are nonlinear function of system state matrix, and $\text{rank}(\mathbf{G}) = n$; $\mathbf{u} \in \mathbb{R}^n$ is the input vector of system.

System (1) is a nonlinear system, this nonlinear model is difficult to design a global control law, because an essentially nonlinear system cannot carry out its global linear model through a global linearization, but it can be represented as a series of local linear system. Fuzzy sliding mode control method combines the advantages of both closely, while maintaining the fuzzy control system does not rely on the characteristics of the model, but it also has the advantage of easy sliding mode control system design and stability analysis theory.

2.2 FUZZY T-S MODEL

T-S fuzzy model is a nonlinear model and easy to express dynamic characteristics of complex systems.

*Corresponding author e-mail:20016109@163.com

Consider nonlinear uncertain systems described by the fuzzy T-S model. The system is described as the following m fuzzy rules, the i th fuzzy inference rule: if z_1 is F_1^i and z_2 is $F_2^i \cdots z_n$ is F_n^i , then

$$\begin{aligned} \dot{x}(t) &= A_i x(t) + B_i u(t) \\ y(t) &= x(t) \end{aligned} \tag{2}$$

In the formula: $z(t) = [z_1(t), z_2(t), \dots, z_n(t)]^T$ is Fuzzy antecedent variables. F_j^i is Fuzzy sets. $x(t) \in R^n$ is State variables. $u(t) \in R^m$ is Input of fuzzy system. $i = 1, 2, \dots, n$ is the number of system input. $y(t) \in R^n$ is Fuzzy output of system. $A_i \in R^{n \times n}$ and $B_i \in R^{n \times m}$ are system matrix and output matrix.

The state equation of fuzzy system from a single point of fuzzification, product inference and centre weighted average and defuzzification is [10-11].

$$\begin{aligned} \dot{x}(t) &= \sum_{i=1}^n \bar{\alpha}_i(z(t)) [A_i x(t) + B_i u(t)] \\ y(t) &= \sum_{i=1}^n \bar{\alpha}_i(z(t)) x(t) \end{aligned} \tag{3}$$

where $\bar{\alpha}_i(z(t)) = \prod_{j=1}^n \mu_{F_j^i}(z_j(t)) / \sum_{i=1}^m \prod_{j=1}^n \mu_{F_j^i}(z_j(t))$, $\mu_{F_j^i}$ is the membership function of $z(t)$ related to the fuzzy sets.

Assume that the fuzzy system is controllable, the fuzzy control rules i is R^i : if z_1 is F_1^i and z_2 is $F_2^i \cdots z_n$ is F_n^i , then $u(t) = -K_i x(t)$

The entire state feedback control law is:

$$u(t) = \sum_{i=1}^n \bar{\alpha}_i(z(t)) K_i x(t) \tag{4}$$

The essentially of T-S fuzzy modelling method is that nonlinear dynamical system model can be regarded as fuzzy multiple local linear approximation model, so in order to strike a T-S model for nonlinear systems, you must obtain each local subsystem linear model.

2.3 CONSTRUCTED THE T-S FUZZY MODEL

For a class of nonlinear dynamic process model:

$$\dot{x} = F(x) + G(x)u \tag{5}$$

In the formula, $x = [x_1, x_2, \dots, x_n]^T$ is the output vector of the system; $F(x) \in R^n$ and $G(x) \in R^n$ are nonlinear

smooth vector function, $u \in R^n$ is the input vector of the system.

If $T(x, u) = F(x) + G(x)u$, the formula (5) can be expressed as:

$$\dot{x} = T(x, u) \tag{6}$$

If $T(x_0, u_0) = F(x_0) + G(x_0)u_0$, and at the operating point (x_0, u_0) , the $T(x, u)$ launched by Taylor methods available:

$$\begin{aligned} \dot{x} &= T(x_0, u_0) + \left. \frac{\partial T}{\partial x} \right|_{\substack{x=x_0 \\ u=u_0}} (x - x_0) \\ &+ \left. \frac{\partial T}{\partial u} \right|_{\substack{x=x_0 \\ u=u_0}} (u - u_0) + \dots \end{aligned} \tag{7}$$

Assuming g_{ij} is the (i, j) -th element of the matrix $G(x)$, then:

$$\left. \frac{\partial T}{\partial x} \right|_{\substack{x=x_0 \\ u=u_0}} (x - x_0) = \left. \frac{\partial T}{\partial x} \right|_{x=x_0} + H(x_0, u_0) \tag{8}$$

In the formula, $\sum_{k=1}^m u_k \left. \frac{\partial g_{ik}(x)}{\partial x_j} \right|_{\substack{x=x_0 \\ u=u_0}}$ is the (i, j) -th element of the matrix $H(x_0, u_0)$, the formula (8) finishing as:

$$\left. \frac{\partial T}{\partial x} \right|_{\substack{x=x_0 \\ u=u_0}} (x - x_0) = G(x_0) \tag{9}$$

If $T(x_0, u_0) = 0$, $(x_0^T, u_0^T)^T$ is the equilibrium point of the formula (6), scilicet the point at (x_0, u_0) has $\dot{x} = 0$.

Assuming $\delta x = x - x_0$, and ignoring the Taylor expansion of the higher order terms, linear model can be obtained at the equilibrium point (x_0, u_0) :

$$\frac{d}{dt} x = Ax + B\delta u \tag{10}$$

In the formula, $A = \left. \frac{\partial T}{\partial x} \right|_{\substack{x=x_0 \\ u=u_0}}$; $B = \left. \frac{\partial T}{\partial u} \right|_{\substack{x=x_0 \\ u=u_0}}$.

In the equilibrium point $x = 0$, it can be constructed locally linear model. In the non-equilibrium point at work, Taylor linearization of the nonlinear model is a result of the radiation model, rather than a linear model.

Assuming for a given operating point x_0 , but it is not necessarily the equilibrium point of the formula (6), construct matrices A and B , so that in the neighbourhood x_0 there are:

$$\begin{aligned} F(x) + G(x)u &\approx Ax + Bu, \quad \forall u \\ F(x_0) + G(x_0)u &= Ax_0 + Bu, \quad \forall u \end{aligned} \tag{11}$$

Because u is arbitrary, so there must be $G(x_0) = B$. Then, in the x_0 's neighbourhood to find a matrix A such that there is:

$$\begin{aligned} F(x) &\approx Ax \\ F(x_0) &= Ax_0 \end{aligned} \tag{12}$$

a_i^T is assumed that the i -th row matrix, F_i is the i -th component of F , the formula (12) may be expressed as:

$$\begin{aligned} F_i(x) &\approx a_i^T x, \quad i=1,2,\dots,n \\ F_i(x_0) &= a_i^T x_0, \quad i=1,2,\dots,n \end{aligned} \tag{13}$$

In points x_0 , expand the left side of the above equation and ignore the second and higher order terms, finishing can be:

$$\nabla^T F_i(x_0)(x - x_0) \approx a_i^T (x - x_0) \tag{14}$$

For vector a_i , it should be kept as close to $\nabla F_i(x_0)$, and to satisfy the constraints $F_i(x_0) = a_i^T x_0$.

Defines $E = \frac{1}{2} \|\nabla F_i(x_0) - a_i\|_2^2$, then a_i can be attributed to the following optimization problem:

$$\min E_{a_i}, \quad a_i^T x_0 = F_i(x_0) \tag{15}$$

Formula (15) is a convex constrained optimization problem. Then, the first condition of the optimization problem is:

$$\begin{aligned} \nabla_{a_i} E + \lambda \nabla_{a_i} (a_i^T x_0 - F_i(x_0)) &= 0 \\ a_i^T x_0 &= F_i(x_0) \end{aligned} \tag{16}$$

In the formula, λ is Lagrange multiplier; The a_i of ∇_{a_i} subscript gradient computation of ∇ is associated with a_i . The formula (16) derived:

$$\begin{aligned} a_i - \nabla F_i(x_0) + \lambda x_0 &= 0 \\ a_i^T x_0 &= F_i(x_0) \end{aligned} \tag{17}$$

Finishing the above formula can be obtained:

$$a_i = \nabla F_i(x_0) + \frac{F_i(x_0) - x_0^T \nabla F_i(x_0)}{\|x_0\|^2} x_0$$

The key to the stability problem of nonlinear systems is to construct the controller, make it in a neighbourhood point within an arbitrary starting point; the trajectory converges to the point of the closed loop system. If the starting point is just the working point, then the trajectories of the closed loop system can stop at this point; expect the controller at after any time. Therefore, the work point must be asymptotically stable equilibrium point of the closed-loop system.

3 Reaching Law Designed

The dynamic quality of sliding movement depends on the switching function and its parameters. In many cases, the reaching law design is features with a simple structure, high quality, it cannot output u limit structure, and only rely on the calculation of its structure and form. When the other variable structure cannot be applied, the reaching law method can still give the results [6, 10].

Reaching law is the following Common.

Equal to the speed of approaching law

$$\dot{s} = -\varepsilon \operatorname{sgn}(s), \tag{18}$$

$\varepsilon > 0$ is the rate of system motion point approach the switching surface. ε is the smaller, slower approach, the adjustment process is slow, ε is the greater, the closer to the fast adjustment process is fast jitter.

Exponential reaching law

$$\dot{s} = -\varepsilon \operatorname{sgn}(s) - ks \quad \varepsilon > 0, k > 0 \tag{19}$$

The reaching law approach speed from the large value gradually reduced to zero, to shorten the approach time, the motion point to the speed of the switching surface is very small. Good by adjusting the parameters can improve the dynamic quality, and reduce the chattering.

Power reaching law:

$$\dot{s} = -k |s|^a \operatorname{sgn}(s) \quad 0 < a < 1, k > 0 \tag{20}$$

The appropriate choice the value of a and k obtained with the exponential reaching law similar results.

Generally reaching law:

$$\dot{s} = -\varepsilon \operatorname{sgn}(s) - f(s) \quad \varepsilon > 0, \tag{21}$$

where $f(0) = 0$, when $s \neq 0$, $sf(s) > 0$.

Exponential reaching law to reach a relatively short time, while the reaching law will be relatively large speed when the motor points to reach the switching surface, causing relatively large buffeting. Power reaching law and the exponential reaching law closer. The general results of reaching law and selection of $f(s)$ is closely related, the selection $f(s)$ of appropriate can effectively eliminate chattering.

In the sliding mode control system, the control structure exists switching characteristics, so that the sliding mode was buffeting form. Thus, scholars proposed saturation function instead of using the sign function. The difference is shown in Figure 1.

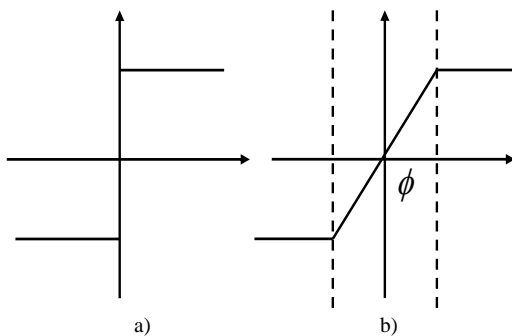


FIGURE 1 The difference between symbolic function and saturation function: a) $\text{sgn}(s)$, b) $\text{sat}(s)$

Saturation function can be expressed as:

$$\text{sat}(s) = \begin{cases} \text{sgn}(s), & s \geq \phi \\ \frac{s}{\phi}, & s < \phi \end{cases}$$

To enable smooth sliding mode presents the movement, to eliminate chattering of sliding movement, Saturation function can be deformed to obtain saturation function modified form, shown in Figure 2.

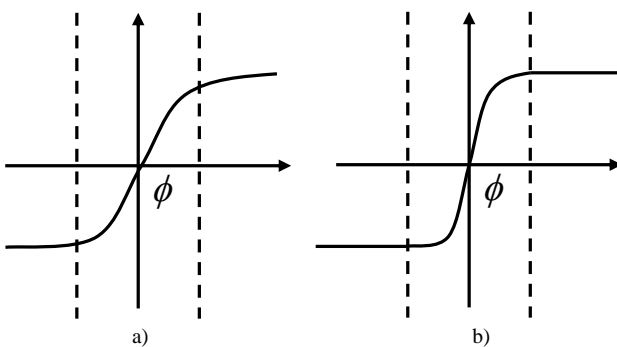


FIGURE 2 Deformation saturation function: a) $\text{sat}_1(s)$, b) $\text{sat}_2(s)$

Deformed saturation function $\text{sat}_1(s)$ is expressed as:

$$\text{sat}_1(s) = \begin{cases} \text{sgn}(s), & s \geq \phi \\ \tanh\left(\frac{2\pi s}{\phi}\right), & s < \phi \end{cases} \quad (22)$$

Deformed saturation function $\text{sat}_2(s)$ is expressed as:

$$\text{sat}_2(s) = \begin{cases} \text{sgn}(s), & s \geq \phi \\ \left(\frac{s}{\phi}\right)^p, & s < \phi \end{cases} \quad (23)$$

For saturation function $\text{sat}_2(s)$:

- 1) When $s < \phi$, that is $s(t)$ when the value is very small, $s(t)$ instead of using s/ϕ , and extended the form of powers reaching law :
 $\dot{s} = -k_2 |s|^\beta \text{sgn}(s)$ $k_2 > 0, 0 < \beta < 1$
- 2) When $s \geq \phi$, that is $s(t)$ when the value is very big, we want a faster rate close to equilibrium, the index reaching law can be at a faster rate approaching equilibrium, so select the form of index reaching law:
 $\dot{s} = -k_1 s |s|^\alpha \text{sgn}(s)$ $k_1 > 0, 0 < \alpha < 1$

Through the above analysis, design new fast reaching law is

$$\dot{s} = -k_1 s |s|^\alpha \text{sgn}(s) - k_2 |s|^\beta \text{sgn}(s), \quad (24)$$

where $k_1 > 0, k_2 > 0, 0 < \alpha < 1, 0 < \beta < 1$.

When the system state away from the sliding surface, the value of a great $s(t)$, the approach speed depends primarily on the first item of the equation (24) in the right; when the system state is close to the sliding surface, that is when the $s(t)$ value is very small, approaching the speed depends on the type (24) in the right.

4 T-S fuzzy sliding mode control

Dynamic quality of sliding movement depends on switching function and its parameters. In many cases, reaching law cannot restrict the output design structure, relying only calculated its structure and form, has a simple structure, high quality features. When other variable structure cannot be applied, reaching law still gives the result.

Consider the nonlinear controlled object, and it can be described by fuzzy TS model, fuzzy TS model to describe the j -th fuzzy inference rules of system are:

$$\begin{aligned} R^j : & \text{IF } x_1 \text{ is } F_1^j \text{ and } x_2 \text{ is } F_2^j \text{ and } \dots \text{ and } x_n \text{ is } F_n^j, \\ & \text{THEN } \dot{x}(t) = A_j x(t) + B_j u(t). \quad (25) \\ & y(t) = C_j x(t). \end{aligned}$$

In the formula, F_j^i is the fuzzy set; x_j the j -th variable system; $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbf{R}^n$ is the input of fuzzy system, $i = 1, 2, 3, \dots, r$ is the number of rules fuzzy inference system; $\mathbf{y} \in \mathbf{R}^m$ is the output of fuzzy system; \mathbf{u} is the control input vector; $\mathbf{A}_i \in \mathbf{R}^{n \times n}$ and $\mathbf{B}_i \in \mathbf{R}^{n \times m}$ are the system matrix and the input matrix; \mathbf{C}_i is the output matrix.

Assuming $\alpha_i(z(t))$ is the membership function, which $x(t)$ relevant in fuzzy set F_j^i , through a single point of blurring, product inference and centre weighted average anti-blur constitute the global fuzzy model can be expressed as:

$$\begin{cases} \dot{\mathbf{x}}(t) = \frac{\sum_{i=1}^r \alpha_i(\mathbf{x}(t))[\mathbf{A}_i \mathbf{x}(t) + \mathbf{B}_i \mathbf{u}(t)]}{\sum_{i=1}^r \alpha_i(\mathbf{x}(t))} \\ \mathbf{y}(t) = \sum_{i=1}^r \bar{\alpha}_i(\mathbf{x}(t)) \mathbf{C}_i \mathbf{x}(t) \end{cases}, \quad (26)$$

Among: $\alpha_i(z(t)) = \prod_{i=1}^r \mu_{F_i^i}(z(t))$;

$$\bar{\alpha}_i(z(t)) = \alpha_i(z(t)) / \sum_{i=1}^r \alpha_i(z(t)), \quad i = 1, 2, \dots, r.$$

4.1 FUZZY SLIDING MODE CONTROL BASED ON REACHING LAW

According to the characteristics of reaching law method, reaching law sliding mode controller is designed based on T-S fuzzy model.

Control law is derived as follows:

$$\mathbf{s} = \mathbf{C}\mathbf{x}. \quad (27)$$

Among: $\mathbf{C} = \text{diag}(c_1, c_2, \dots, c_{13}), c_i (i = 1, 2, \dots, n)$ is constant.

Assuming slaw is reaching law, by derivation and organizes the formula can get:

$$\dot{\mathbf{s}} = \mathbf{C}\dot{\mathbf{x}} = \text{slaw}. \quad (28)$$

Through state equation (25) into equation (28), we obtain T-S fuzzy sliding mode controller is:

$$\begin{aligned} \mathbf{R}^i : \text{IF } x_1 \text{ is } F_1^i \text{ and } x_2 \text{ is } F_2^i \text{ and } \dots \text{ and } x_n \text{ is } F_n^i, \\ \mathbf{u}_i = (\mathbf{C}\mathbf{B}_i)^{-1}(-\mathbf{C}\mathbf{A}_i \mathbf{x}(t) + \text{slaw}). \end{aligned}$$

Through new fast reaching law (24) into the above formula, T-S fuzzy sliding mode controller based on the new reaching law is:

$$\begin{aligned} \mathbf{R}^i : \text{IF } x_1 \text{ is } F_1^i \text{ and } x_2 \text{ is } F_2^i \text{ and } \dots \text{ and } x_n \text{ is } F_n^i, \\ \mathbf{u}_i = (\mathbf{C}\mathbf{B}_i)^{-1}(-\mathbf{C}\mathbf{A}_i \mathbf{x}(t) - K_1 \mathbf{S} |\mathbf{S}|^\alpha \text{sgn}(\mathbf{S}) - K_2 |\mathbf{S}|^\beta \text{sgn}(\mathbf{S})), \quad (29) \end{aligned}$$

Among: $0 < \alpha < 1, 0 < \beta < 1, K_1 = \text{diag}[k_{11} \ k_{12} \ k_{13}], K_2 = \text{diag}[k_{21} \ k_{22} \ k_{23}], k_{ji} > 0, (i = 1, 2, 3; j = 1, 2).$

Stability proof:

Take Lyapunov function is $\mathbf{V} = \frac{1}{2} \mathbf{s}^T \mathbf{s}$. Derivation of this formula and equation (24) into it, it can be obtained: $\dot{\mathbf{V}} = \mathbf{s}^T \dot{\mathbf{s}} = -K_1 \mathbf{s}^T \mathbf{s} |\mathbf{s}|^\alpha \text{sgn}(\mathbf{s}) - K_2 \mathbf{s}^T \mathbf{s} |\mathbf{s}|^\beta \text{sgn}(\mathbf{s}) < 0$.

Each fuzzy sliding mode subsystem is asymptotically stable; each local subsystem global controller is a weighted sum of the control law, based on the above model analysis can be obtained: $\mathbf{u} = \sum_{i=1}^r \alpha_i \mathbf{u}_i$.

Take the following Lyapunov function is $\mathbf{V} = \mathbf{s}^2 / 2$. Derivation of this formula, it can get:

$$\begin{aligned} \dot{\mathbf{V}} &= \mathbf{s} \dot{\mathbf{s}} \\ &= -\sum_{i=1}^r \alpha_i [K_1 \mathbf{S} |\mathbf{S}|^\alpha \text{sgn}(\mathbf{S}) + K_2 |\mathbf{S}|^\beta \text{sgn}(\mathbf{S})] \cdot \\ &< 0 \end{aligned} \quad (30)$$

Therefore, global T-S fuzzy model is asymptotically stable.

4.2 FUZZY SLIDING MODE TRACKING CONTROL

According to the characteristics of reaching law approach. Reaching law sliding mode controller is designed based on T-S fuzzy model. Assuming a given input signal is \mathbf{x}_d , error vector is defined as:

Control law is derived as follows:

$$\mathbf{s} = \mathbf{C}\mathbf{E}. \quad (31)$$

Among: $\mathbf{C} = \text{diag}(c_1, c_2, \dots, c_n), c_i (i = 1, 2, \dots, n)$ is constant.

Assuming slaw is reaching law, By derivation and organize the formula (31) can get: $\dot{\mathbf{s}} = \mathbf{C}\dot{\mathbf{E}} = \text{slaw}$.

Through state equation (25) into the above formula, we obtain T-S fuzzy sliding mode controller is $\mathbf{R}^i : \text{IF } x_1 \text{ is } F_1^i \text{ and } x_2 \text{ is } F_2^i \text{ and } \dots \text{ and } x_n \text{ is } F_n^i,$

$$\mathbf{u}_i = (\mathbf{C}\mathbf{B}_i)^{-1}(\mathbf{C}\dot{\mathbf{x}}_d - \mathbf{C}\mathbf{A}_i \mathbf{x}(t) - \text{slaw})$$

Through new fast reaching law (24) into the above formula, T-S fuzzy sliding mode controller based on the new reaching law is:

$$R^i : \text{IF } x_1 \text{ is } F_1^i \text{ and } x_2 \text{ is } F_2^i \text{ and } \dots \text{ and } x_n \text{ is } F_n^i, \tag{32}$$

$$u_i = (CB_i)^{-1} (C\dot{x}_d - CA_1x(t) + K_1S|S|^{\alpha} \text{sgn}(S) + K_2|S|^{\beta} \text{sgn}(S))$$

Among: $0 < \alpha < 1, 0 < \beta < 1, K_1 = \text{diag}[k_{11} \ k_{12} \ k_{13}], K_2 = \text{diag}[k_{21} \ k_{22} \ k_{23}], k_{ji} > 0, (i = 1, 2, 3; j = 1, 2).$

Stability proof:

Take Lyapunov function is $V = \frac{1}{2} s^T s$.

Derivation of this formula and equation (24) into above formula, it can be obtained:

$$\dot{V} = s^T \dot{s} = -K_1 s^T s |s|^{\alpha} \text{sgn}(s) - K_2 s^T |s|^{\beta} \text{sgn}(s) < 0 \tag{33}$$

Each fuzzy sliding mode subsystem is asymptotically stable; each local subsystem global controller is a weighted sum of the control law, based on the above model analysis can be obtained: $u = \sum_{i=1}^r \alpha_i u_i$.

Take the following Lyapunov function is $V = s^2/2$. Derivation of this formula, it can get:

$$\dot{V} = s\dot{s} = -\sum_{i=1}^r \alpha_i [K_1 S |S|^{\alpha} \text{sgn}(S) + K_2 |S|^{\beta} \text{sgn}(S)] < 0 \tag{34}$$

Therefore, global TS fuzzy model is asymptotically stable.

5 Controller Designed and Stability Analysis

5.1 CONTROL MODEL OF SYSTEM

Body coordinates, the BTT cruise aircraft control system simplified mathematical model as follows [4, 12]:

Roll channel mathematical model

$$\begin{cases} \dot{\gamma} = \omega_x \\ \dot{\omega}_x = -c_1\omega_x - c_2\beta - c_3\delta_x - c_4\omega_y - c_5\delta_y \end{cases} \tag{35}$$

Yaw channel mathematical model:

$$\begin{cases} \dot{\beta} = -b_4\beta + \omega_y + \omega_x\alpha - b_3\delta_y - b_7\delta_x \\ \dot{\omega}_y = -b_2\beta - b_1\omega_y + b_6\omega_x - b_3\delta_y - b_8\delta_x + b_7\omega_x\omega_z \end{cases} \tag{36}$$

Roll channel Mathematical model:

$$\begin{cases} \dot{\gamma} = \omega_x \\ \dot{\omega}_x = -c_1\omega_x - c_2\beta - c_3\delta_x - c_4\omega_y - c_5\delta_y \end{cases} \tag{37}$$

Consider the control of the BTT aircraft model (35) ~ (37) If the aircraft longitudinal and lateral loading as the output of the system, the system dynamics do not expect the non-minimum phase phenomenon, at this time feedback linearization theory cannot be directly used for non-minimum phase systems, it will easily lead to system instability.

Therefore select the output vector is $y = [\gamma, \beta, \alpha]^T$. To simplify the problem, ignoring control the lift and side force generated by the rudder, and assuming that the BTT aircraft control model in dynamic coefficient changes with time for a small amount of time to change the system state variables, for each output variable and on derivative of the time variable, Until at least one control volume is $u = [\delta_x, \delta_y, \delta_z]^T$. $x = [w_x, w_y, w_z, \gamma, \beta, \alpha]^T$ is State variables. Using formula (35) - (37) to get:

$$\ddot{y} = \begin{bmatrix} \ddot{\gamma} \\ \ddot{\beta} \\ \ddot{\alpha} \end{bmatrix} = F(x) + G(x)u, \tag{38}$$

where $F(x) = [f_1(x) \ f_2(x) \ f_3(x)]^T$,

$$f_1(x) = -c_1\omega_x - c_2\beta - c_4\omega_y,$$

$$f_2(x) = (b_4^2 - b_2 - c_2\alpha)\beta - (b_4\alpha - b_6 + c_1\alpha - a_4\alpha + \omega_z)\omega_x - (b_1 + b_4 + c_4\alpha)\omega_y,$$

$$f_3(x) = (a_4^2 - a_2)\alpha - (a_4 + a_1)\omega_z,$$

$$G(x) = \begin{bmatrix} -c_3 & 0 & 0 \\ 0 & b_4b_5 - b_3 & -c_3\alpha \\ 0 & 0 & a_4a_5 - a_3 \end{bmatrix}.$$

Other relevant parameters can refer to the literature [1].

5.2 CONTROLER DESIGNED

Consider the BTT aircraft control system model (38), the application of section one section of T-S fuzzy modelling method of BTT aircraft control system, T-S fuzzy model is established.

Select the feature points in the entire flight airspace $t_i (i = 1, 2, \dots, r)$, Feature points on the local subsystem is:

$$\ddot{y} = F_i(X) + G_i(X)u. \tag{39}$$

Following r rules to describe the dynamic behaviour of the system, i fuzzy inference rules are:

$$R^i : \text{IF } t_i \text{ is } F^i, \text{ THEN} \\ \ddot{\mathbf{y}} = \mathbf{F}_i(\mathbf{X}) + \mathbf{G}_i(\mathbf{X})\mathbf{u} \quad (40)$$

where $\mathbf{X} = [\gamma, \beta, \alpha, \dot{\gamma}, \dot{\beta}, \dot{\alpha}]^T$.

Set $\alpha_i(t_i)$ is t_i related to the fuzzy sets membership function of F^i global T-S fuzzy model of BTT aircraft control system is:

$$\ddot{\mathbf{y}} = \sum_{i=1}^r \alpha_i(t_i)(\mathbf{F}_i(\mathbf{X}) + \mathbf{G}_i(\mathbf{X})\mathbf{u}) \quad (41)$$

Conversion aircraft control system model (38). Make $\mathbf{x}_1 = \mathbf{y}$, $\mathbf{x}_2 = \dot{\mathbf{y}}$, the system model can be expressed as:

$$\begin{cases} \dot{\mathbf{x}}_1 = \mathbf{x}_2 \\ \dot{\mathbf{x}}_2 = \mathbf{F}(\mathbf{X}) + \mathbf{G}(\mathbf{X})\mathbf{u} \end{cases} \quad (42)$$

where $\mathbf{x}_1 = \mathbf{y} = [\gamma, \beta, \alpha]^T$, $\mathbf{X} = [\mathbf{x}_1^T, \mathbf{x}_2^T]^T$.

Given input signal is $\mathbf{X}_d = [\mathbf{x}_{1d}^T, \mathbf{x}_{2d}^T]^T$, through the design of control laws, the state \mathbf{X} track the desired state \mathbf{X}_d . Define the error vector as follows:

$$\mathbf{E} = \mathbf{X} - \mathbf{X}_d = [\mathbf{e}^T \ \dot{\mathbf{e}}^T]^T, \quad (43)$$

where $\mathbf{e} = \mathbf{x}_{1d} - \mathbf{x}_1 = [e_1 \ e_2]^T$.

Using reaching law control model, the control law is pushed to the following:

$$\mathbf{S} = \mathbf{C}\mathbf{E}, \quad (44)$$

where $\mathbf{C} = [\mathbf{C}_1 \ \text{diag}(1,1,1)]$ is the matrix, $\mathbf{C}_1 = \text{diag}(c_{11}, c_{12}, c_{13})$, $c_{1i} (i=1,2,3)$ is the normal number.

On the type derivation and finishing available:

$$\begin{aligned} \dot{\mathbf{S}} &= \mathbf{C} \begin{bmatrix} \dot{\mathbf{X}}_d - \dot{\mathbf{x}}_1 \\ \ddot{\mathbf{X}}_d - \ddot{\mathbf{x}}_2 \end{bmatrix} \\ &= \mathbf{C} \begin{bmatrix} \dot{\mathbf{X}}_d \\ \ddot{\mathbf{X}}_d \end{bmatrix} - \mathbf{C}_1\mathbf{x}_2 - \ddot{\mathbf{x}}_2 = \text{slaw} \end{aligned} \quad (45)$$

Finishing (42) and (45) and substituted into the T-S fuzzy inference rules (40), the T-S-type fuzzy sliding mode controller is:

$$R^i : \text{IF } t_i \text{ is } F^i, \text{ THEN} \\ \mathbf{u}_i = \frac{1}{\mathbf{G}_i} \left[\mathbf{C} \begin{bmatrix} \dot{\mathbf{X}}_d \\ \ddot{\mathbf{X}}_d \end{bmatrix} - \mathbf{C}_1\mathbf{x}_2 - \mathbf{F}_i(\mathbf{X}) - \text{slaw} \right] \quad (46)$$

The new fast reaching law (24) into the above equation, T-S fuzzy sliding mode controller based on the new reaching law for BTT Aircraft is:

$$R^i : \text{IF } t_i \text{ is } F^i, \text{ THEN} \\ \mathbf{u}_i = \frac{1}{\mathbf{G}_i} \left[\mathbf{C} \begin{bmatrix} \dot{\mathbf{X}}_d \\ \ddot{\mathbf{X}}_d \end{bmatrix} - \mathbf{C}_1\mathbf{x}_2 - \mathbf{F}_i(\mathbf{X}) + K_1|\mathbf{S}|^\alpha \text{sgn}(\mathbf{S}) + K_2|\mathbf{S}|^\beta \text{sgn}(\mathbf{S}) \right], \quad (47)$$

where $k_{ji} > 0$, $K_2 = \text{diag}[k_{21} \ k_{22} \ k_{23}]$, $(i=1,2,3; j=1,2)$; $0 < \alpha < 1$, $0 < \beta < 0$.

Take the following Lyapunov function $V = s^2/2$.

This derivation and reaching law formula (24) into the above formula, you can have:

$$\begin{aligned} \dot{V} &= s\dot{s} \\ &= -K_1\mathbf{S}|\mathbf{S}|^\alpha \text{sgn}(\mathbf{S}) - K_2|\mathbf{S}|^\beta \text{sgn}(\mathbf{S}) \\ &< 0 \end{aligned} \quad (48)$$

Each fuzzy sliding mode subsystem is asymptotically stable; the global controller for each local subsystem control law is weighted and sum:

$$\mathbf{u} = \sum_{i=1}^r \alpha_i \mathbf{u}_i \quad (49)$$

Take the following Lyapunov function $V = \frac{1}{2}s^2$. This type derivation, we can obtain:

$$\begin{aligned} \dot{V} &= s\dot{s} \\ &= -\sum_{i=1}^r \alpha_i [K_1\mathbf{S}|\mathbf{S}|^\alpha \text{sgn}(\mathbf{S}) + K_2|\mathbf{S}|^\beta \text{sgn}(\mathbf{S})] \\ &< 0 \end{aligned} \quad (50)$$

Therefore, global T-S fuzzy model is asymptotically stable.

6 Controller simulation

In order to verify the tracking performance of the system, get parameters and mathematical models a certain type of BTT aircraft aerodynamics during the flight. Set the roll angle tracking instructions is $\gamma = 120^\circ$, 0.25Hz square wave signal; angle of attack tracking instructions for the amplitude is 0.25Hz square wave signal, $\alpha = 10^\circ$, Sideslip angle $\beta = 0$.

Throughout the flight, Select five characteristic points constitute fuzzy T-S model of aircraft system; Design membership function is triangular membership functions, According to the T-S fuzzy model rules, as in Figure 3.

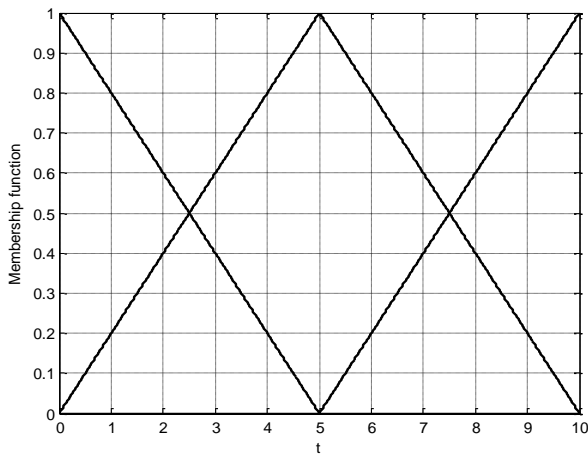


FIGURE 3 Fuzzy membership function in simulation

Using pole placement method, obtain the matrix $C = [C_1 \ C_2]$, $C_1 = \text{diag}[10 \ 2 \ 5.2]$, $C_2 = \text{diag}[1 \ 1 \ 1]$. Select $K_1 = \text{diag}(18, 6, 18)$, $K_2 = \text{diag}(2, 1, 2)$.

The simulation results shown in Figure 4 and Figure 5. The simulation results show that the roll angle γ and α angle of attack can quickly track the command signal.

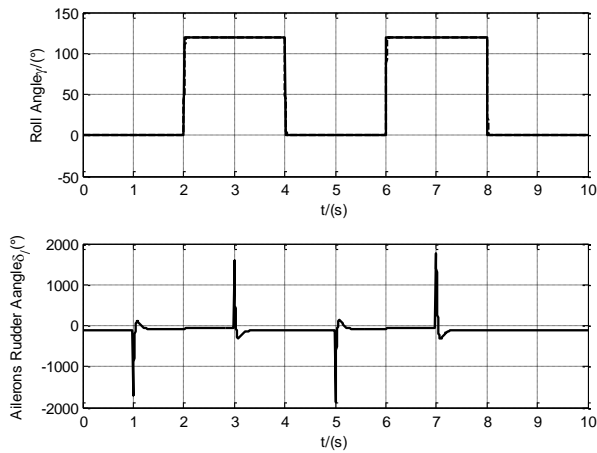


FIGURE 4 Roll angle tracking and aileron rudder angle response curve

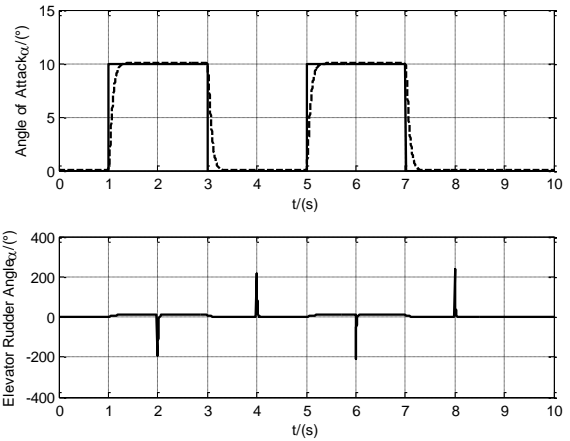


FIGURE 5 Angle of attack tracking and pitch rudder angle response curve

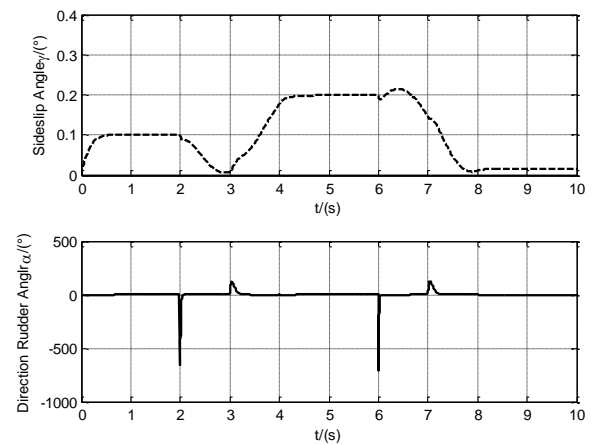


FIGURE 6 Sideslip angle tracking response and yaw rudder angle response curve

It can be seen from Figure 6, the sideslip angle β remain at $|\beta| \leq 3^\circ$ on to meet the performance requirements of the system, it can effectively control for BTT aircraft.

7 Conclusion

BTT aircraft is a multi-variable nonlinear coupling object, to solve this problem, a new BTT aircraft controller is been designed based on Control method of sliding mode control theory and fuzzy T-S. Reaching law proposed by the T-S fuzzy sliding mode controller design method applied to BTT aircraft controller design is feasible, simulation results show that Control effect is better in the aircraft control, and it is able to fast track Guided instruction and Meet the design requirements.

References

[1] Yu Xiuping, Guan Shuangying. 2008 Research on autopilot of BTT aircraft based on H_∞ control theory *Systems Engineering and Electronics* 30(5) 503-7 (in Chinese)
 [2] Zang You'an, Hu Yun'an 2003 *Nonlinear design approaches for aircraft control and guidance* Beijing: Industrial Publisher 68-74

[3] Lee C C A 1991 *International Journal of Intelligent Systems* 6(1) 71-93
 [4] Zheng Wenda, Liu Gang 2010 Bank-to-Turn Aircraft Control System Design Based on Improved Fuzzy Basis Function Networks *2010 International Conference on Future Information Technology and Management Engineering* 484-7
 [5] Ying H 1998 *IEEE Trans Fuzzy Syst* 6 582-7
 [6] Zhen Qi, McInroy J E, Jafari F 2007 *Journal of Intelligent and Robotic Systems* 48 (3) 333-56

[7] Zhou Jun, Wang Tin 2007 Integrated Guidance Control System for Beam-riding Guidance Aircraft Based on Second Order Sliding Mode Control *Journal of Astronautics* 6(18) 1632-7 (in Chinese)

[8] Lin C K 2009 H-infinity reinforcement learning control of robot manipulators using fuzzy wavelet networks *Fuzzy Sets and Systems* 160(12) 1765-86

[9] Boulkroune A, M'saad M, Farza M 2011 Adaptive fuzzy controller for multivariable nonlinear state time-varying delay systems subject to input nonlinearities *Fuzzy Sets and Systems* 164(1) 45-65

[10] Jiang Jun, Chen Qing wei, Guo Jian et al 1996 Sliding mode stable tracking control for mobile satellite communication system based on a new reaching law *Control and decision* 19(5) 1056-63 (in Chinese)

[11] Tong Shaocheng, Wang Tao, Wang Yanping. 2004 *Fuzzy control system design and stability analysis* Beijing: Science Publishing house 17-21(in Chinese)

[12] Dong Xiangyu, Ren li, Liu Yunfei, et al. 2009 Simulation of BTT aircraft based on variable structure adaptive control *Computer Simulation* 11(26) 65-8 (in Chinese)

[13] Mirseidova Sh, Inoue A, Atymtayeva L 2012 Evaluation of fair market price of resources in oil and gas industry using fuzzy sets and logics *Computer Modelling and New Technologies* 16(1) 30-4

Authors	
	<p>Hongqing Hou, born on November, 1979, Xi'an City, Shan'xi Province, P.R. China</p> <p>Current position, grades: Postdoctoral of Xi'an Research Inst. of hi-Tech hongqing Town, China. University studies: received his D.E. in Control Science and Engineering from Xi'an Research Inst. of hi-Tech hongqing Town in China. Scientific interest: His research interest fields include precision-guided and control, mechanotronics. Publications: more than 10 papers published in various journals. Experience: He has completed three scientific research projects.</p>
	<p>Qian Miao, born on January, 1985, Xianyang City, Shan'xi Province, P.R. China</p> <p>Current position, grades: lecturer of Xi'an Research Inst. of hi-Tech hongqing Town, China. University studies: received her M.E. in Software Engineering from North Western Polytechnical University in China. Scientific interest: her research interest fields include Software Engineering, Data Structure. Publications: more than 8 papers published in various journals. Experience: She has teaching experience of 3 years, has completed two scientific research projects.</p>
	<p>Chuanqiang Yu, born on January, 1975, Yantai City, Shandong Province, P.R. China</p> <p>Current position, grades: the Associate Professor of Xi'an Research Inst. of hi-Tech hongqing Town, China. University studies: received his D.E. in Armament Science and Technology from Xi'an Research Inst. of hi-Tech hongqing Town in China. Scientific interest: His research interest fields include Motor and Electrical Technology, Hydraulic Technology. Publications: more than 30 papers published in various journals. Experience: He has teaching experience of 8 years, has completed ten scientific research projects.</p>
	<p>Qinhe Gao, born on July, 1968, Yinan City, Shandong Province, P.R. China</p> <p>Current position, grades: Professor of Xi'an Research Inst. of hi-Tech hongqing Town, China. University studies: received his D.E. in Armament Science and Technology from Xi'an Research Inst. of hi-Tech hongqing Town in China. Scientific interest: His research interest fields include Motor and Electrical Technology, Fault Detection and Diagnosis. Publications: more than 80 papers published in various journals. Experience: He has teaching experience of 16 years, has completed nineteen scientific research projects.</p>

A printer reverse characterization model based on BP neural network

Lei Zhao, Guangxue Chen*

State Key Laboratory of Pulp and Paper Engineering, South China University of Technology, 510640, Guangzhou, P.R.China

Received 1 March 2014, www.tsi.lv

Abstract

For colour printer, there are very complicated nonlinear relation between its printed colour chromatic values and input digital image pixel values. In the research, data sets of printed colour chromatic values and their digital image pixel values are classified by hue angle range, the data in each hue angle range is taken as learning samples to create BP neural network. With improved combined method of additional momentum factor and variable learning rate, BP neural network of each hue angle range is trained and created. The experiment result shows that, with appropriate structure and classified learning samples, the reverse characterization model based on ten BP neural networks can be trained in relative short time; the colour errors between the experimental printed colour chromatic values and computed printed chromatic values are far less than the threshold of human eyes, i.e. the reverse characterization model achieves rather high accuracy.

Keywords: BP neural network, Hue angle range, Data classification, Colour management

1 Introduction

The working principle of printer is to output colour according to the input drive value [1]. There are two device models to describe the relationship of the input drive values and the printed colour chromatic values of printer. One is from input drive values to printed colour chromatic values, it predicts printed colour chromatic values according to input drive values, and it is characterization model. Another one is from printed colour chromatic values to input drive values, it lookups input drive values according to printed colour chromatic values; it is reverse characterization model [2]. At present, both the models are mainly based on linear fitting method, polynomial regression method and look-up table method etc. These methods have made a certain effect, in spite of their own limitations. Linear fitting method has low conversion accuracy and practicality [3], the coefficients of polynomial regression method are difficult to determine, and the computation is intensive [4], the look-up table method requires a very large amount of data and high system resources occupancy to create model and to attain its higher accuracy [5].

Printers differs in working principle, interior structure, and driver and so on, so a same digital image pixel is inputted, different printers are likely to show different colours. The colour range represented by printer shows different area sizes and shapes as gamuts [6]. A colour management system adopts both the models [7]. In the research, colour chromatic value is described by CIE 1976 $L^*a^*b^*$ chromatic system.

2 Basic principle of BP neural network

The artificial neural network includes mass interconnected simple components, it is built to achieve a certain function on the basis of mathematics knowledge and how human brain works and constitutes.

2.1 NEURON

Neuron is the basic unit of neural network; it is generally nonlinear component with multi-input and single-output. A typical neuron model is showed in Figure 1, containing R input components and an offset b , the output of activation function f is a . The input of the activation function f includes the input components $p_j(j=1, 2, 3, \dots, R)$ and the weight value component $w_j(j=1, 2, 3, \dots, R)$, the net input of the activation function f is $n=Wp+b$. Wherein, W and p are respectively the vector form of the weight value component w_j and the input component p_j .

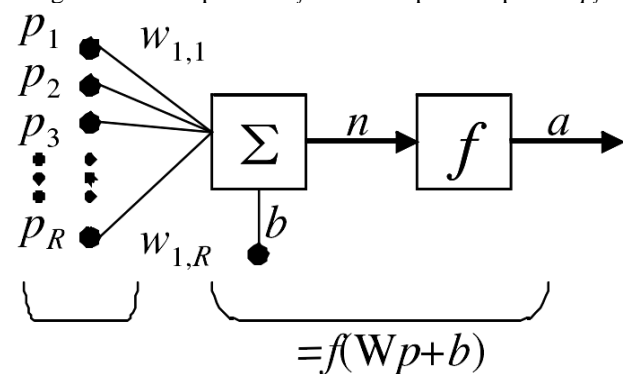


FIGURE 1 Neuron Model

*Corresponding author e-mail: chengx@scut.edu.cn

The vector form of neuron output a is:

$$a = f(Wp + b) = f\left(\sum_{j=1}^r w_j p_j + b\right). \tag{1}$$

Typically, the offset b is a fixed constant as 1; it plays an important role in the neural network [8]. In the study, all the BP neural networks contain offset.

2.2 ACTIVATION FUNCTION OF NEURONS

Activation function is the key factor of neurons and neural networks; it controls the activation, conversion and output of neuron input component, and converts infinite input to limited output. It determines the nature and capacity of neurons and neural networks.

Common activation function includes linear function, threshold function, Sigmoid function and hyperbolic tangent function, which can be selected according to the problem to be solved. In the research, nonlinear activation functions are used to convert colour chromatic value to digital image pixel value [9].

2.3 BP NEURAL NETWORK

BP neural network, i.e. back propagation neural network, belongs to feedback neural networks. There is full



FIGURE 2 Printed Colour Patches

Image with colour patches are printed on photo-printing paper by EPSON Stylus Pro 7600 high-resolution printer, with black, shiny black, magenta, light magenta, green, bright blue, yellow ink, at 720×720dpi resolution. After ink drying, under a D65 illuminating 10° visual field, chromatic values of colour patches are measured by X-Rite 528 spectral density meter and recorded as Excel 2003 files.

3.2 DETERMINATION OF BP NEURAL NETWORK STRUCTURE AND TRAINING PARAMETERS

According to BP neural network theory, its simulation and approximation ability are closely related with its hidden layer number and its hidden layer neurons

interconnection between each adjacent layer neurons, but there is no direct connection between neurons of the same layer and between the input and output layers. In its training process, according to the error between the network output and target values, Wildrow-Hoff generalized learning rule is used to modify nonlinear differential function weights repeatedly, until it reaches the set target.

The neurons number of input layer and output layer is determined by specific problems. Besides the input layer and output layer, several hidden layers are introduced to enhance the BP neural network simulation and computing ability. In theory, as long as it has sufficient hidden layers and hidden layers have enough neurons, it can approximate any function or data relationship [10].

3 Experiment

3.1 LEARNING SAMPLE ACQUISITION OF BP NEURAL NETWORK

Based on 24-bit deep BMP format RGB digital image pixel value, colour patches are generated. Pixel values are selected as approximate equal interval of ten points, which are 0, 28, 56, 84, 112, 140, 168, 196, 224 and 255 respectively. By combining the pixel values of the three RGB channels, 1000 patches are shown as Figure 2.

number. In general, the more the hidden layer number and the hidden layer neurons number is, the stronger its data processing and simulation capability is [10]. In the study, BP neural network will convert printed colour patches chromatic values into digital image pixel values. Printed colour patches chromatic values $L^*a^*b^*$ distribute in the range of (0,100), (-120,120) and (-120,120), meanwhile the digital image pixel values RGB distribute in the range of (0,255), (0,255) and (0,255). Both the hidden layer number and the hidden layer neurons number of BP neural network are necessary to be increased appropriately to enhance the simulation and computation ability, to describe the relation between the two data sets with such feature.

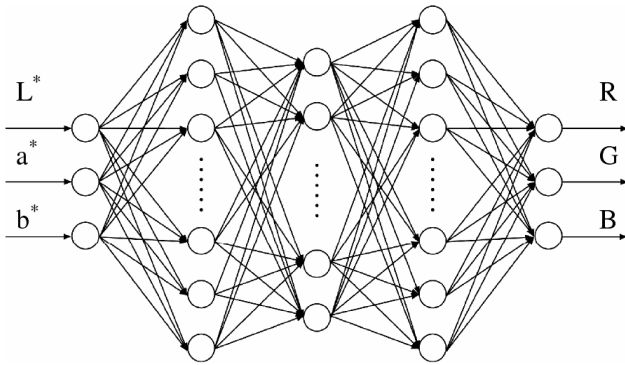


FIGURE 3 BP network structure

In the study, BP neural networks with one hidden layer are not adopted. Referring to the literatures of BP neural network in colour management [11], considering the factors of computing ability and learning sample characteristics, the BP neural network is determined and shown in Figure 3. The neurons of each layers is 3-30-20-30-3. Both the neurons number of input layer and output layer are three, which corresponds to printed colour chromatic value $L^*a^*b^*$ and digital image pixel value RGB respectively. BP neural networks are created by MatLab 12.0 neural network tools, its parameters are shown in Table 1.

TABLE 1 Parameters of BP neural network

Items of BP Neural Network		Parameters
Structure	Input Layer Data	$L^* a^* b^*$
	Input Neuron Number	3
	Output Layer Data	R G B
	Output Neuron Number	3
Functions	Hidden Layer1	Tansig
	Hidden Layer2	Tansig
	Hidden Layer3	Tansig
	Output Layer	Purelin
	LearningFcn	Traingdx
	performFcn	MSE function
	trainParam.lr	0.05
	trainParam.lr_inc	1.05
	trainParam.lr_dec	0.7
	trainParam.max_fail	5
Learning Process	trainParam.max_perf_inc	1.04
	trainParam.mc	0.9
	trainParam.min_grad	1e-10
	trainParam.epochs	1000000
	trainParam.goal	0.0005
	trainParam.time	Inf
	show	10

3.3 CLASSIFICATION OF LEARNING SAMPLE AND TRAINING OF BP NEURAL NETWORK

3.3.1 Preliminary training and analysis of BP neural network

Under the supervision of learning sample data, the BP neural network is trained by all learning sample in Matlab 12.0. In the training, the objective function error converged slowly for a long time or even stagnated, so that the BP neural network cannot be trained

successfully. According to neural network theory and practical experience [12], the BP neural network has stronger capability; but the amount of learning sample is too large to train the network successfully. In addition, how learning samples distribute in three-dimensional space can also explain the reason of training failure. The distribution of printed colour patches chromatic values and pixel values in the three-dimensional space are shown in Figure 4a and Figure 4b respectively. It can be seen from Figure 4, the digital image pixel values distribute in the space of a cube evenly, but printed colour patches chromatic values distribute in the specification space unevenly.

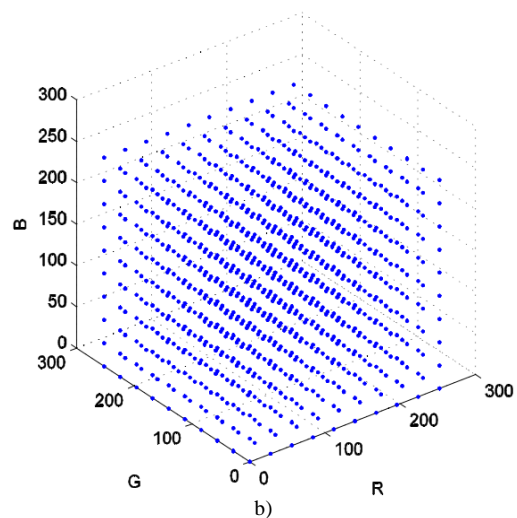
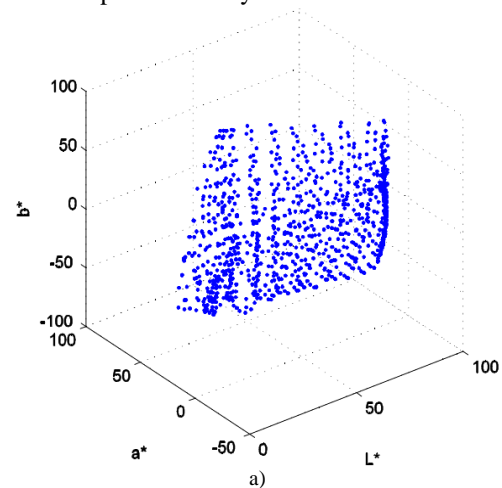


FIGURE 4 Distribution of learning samples in three dimension space: a) Printed colour patches chromatic values; b) Digital image pixel values

Printer reverse characterization model converts printed colour patch chromatic value into pixel value. As can be seen from Figure 4, if a single BP neural network is used to convert unevenly distributed data into evenly distributed data, both the hidden layer number and hidden layer neurons number should be increased. Even if the BP neural network is trained successfully, the neural network training time will increase and the operating efficiency will reduce when the model works.

3.3.2 Learning samples classification by hue angle range

In order to improve the training success rate and shorten the training time, referring to the learning sample classification method by lightness [13], all the learning samples are classified by chromatic value of printed colour patches. Hue angle of learning sample is calculated as formula:

$$\alpha = \tan^{-1}(b^*/a^*) \tag{2}$$

Take L* axis as centre, for instance, CIE 1976 L* a* b* colour space is divided into 8 equal parts according to hue angle range. The view on the positive direction of the L* axis is shown in Figure 5, wherein the hatched line portion represents a hue angle range.

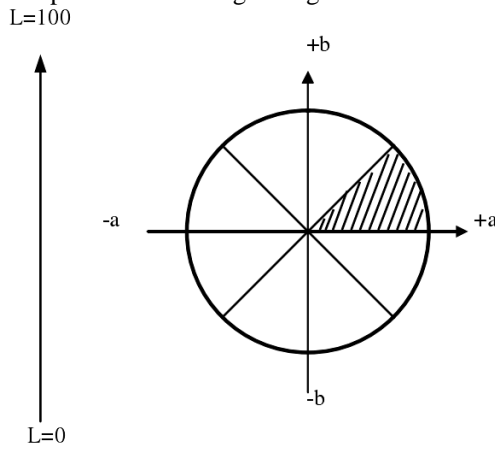


FIGURE 5 Hue angle range

In the study, CIE 1976 L* a* b* colour space is divided into 10 equal parts; each part is taken as the main part of a hue angle range. The hue angle range radian of No. $i(i=2, 3, 4 \dots 9)$ is:

$$[2\pi((i-1)/10) - TFactor, 2\pi(i/10) + TFactor] \tag{3}$$

where TFactor is the redundancy factor. It is taken as one tenth of a single hue angle range, i.e. $(1/100) * 2\pi$. For instance, the radian of 6th hue angle range is:

$$[2\pi(5/10) - 0.02\pi, 2\pi(6/10) + 0.02\pi] \tag{4}$$

Meanwhile, the radians of 1st and 10th hue angle range are respectively:

$$[2\pi - 0.02\pi, 2\pi] \cup [0, 2\pi(1/10) + 0.02\pi] \tag{5}$$

$$[2\pi(9/10) - 0.02\pi, 2\pi] \cup [0, 0.02\pi] \tag{6}$$

The redundancy factor is introduced to train network in a relative wider range, and the BP neural network works in relative narrower range, in order to improve neural network accuracy [13].

The learning sample is classified into ten sets, as learning samples of the BP neural networks within each

hue angle respectively. Thus, the printer reverse characterization model consists of ten sub-models. When the reverse characterized model is implemented, hue angle is obtained according to formula (2) to determine the hue angle range where printed colour patches chromatic values locates, and then digital image pixels values is calculated by the sub-model which corresponds to the hue angle range.

How the learning samples distribute in the ten hue angle ranges of CIE 1976 L* a* b* colour space is shown in Table 2.

TABLE 2 Distribution of learning samples in each hue angle range

No. of hue angle range	Hue angle range (Unit: radian)	Learning samples number
1	$(6.220, 2\pi) \cup (0, 0.691)$	75
2	$(0.565, 1.319)$	101
3	$(1.193, 1.948)$	132
4	$(1.822, 2.576)$	232
5	$(2.450, 3.204)$	98
6	$(3.078, 3.832)$	58
7	$(3.707, 4.461)$	53
8	$(4.335, 5.089)$	88
9	$(4.964, 5.718)$	217
10	$(0, 0.062) \cup (5.592, 2\pi)$	135

It can be seen from Table 2, due to the introduction of redundancy factor, learning sample amount in all hue angle ranges is 1189, which is more than the learning samples amount 1000. It is because part data is in the hue angle range edged, which belongs to two adjacent hue angle ranges at the same time.

3.3.3 Distribution of classified learning sample in three-dimensional space

How the 10 sets of learning sample data distribute in three-dimensional space is shown in Table 3. On the view of learning samples distribution in three-dimensional space, the relation of all learning samples is converted into the relation of learning samples in each hue angle range.

4 Experimental results and analysis

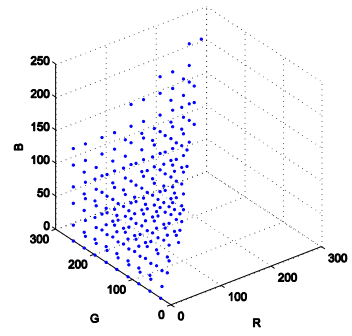
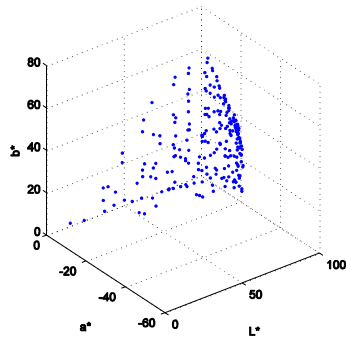
4.1 TRAINING PROCESS EVALUATION OF BP NEURAL NETWORK IN EACH HUE ANGLE

In order to improve the BP neural networks training speed, learning samples in each hue angle range are normalized to the range (-1,1) [13], combined training method of additional momentum factor and variable learning rate are adopted. The training process parameters and training process features of the ten BP neural networks are shown as Table 4 and Table 5.

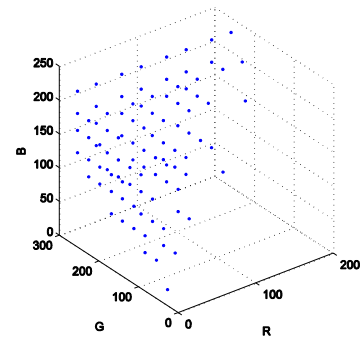
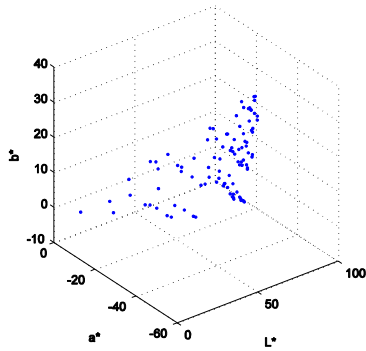
TABLE 3 Learning Samples distribution in three-dimensional space in each hue angle range

Hue angle range No.	Printed Patches $L^*a^*b^*$ Values	Digital Image Pixel RGB Values
1		
2		
3		

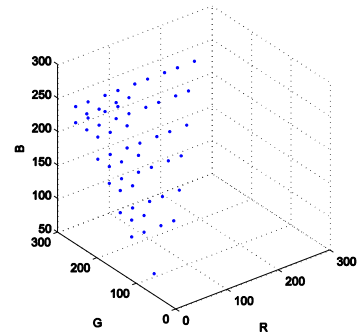
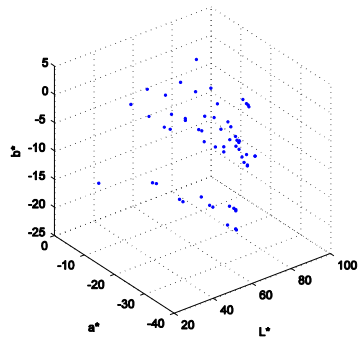
4



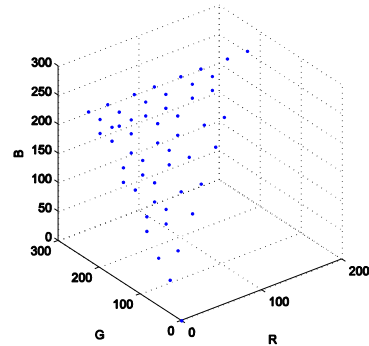
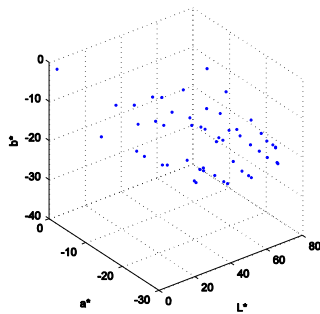
5



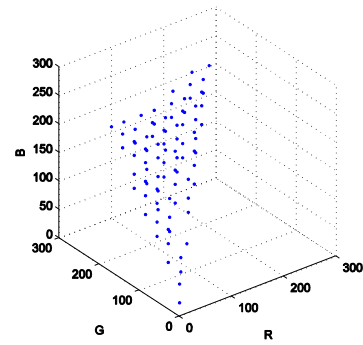
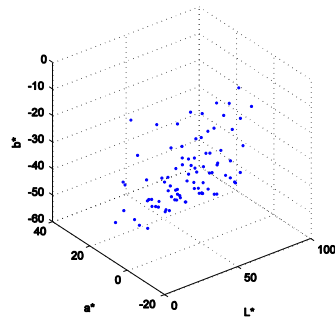
6



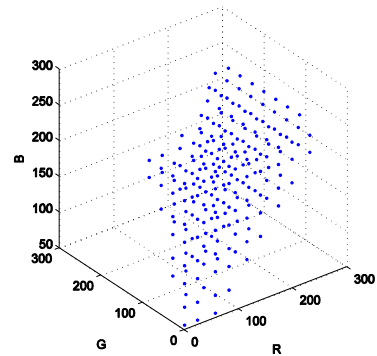
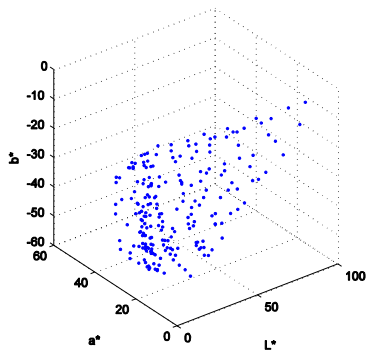
7



8



9



10

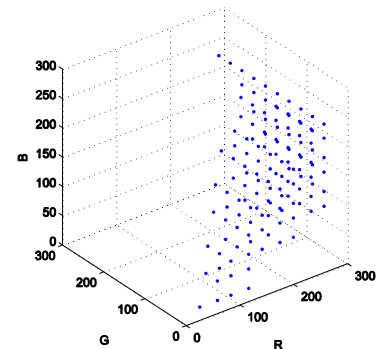
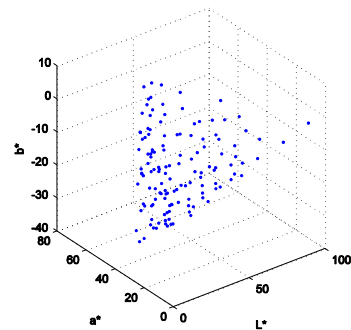


TABLE 4 Training process parameters of the BP neural network of each hue angle range

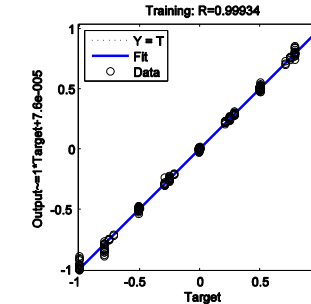
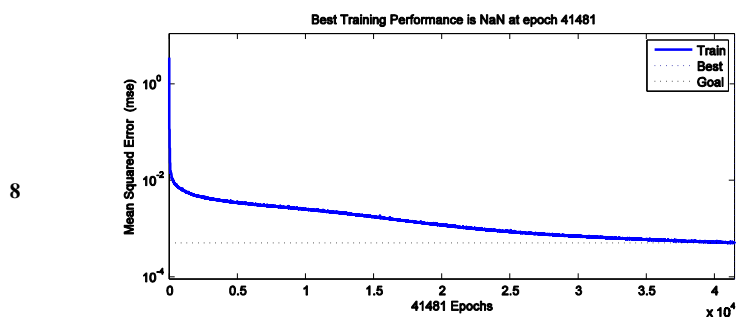
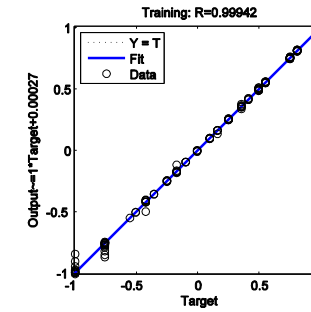
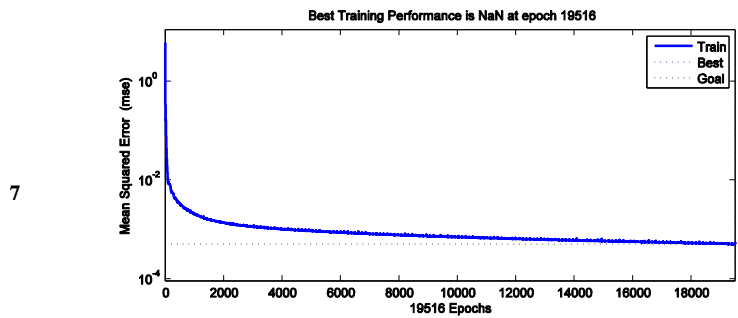
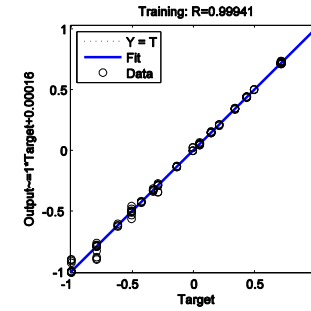
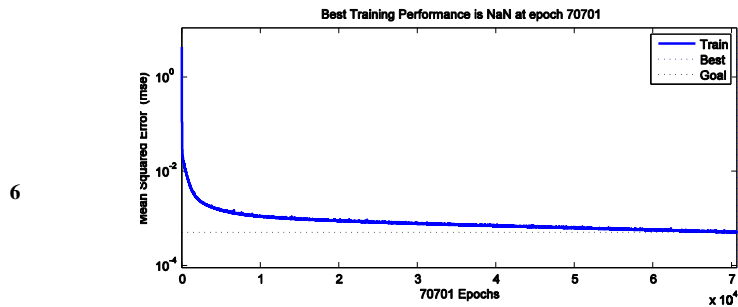
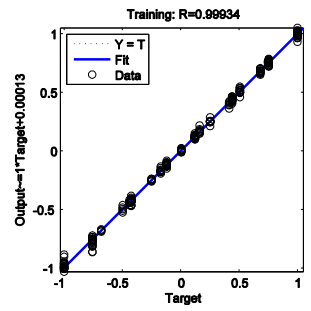
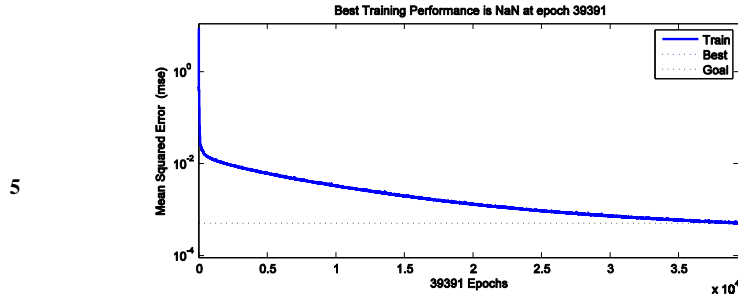
Hue angle range No.	Epoch number	Training time	MSE function performance
1	40768	0:05:24	5.08
2	37281	0:05:12	10.4
3	17995	0:02:22	4.17
4	40582	0:06:45	8.90
5	39391	0:05:05	8.80
6	70701	0:11:24	4.28
7	19516	0:02:07	5.99
8	41481	0:07:08	3.49
9	143060	0:47:03	5.46
10	224598	1:12:10	4.95

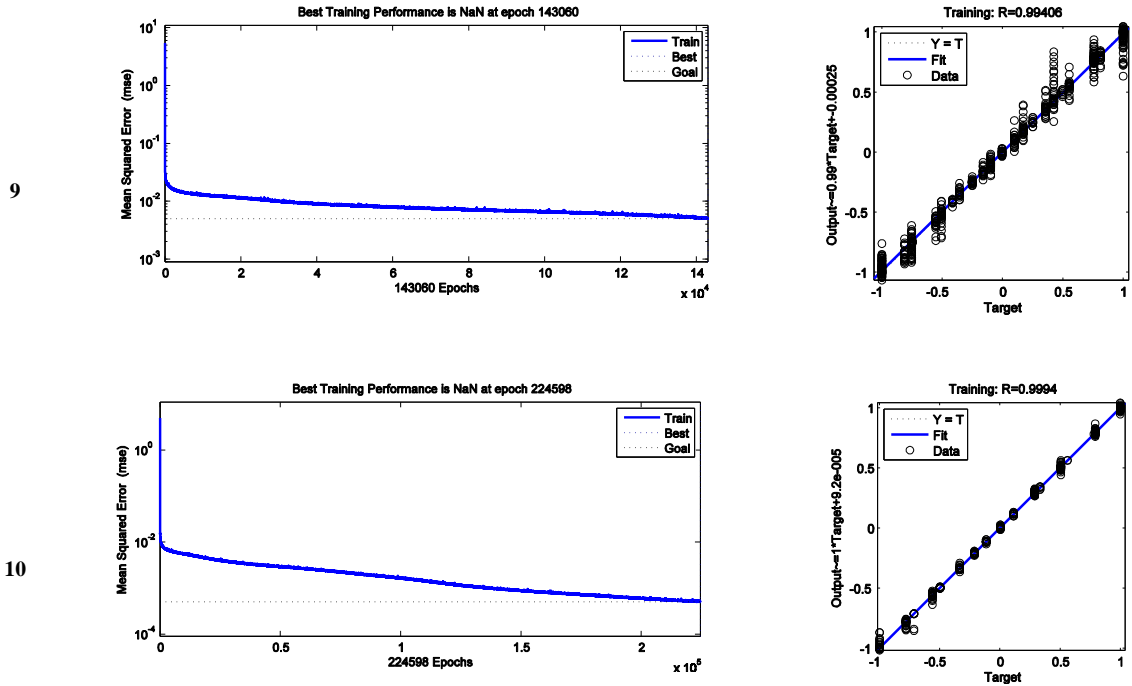
As can be seen from Table 4, all the ten BP neural networks trained successfully, and their training time distribute in the range (0:02:07, 1:12:10). In general, the learning sample amount affects the training time of the

neural network directly, so does how the learning samples distribute in three-dimensional space. Take the neural network with shortest and longest training time as example, their training time is 0:02:07 and 1:12:10, their neural network corresponds to the hue angle rang of 7 and 10 respectively. The learning sample amount of the neural network of the shortest training time is indeed the least, it is 53. Meanwhile, the learning sample amount of the neural network of the longest training time is not the most; it is 135, which is far less than that of hue angle range 4. Though the learning sample amount of hue angle range 4 is the most, its training time of is only 0:06:45. So that, how the learning samples distribute in three-dimensional space affects the training time more than the learning sample amount does. Thus, qualitative relationship between the training time and the learning sample distribution in the three-dimensional space can be concluded from Table 4.

TABLE 5 Training process characteristic of the BP neural networks of each hue angle range

Hue angle range No.	MSEREG of the network during learning	Correlation coefficient of target data and output data
1		
2		
3		
4		





As can be seen from Table 4 and Table 5, the learning samples of the networks with shorter training time distribute more concentrated and has a more consistent distribution characteristics. While the learning samples of the networks with longer training time distribute more dispersed, and the distribution characteristics is less consistently, which is the root cause that result in different training time. It is also shown in the MSE functions, i.e. generally, the larger the performance values are and the smaller the learning samples amount are, the shorter the training time is, and vice versa. The more important factor that affects the training time is the distribution characteristics in three-dimensional space of the learning sample.

In addition, from Table 5, to the fitting diagram of BP neural network output value and output learning sample, in spite of all neural network with the same parameters and structure, the processing power of learning samples of different numbers is closely related to the learning samples amount. Limited by the same neural network structure, the training data of BP neural network of hue angle range 4 and 9 with maximum learning sample amount is most dispersed, their fitting is the worst of all the ten BP neural networks, and vice versa, the less the learning sample amount is, the closer the neural network output value and the output learning samples are, i.e. the better the fitting is.

The gradient decline curve of MSE function reflects the complex relationship of the number of iterations

epochs, training time and MSE function performance. The smoother the curves are, the shorter the training time is. MSE function also shows relatively good performance, such as the networks in hue angle range 3 and 7, and vice versa, the longer the training time is, such as the networks in hue angle range 9 and 10. The MSE function curves could be used to determine whether training process should be terminated timely and re-started, without spending unnecessary time waiting for the training failure.

4.2 EVALUATION OF PRINTER REVERSE CHARACTERIZATION MODEL

81 sets of chromatic values are selected from the printer gamut randomly, and input into printer reverse characterization model. Colour patches contain calculated pixel values are printed and their chromatic values are measured. The colour errors between colour patches to be printed and the printed colour patches are calculated and shown as Figure 6 according to their colour patches No.. The maximum colour error, minimum colour error and average colour error are 5.126, 0.576 and 2.137 respectively. As can be seen from Figure 6, the colour patches number whose colour error less than 4 is 76, and most of them concentrate within 3. Considering the printer repeatability, printer reverse characterization model shows rather high conversion accuracy.

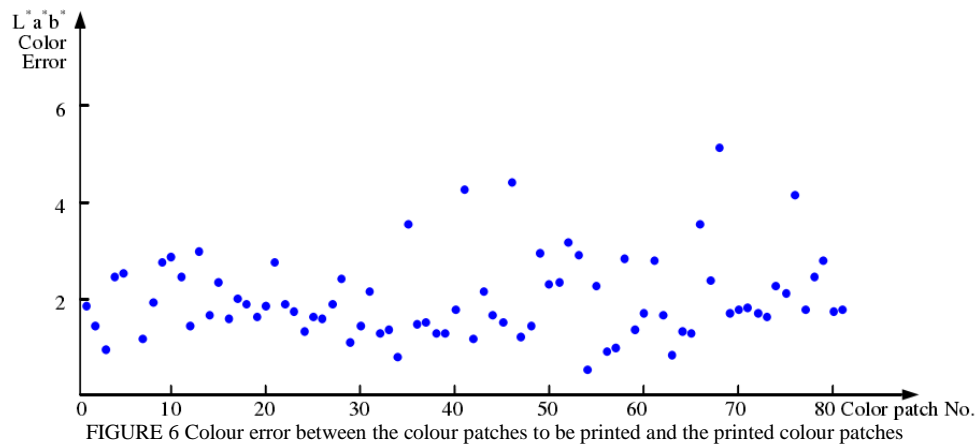


FIGURE 6 Colour error between the colour patches to be printed and the printed colour patches

5 Conclusions

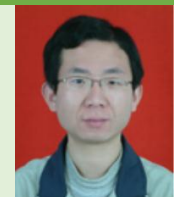
In the study, according to the distribution characteristics of the learning samples in three-dimensional space, the reason why a single BP neural network cannot be trained successfully is analysed. With improved combined training methods of additional momentum factor and variable learning rate, a printer reverse characterization model based on ten BP neural networks is proposed, whose learning samples are classified by hue angle range.

The learning sample classified method not only solves the problem of training failure, but also shortens the training time. Due to the good nonlinear approximation characteristics of BP neural network, the reverse characterization model reaches a high accuracy, and the average colour error between the experimental measurement values and the model calculated values is far less than the threshold that the human eye can perceive.

References

- [1] Liao N F, Shi J S 2009 *Digital Graphic Image Colour Management System Introduction* Beijing Institute of Technology Press: Beijing China Ch 7 (in Chinese)
- [2] Zhou S S 2013 *Printing colour science* Printing Industry on Press: Beijing China Ch 9(in Chinese)
- [3] Xu Y F, Liu W Y 2004 *Opt. Prec.Eng.* **12**(3) 265-9(in Chinese)
- [4] Han K 1998 Characterizing the desktop colour printer with polynomial regression *IEEE Proc. Syst. Man Cy.* **5** 4369-72
- [5] Huang P 1993 *Electron. Imaging* **2**(1) 53-61
- [6] Li X Z, Chen G X, Chen Q F 2010 Digital printing gamut mapping technology *China Printing Pack. Res.* **2**(3) 8-13(in Chinese)
- [7] Jan M 2008 *Colour Gamut Mapping* John Wiley & Sons Ltd: Chichester Ch 8
- [8] Ma R 2012 *Artificial Neural Network*, Mechanical Industry Press: Beijing China Ch 8 (in Chinese)
- [9] Wang X 2007 *Artificial Neural Network Theory and Application* Northeastern Univ. Press: Shenyang Ch 5(in Chinese)
- [10] Zhou K L and Kang Y H 2010 *Neural Network Model and Its Matlab Simulation Program Design* Tsinghua Univ. Press: Beijing Ch 6(in Chinese)
- [11] Liao N F 2000 CRT Chromaticity transformation based on multi-layers BP neural network *J. Image Graph.* **5**(6) 470-3(in Chinese)
- [12] Zhou S Q, Zhao D Z 2000 Printer colour control technology based on BP neural network *Opt. Technol.* **26**(1) 49-51(in Chinese)
- [13] Huang Q M, Zhao D Z 2003 Artificial neural network as printer colour space conversion tool *Opt. Technol.* **29**(2) 146-8(in Chinese)

Authors



Lei Zhao, born on May10, 1978, Hebei, P.R.China

Current position, grades: South China University of Technology, Doctor
University studies: Xi'an University of Technology (Faculty of Printing and Packaging engineering), 1996
Master degree in printing engineering, Xi'an University of Technology, Xi'an China, 2005
Scientific interest: Digital Media technology, Color Management
Publications: more than 80 papers in academic journals



Guangxue Chen, born on February 13, 1963, Qi County, Henan, P.R.China

Current position, grades: South China University of Technology, Professor
University studies: Zhengzhou Institute of Surveying and Mapping(Dept. of Cartography and printing), 1983;Master degree, Northwestern University, Xi'an China, 1995; Doctor degree, P.L.A Information Engineering University, Zhengzhou China, 2005
Scientific interest: Cartography and printing maps electronic publishing, digital printing, information visualization, geographic information technology
Publications: more than 10 papers in academic journals

Communication technology in the application of the smart grid

LiLi Chu*, LiLi Qin

School of Electronics and Information Engineering, Liaoning University of Technology, Jinzhou, 121001, China

Received 1 March 2014, www.tsi.lv

Abstract

By studying the significance of the smart grid, combined with a regional substation point location, line conditions, existing and future business development, existing communications equipment status, etc., the author initially sets the smart grid communications infrastructure deployment and network planning, in order to use the most reasonable communication technologies to support rapid development of smart grid. In a certain city with electric power communication network to the actual construction goal, we should complete the city power system communication network covering the whole deployment. At the same time combined with network energy efficiency project, we should analyse the already formed network, provide effective optimization model of energy efficiency and practical algorithm, and analyse its rationality through the simulation analysis, further improve the overall network in order to make it efficient to run.

Keywords: Smart grid, Data network, Network energy efficiency, Energy efficient routing strategy

1 The meaning of the smart grid communication network construction

The smart grid is the high integration of traditional electric power industry and the information communication. It will become a comprehensive configuration platform of energy and information, makes the grid infrastructure in a larger scope, and serves the society in more ways. Among them, intelligent power grid transformation is supported by information and communication technology, based on intelligent control means, covering all aspects of power grid. We should improve its informatization, automation, interactive level in order to meet the demand for electricity on the basis of economic and social development and people's lives.

Electric power communication network is an important part of strong smart grid. With the continuous development of smart grid technologies, the business application system has been improved gradually, and higher requirements on the network bandwidth and reliability etc. are put forward. Communication network is an important component of the electric power communication network platform and is an extension of the electric power communication network backbone. However, because of its wide coverage, multiple nodes, difficult construction and other reasons, it has been lack of applicable communication technology and construction mode for the electric power communication network, and has become the bottle neck restricting the application and development of electricity business (see, e.g. [1-3]). Power communication network is an important foundation to promote the intelligent power grid construction, protection grid production, operation, management and supply services. Because the units

unable to fully grasp the modern communication technology and the development trend, there are many problems in the construction of electric power communication network and in grasping the technical direction in the current and different regional electric power communication network construction mode, the level of equipment, as well as the application effect are different; Companies generally do not pay attention to the choice of communication technology and network, the work way of thinking is not clear, communication network construction is lack of system planning, the cost of construction and operation is high, the input and output does not match, the overall level of communication network lag behind the development requirements; Some companies hire public network channel resources to transport important business information, which will bring risks to the power grid operation and management services. All of these were not adapted to the development of strong smart grid, to strengthen and perfect the construction of electric power communication network is very urgent.

For carrying out the future development of power grid planning and the future development of the communication network planning to China State Grid Corp, to meet the needs of smart grid construction, the ability of supporting construction should be focused and improved, the backbone communication network will be extended to the 66 kV side, and the construction of communication network should be further strengthened. As an important part of communication network, the electric power communication network, need to change the traditional mode of development, accelerating the pace of construction and upgrading, efforts to solve the influence and restriction of power automation communication

*Corresponding author e-mail: chulili902@163.com

problems, and the intelligent application of management information system, and provide important technical support to improve the power grid safe and economic operation level and the management level of enterprises.

Because the characteristics of the network is flexible, chain network, tree network, star network, ring network, hand in hand net basic network, and the combination of network by any of these basic network structure can be formed. The use of passive optical network can be based entirely on a grid structure for low-voltage networking communications network, and once with the low-voltage grid structure adjustment to achieve rapid adjustment of the communication network. Passive optical network having a multi-node failure characteristic, the passive optical network communication terminal of any one failure will not affect the normal operation of other communication terminals. It is much in line with large communication terminal distribution automation, information collection requirements on the reliability of communication. Passive optical network technology in China Telecom has been running for many years and there are many manufacturers to provide equipment and technical support. Passive Optical Network in Power Line Communication as a relatively new technology in the power system pilot cities have been widely used, and achieved good results.

2 Access network planning and research

Smart grid access network communication network construction is the construction of distribution network, considering the PON access technology has been widely used in the operator access networks, the smart grid access network side (i.e. distribution communication network side) of PON access will be focused research.

In the distribution network, the communication backbone layer distribution master station to the substation used SDH backbone optical transmission network which has been built. In access layer, substation to ring network cabinet, column switch and other automated information collection site selected access scheme in EPON optical fibre network technology. EPON communication is based on the network structure of hand in hand, and it support dual OLT uplink mode and support for 1 + 1 and 1:1 protection way.

1) OLT equipment is installed in the substation, interconnected existing SDH transmission devices via GE or FE, completing the distribution information to the master transfers. OLT support EPON/GPON platform, support 10G GE, STM-1, E1 and other upstream, to facilitate future expansion application and cost savings. It also supports dual master control, dual power supply and the uplink port redundancy backup, to ensure the system security.

2) The ONU equipment is installed in the information collection points to achieve related equipment information uploaded to the substation. ONU devices use double PON port equipment to achieve full protection

self-healing, using industrial grade equipment, case closed, aluminium material. It's surface without holes to prevent water leakage accident. It support environment temperature -40 degrees to +85 degrees to meet the operating environment more hostile scene. It also support AC, DC power supply, DC power supply support adaptive 9V~60V. Convenient installation and maintenance, saving the spare parts and saving investment are the main features of ONU equipment.

3) The logical table of telecommunication system in electric is shown in Fig. 1.

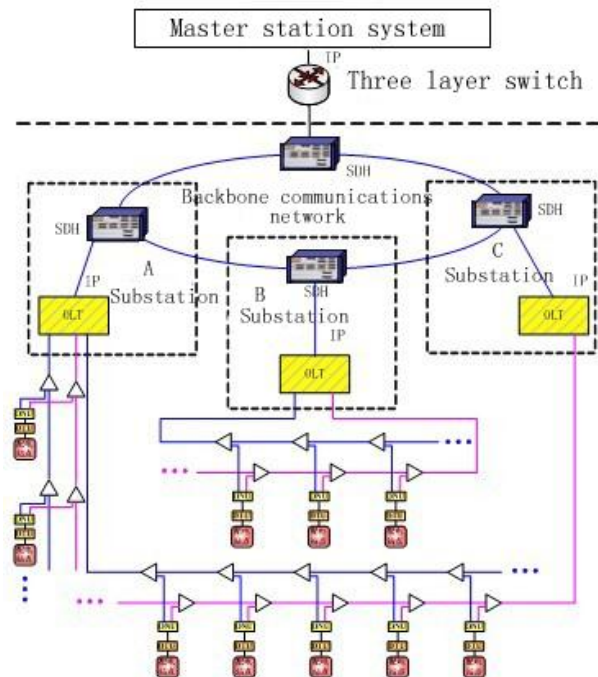


FIGURE 1 Logical table of telecommunication system in electric

Now combined distribution network characteristics and advantages of PON network, EPON network deployment recommendations are summarized as follows:

Suggestions for the development of OLT

1) OLT location and coverage area

According to coverage, select the appropriate room. Urban areas of OLT coverage should be controlled within 5 km, rural area can be properly extended to 10 km.

2) OLT configuration recommendations

In the end, the installation of OLT equipment should be used in large capacity and OLT rack mounted equipment. The OLT equipment deployed in the module Bureau and the access point should be based on the final number of users to select the appropriate capacity equipment and should be appropriate use fixed port OLT devices with small capacity. OLT equipment should have at least two GE port, which with the increase of users. We should increase the port GE as the user speed gradually, and ensure OLT port has at least one backup GE.

Suggestions for the development of ODN [4]

1) How to set the beam splitter position

City area splitter position should mainly adopt the centralized placement, set in the area of optical node,

such as small room, small roadside or corridor in order to improve the trunk fibre utilization rate and try to avoid placing OLT and the splitter in the same end office room. Rural area splitter position can be flexibly placed in the township end office, cable transfer box or optical cable connector box etc. according to the actual situation.

2) How to choose the beam splitter

POS (Passive Optical Splitter) optical branching ratio should be used 1:2, 1:4, 1:8, 1:16, 1:32 the five common types according to the selection, the number of broadband users of one PON port should not be more than 512. The 2: N optical splitter can be used in the case there is a need of protection. The PLC (Planar Light wave Circuit) splitter should be the main choice in current. In the low branching ratio (1:2, 1:4) situation based on the cost we may be appropriate to consider the use of FBT (Fused Biconical Tap) type optical splitter.

The total splitting ratio of ODN should be determined according to the users' requirements, the optical link bandwidth attenuation factors and other factors. Single-stage method should be mainly used in the urban area and the rural areas. Multistage method can be used based on the multistage light cable resources and village distribution, but it should be controlled in three stages.

Suggestions for the development of ONU

In the network construction, enterprises should select different types of ONU according to the users' needs, but they should control total types, avoiding excessive ONU models in the access network. Products, which can replace the port of PON module, should be select. The ratio of broadband and narrowband users should be considered, so enterprises should try to choose products that can be flexibly adjusted ratio. For the ONU supported the type of DSL, consideration should be given to the board and the general DSLAM network equipment in order to use network to adjust the DSL card.

Suggestions for the development of network management system

A unified PON network management system should be established while in the PON network. The system should support for OLT and ONU configuration, fault, performance, security and other management functions, and requires an ONU that can support automatic discovery. Through the development of connecting single system, service system and accounting system interface, it can gradually complete the job automatically accepted, business automatic opening function. The PON management should be incorporated into the integrated network management system that already has broadband data integrated network management system.

3 The construction of the access network

Access network construction combined with SDH transmission ring network can provide network load for the automation system of smart grid. The sub regional centre and the 66 kV substation have been built the SDH loop network, so the OLT can through the Ethernet port

access SDH transmission equipment in substation. When connected to the FE electrical interface in the same SDH, the OLT can automatic or manually turn its rate to one Mbps.

The network of the telecommunication backbone is shown in Fig. 2.

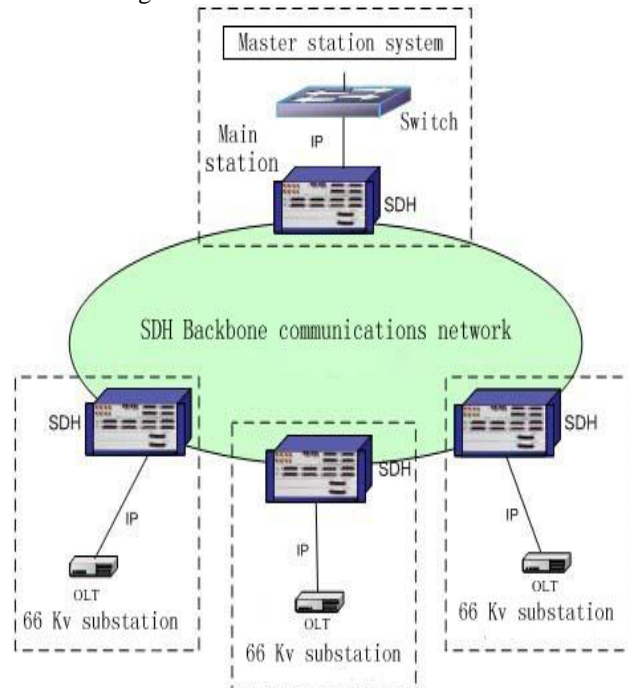


FIGURE 2 The network of telecommunication backbone

A three layer of Ethernet switches was placed in the centre of the network to gather all the data of distribution automation from each substation. Taking into account the future network communication system with numerous communication terminal equipment carried by each substation, isolation and stratification can be achieved through the three layer routing.

Considering the redundancy and other factors, each 10 kV line needs 24 core cables, 16 cores used to transmit distribution information, 8 cores used to transmit electricity information. Fibre optic cable is interrupted and fused in a set of light distribution which installed on each ring network cabinet (or switch on the pillars), and then connect it to the ONU equipment of the distribution station through a splitter pigtail.

All information collection points involve in the project adopt the fibre EPON communications system and use the ONU with double PON ports as information collection terminals. Routings use the "hand in hand" type of protection way combined with the actual trend of distribution lines in the selected area.

The hand-in-hand two-point structure of EPON as shown in Fig. 3, OLT1 and OLT2 were installed in the different substation and the ONU equipment were installed in each information collection point. When the cable break or individual OLT equipment fail to work, the ONU equipment can select different access point to connect the OLT in order to protect the system.

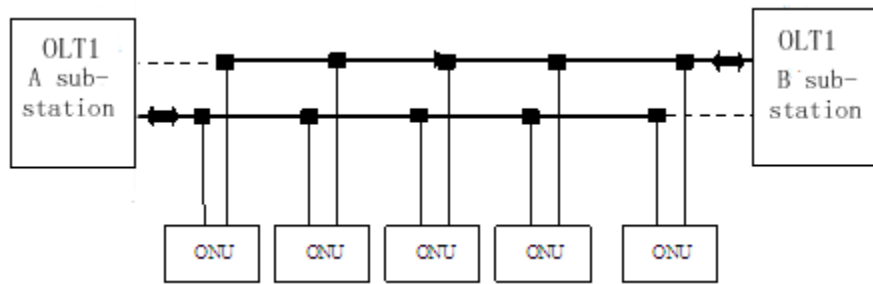


FIGURE 3 Hand-in-hand two-point structure of EPON

According to the OLT to the ONU channel attenuation's range and the experience of the ONU configuration, combined with the actual situation of City Bureau distribution network, and give full consideration to the future network expansion, the transformation and upgrade, the early network should be set aside enough light power capacity. For wiring way hand-in-hand, the number of ONU carried by PON ports of the OLT should be not more than 12 in the early planning, the highest level of optical-splitting is eight stage optical-splitting in the backbone. When the number of nodes exceeds the number that the ONU equipment and the splitters can control, enterprises can take apart the network and add

new light core to the network in order to reduce the beam splitter series and the channel attenuation. In the design phase of scale, enterprises should take the above design as the basis and combine with the distribution network connection status and the long-term planning of the specific situation on the channel attenuation to analyse and calculate the channel attenuation in order to determine the ONU of each line and the actual configuration of the beam splitter.

The following is a concrete example, a network structure corresponding to a hand-in-hand line. The architecture of hand-in-hand joint-access is as shown in Fig. 4.

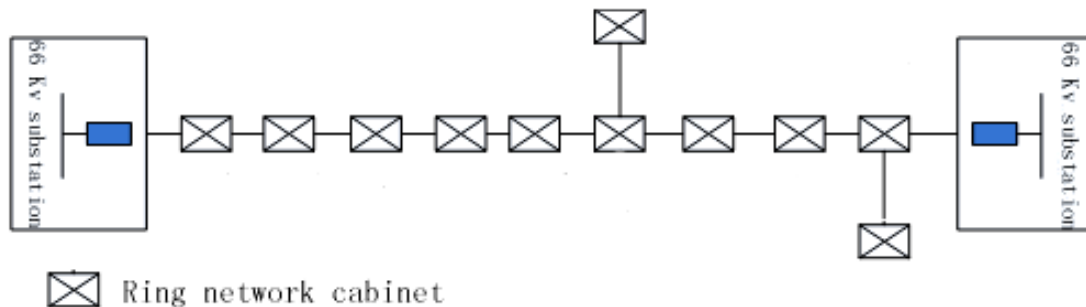


FIGURE 4 The architecture of hand-in-hand joint-access

a. Line's ends were ended in two substations OLT device based on the construction of distribution automation project line information point within the region distribution, respectively, which is similar to the above line grid structure with hand-in-hand network structure.

b. The network topology may change in the future, considering the expansion, transformation and upgrading of network, so enterprises need to reserve space for expansion. Each splitter should reserve at least one expansion port, if the recent expansion of the branch is

large, enterprises can use the splitter with large splitting ratio splitter to get more expansion ports.

c. All lines in the network use hand-in-hand type structure, trunk lines use the beam splitter whose spectral ratio is 10:90 to make the network realize five stage optical-splitting modes or six stage optical-splitting modes. ONU access different PON ports of OLT in the different substation point, respectively. The typical circuit occupies two PON ports of each site, four PON ports in the network, and occupies four core optical fibres of the trunk line. The hand-in-hand communication structure is shown in Fig. 5.

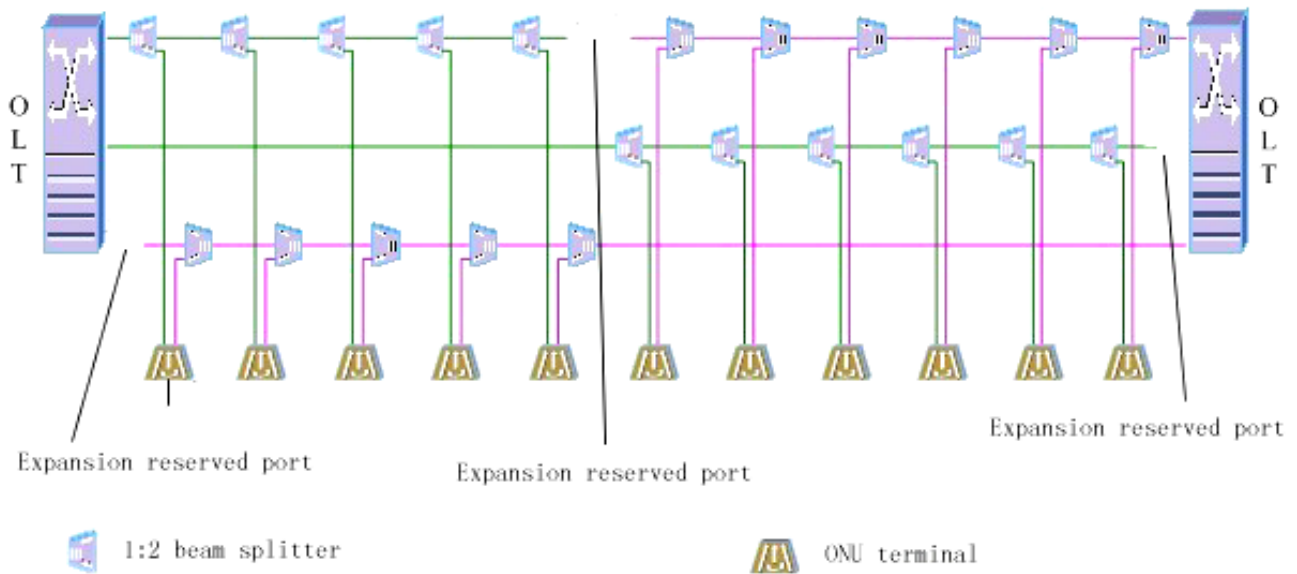


FIGURE 5 The hand-in-hand telecommunication

4 Energy efficiency analysis of the network

4.1 THE INDICATORS OF ENERGY EFFICIENCY EVALUATION

In order to evaluate the energy efficiency performance of the network, this section will introduce a series of performance indicators. They are mainly used to illustrate the superiority of the proposed algorithm in the simulation analysis process.

4.1.1 The link utilization [5]

First, the author defines the link utilization. In this paper, $\overline{u_{ij}}$ indicates the link utilization. The link utilization L_{ij} is the ratio of average flow on the link and the link capacity, the formula is shown as follows:

$$\overline{u_{ij}} = \frac{\overline{f_{ij}}}{C_{ij}}, \tag{1}$$

The link utilization is mainly used to measure the status of the link utilization, for example, if the $\overline{u_{ij}}$ of the link L_{ij} is more than 80%, then the link rate is higher, and the network's risk that other services through this link brings is bigger, which is more easily lead to link congestion and denial of service condition. We should avoid the link utilization rate is too high. However, if the network utilization rate is too low, then the network energy efficiency would decline. So in order to achieve energy efficiency network, the link utilization should be controlled in a safe range by using a reasonable mechanism.

4.1.2 The link utilization variation

U is the matrix of the network link utilization variation within the time T after several users' request, $\overline{u_{0ij}}$ and $\overline{u_{ij}}$ indicates the initial link utilization rate and the request link utilization, respectively. The formula is expressed as:

$$U_{ij} = \overline{u_{ij}} - \overline{u_{0ij}}, L_{ij} \in E. \tag{2}$$

Link utilization variation matrix is used to measure the link utilization of the entire network and to constraint the performance indicators of the network. It is mainly used in algorithms (1) and (3).

4.1.3 The average node degree

Due to network node degree is inherent in a network, the average node degree of network is large, the connectivity of the network is strong, and the network can adapt to the negative load state of the business. So the average node degree will also affect the results of the algorithm simulation, the average node degree is defined as follows:

$$D = \frac{\sum_{i \in V} d_i}{\|V\|}, \tag{3}$$

d_i indicates the degrees of node i , $\|V\|$ denotes the number of nodes in the network. The indicator is used in the simulation analysis and mainly embodies that average node degree of different network has different effects on algorithm performance.

4.1.4 The average activation link utilization

The average activation link utilization is similar to the link utilization. This definition is specially used for the network of energy efficiency, because only in the network of energy efficiency it can sleep for the redundant links. In this paper, the average activation link utilization is the ratio of the sum of the average utilization of link utilization and the number of the activation link. The formula is shown as follows:

$$U_l = \frac{\sum_{l \in E} u_l}{\|E\| - S_n}, \tag{4}$$

u_l is the link utilization of l , $\|E\|$ indicates the total number of links in the network, S_n indicates the dormancy link number. The average activation link utilization is used to measure network performance after dormancy.

4.1.5 The denial service rate

The denial service rate is the ratio of the number of refused service and the number of the total requests in network traffic when the network link residual bandwidth is shortage in the case of the network routing is successful. In this paper, the author use the BUG to indicates the denial service rate. The formula is shown as follows:

$$Br = \frac{Nr}{NR}, \tag{5}$$

Br indicates the denial service rate, Nr indicates the number of refused service and NR indicates the number of the total requests in network traffic.

This article defines the above indicators used to measure the performance of network and used in the simulation analysis, deeply describes the network of the energy efficiency performance.

4.2 RESULTS AND ANALYSIS OF THE SIMULATION

The network energy efficiency of A, B, C and D with ring network mode were analysed by using MATLAB in this paper. For dormancy algorithm, we use the EAR algorithm, the traditional dormancy algorithm to compare the energy efficiency of four kinds of networking solutions and network characteristic properties under the different routing algorithms.

The simulation results are shown in Fig. 6, the cumulative number of denial of service distribution in 10 independent times is depicted in the figure. We can find that the two routing algorithms (OSPF and MHA) have

little effect on the denial service times under the same network mode. Only in A ring network the routing algorithms have effect on the denial service times.

The simulation results are shown in Fig. 7. The figure shows cumulative distribution of the denial service rate. The x axis represents the different link, the y axis represents the average link usage calculated in the 10 times of the simulation experiments, the curves of different colour in the figure represent the changes of each link usage obtaining by using different routing algorithms. The horizontal line presents the average link usage of the network; we can see the routing algorithms have no effect on link usage from the simulation diagram. In addition, the link order is A, B, C and D on the basis of the average usage from large to small.

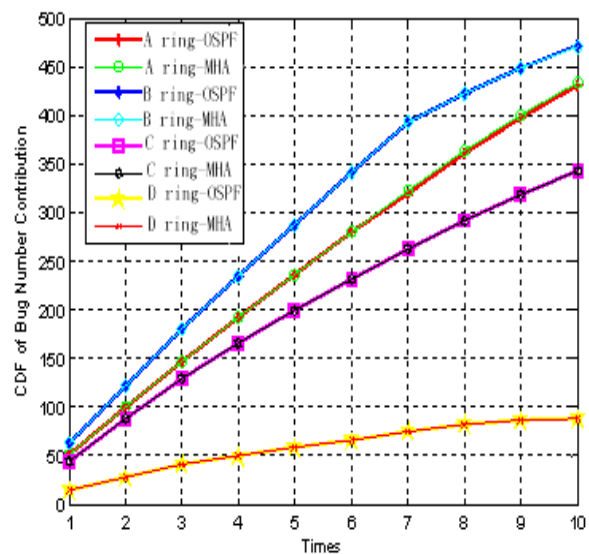


FIGURE 6 The accumulation distribution figure of devotion rate of refusal to service in different route algorithm

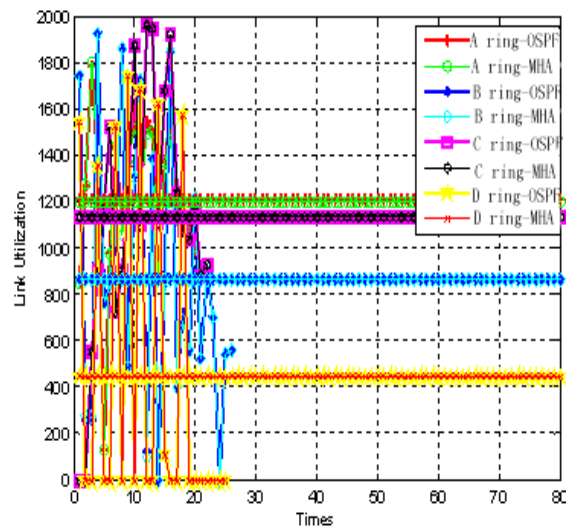


FIGURE 7 The consumption in average link with different route strategies

The values of network power-efficiency with different emulation labs as shown in Figure 8. It can be seen that the difference between OSPF algorithm and MHA

algorithm on A ring network energy efficiency is bigger, but the two algorithms have little influence on other network efficiency.

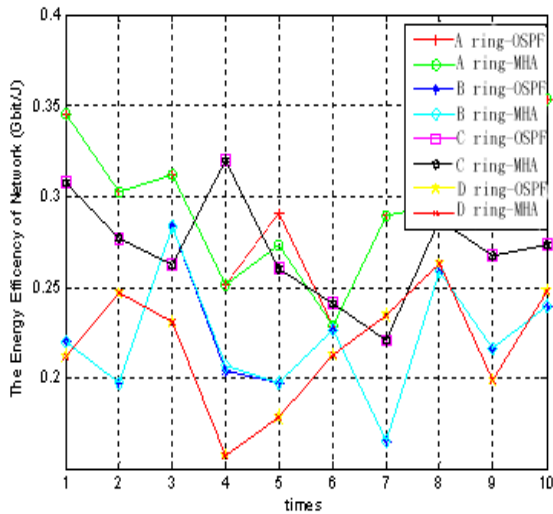


FIGURE 8 The value of network power-efficiency with different emulation labs

5 Conclusions

For example, we can carry out the actual communication network construction in a northern city, including the detailed content of construction and the equipment configuration of cable, transmission, access network and data network. Combining with the classic network performance analysis, we should construct the energy efficiency analysis model, carry out the simulation on the basis of the energy efficiency index set for the simulation, analyse the experimental results deeply and complete network performance optimization for deployment through regional pilot in the process of designing network operations in order to support the development of electric power communication network better.

References



[1] Li T Z 2012 Flexible interactive technical connotation and development trend of intelligent electricity *Automat. Electr. Pow. Syst.* **2** 11-7 (in Chinese)

[2] Liu G T, Sun Z Q 2013 Intelligent distribution network communications technology development review *J. Internet Technol.* **1** 49-53(in Chinese)

[3] Cao Y, Yao J G 2011 The latest progress of smart grid core standard IEC61970 *Automat. Electr. Pow. Syst.* **35**(17) 1-4 (in Chinese)

[4] Liu W, Yang H X, Zhu B 2012 Review of research on smart grid technical standard system, *Pow. Syst. Protect. Control* **40**(10) 120-6 (in Chinese)

[5] Yao J G, Yang S C 2012 Smart grid “net charge” interactive operation control concept and research framework *Automat. Electr.Pow. Syst.* **21** 1-6 (in Chinese)

Authors	
	<p>Lili Chu, born in November 30, 1970, Jinzhou, Liaoning, China</p> <p>Current positions, grades: Liaoning University of Technology, Professor</p> <p>University study: Liaoning University of Technology (Faculty of Application of electronic technology), 1994; Master’s degree in communication and information system. Liaoning University of Technology,2000; Ph.D. degree in Liaoning Technical University, 2012</p> <p>Research activities: Modern communication network technology, Computer Network, Signal & System, Coding Theory</p> <p>Publications: 20</p> <p>Experience: Chaired or participated in the National Natural Science Foundation and Natural Science Foundation of Liaoning</p>
	<p>Lili Qin</p> <p>Current positions, grades: State Grid Liaoyang Power Co., Ltd, assistant engineer, at present, studies at Liaoning Industrial University, Master of Electronics and Communications Engineering, will graduate in July 2014.</p> <p>University study: Bachelor’s degree in Shenyang Institute of Engineering, Power Systems and Automation, 2009.</p> <p>Experience: Participated in construction engineering of Liaoyang Gexi River 66 kV substation, Liaoyang Dongsheng 66 kV substation, Liaoyang Hedong 66 kV substation and so on, (2009 - 2014)</p>

A distributed multicast routing algorithm based on bone node set for mobile IP

Ling Zhou*, De Feng Zhang

Department of Computer Science and Technology, Foshan University, Foshan 528000, Guangdong, China

Received 1 March 2014, www.tsi.lv

Abstract

Multicast routing is an important issue in network communication. In order to optimize the multicast routing cost and lessen the transmission delay for mobile IP communication, an idea of bone node set is introduced and the distributed multicast routing algorithm is designed based on the idea firstly. At the same time, the algorithm is implemented according to centre version and distributed version in detail, respectively. Then its necessary data structures, time complexity and message complexity are analysed in theories according to order of sequence for distributed operation. At last, simulation experiments are done in a 7×7 mesh topology and the results show that the designed algorithm can optimize the routing cost for multicast routing and reduce the transmission delay greatly compared to some same type algorithms. The distributed routing algorithm with the simple complexity can be efficiently used in large-scale mobile IP network.

Keywords: Distributed routing, mobile IP, bone node set, performance analysis, simulation

1 Introduction

Multicast is a one-to-many (or many-to-many) data communication method. By multicast communication the sender only needs to send a packet and all the destinations will receive the packet with only one copy of the packet in each link. The key factor to realize multicast is to design a reasonable multicast routing algorithm to construct the data forward tree. Routing cost, time delay and scalability are the three most important parameters to evaluate the multicast routing algorithm. Domestic and foreign scholars have raised some excellent multicast routing algorithm based on traditional IP network and the fixed topology network structure. Recently with the occurrence of mobile nodes (MN), there are some serious challenges to those traditional multicast routing algorithms, which are no longer fitted, to the new mobile Internet environment. Therefore, how to design a multicast routing algorithm for mobile IP (MIP) is becoming a more and more important problem in the current mobile IP research area [1].

IETF has proposed two basic methods to solve the application problem of multicast in mobile IP, that is, the bi-directional tunnel (BT) and remote subscription (RS) [2]. And some scholars have also put forward several improved algorithm according to BT and RS.

With BT algorithm, when the mobile node moves to a foreign subnet, it will acquire a care of address (CoA) from the visiting foreign subnet and register the CoA to its home agent (HA), after that the MN will forward data packet by HA through bi-directional tunnel. Therefore, by BT mode, the MN's movement and handover are

transparent to all the other network nodes and it is fit to the current mobile IP architecture. However, BT introduces a tri-angle routing problem and so its routing path is not the most optimization one. In addition, there exists the problem of a longer joint delay due to the long tunnel path, and the focus problem because of different home agents, etc.

With RS algorithm, when a mobile node moves to the visiting foreign subnet, it will apply to join the multicast group by access router (AR) in the foreign subnet and directly receives data packet from AR in foreign subnet. Therefore, in RS method its routing path is the most optimization one. However, RS also introduces some problems. For examples, RS leads to a frequent path reconstruction for multicast routing, and brings about a serious packet loss rate due to high delay in handover.

To reach a compromise between routing path optimization and reconstruction frequency, a range-based mobile multicast (RBMoM) routing algorithm is proposed by Lin C R and Wang K M [3]. The main contribution in RBMoM is that the multicast home agent (MHA) is introduced into mobile IP architecture and the concept of service range is put forward. RBMoM can optimize routing path by the choice of MHA and reduce the frequency of reconstruction by changing the service range. The main disadvantages of RBMoM are the uncertainty of service range and the extra election of MHA.

In order to optimize the handover delay and reduce the mobile IP multicast packet loss rate, a multicast routing algorithm based on mobility support agent (MSA) is proposed by Wu J. MSA is a router in the MN's

*Corresponding author e-mail: cszhouling@sohu.com

foreign subnet [4], and it is responsible for pre-registration in a visiting foreign subnet when the handover happens. Therefore, MSA can optimize the join delay when handover occurs and reduce packet loss rate. The key work for the algorithm is to predict what time the handover will happen and which direction the MN will move according to.

By trying to solve the problem of packet loss between handover, a multicast routing algorithm called multicast by multicast agent (MMA) is also put forward by Suh Y [5]. The main idea in MMA is to introduce a multicast forwarder (MF) entity, which is a data packet forward router for multicast. When handover happens, MF will buffer the communication data for multicast and later forwards the data to the new subnet. By the collaboration of MF, the packet loss ratio can greatly reduce between the old subnet and the new subnet when the handover occurs.

Handover brings many problems in mobile IP. Beside above strategies to optimize the handover, Tan C L develops an algorithm called MobiCast [6] to handle with it, which is mainly based on the idea of hierarchy, that is, the network is divided into macro layer and micro layer. A macro layer may include several micro layers, and each micro introduces a domain foreign agent (DFA) entity to manage some subnets in the same micro layer. When the MN moves between different subnets (belong to the same micro layer), MobiCast need not to reconstruct the multicast tree and these problems coming from handover can be avoided. When moving between macro layers, DFA is responsible for MN to join the multicast tree, and to reconstruct multicast routing tree.

Above several solutions have been put forward to solve the mobile multicast routing problem by some scholars. Nevertheless, most of all change the architecture of mobile IP by introducing a new network entity to optimize the handover, and few optimizes the network cost and reduces the handover delay according to itself characteristics of multicast [7]. At present, there are still the next important problems to be solved in mobile IP for multicast routing communication. 1) The construction of low cost multicast routing tree. The tree cost computing by the multicast routing algorithm is not the most optimized one and so consumes too much bandwidth to forward data packet; especially in mobile wireless environment, it is difficult to apply. 2) The seamless handover problem. Multicast routing algorithms introduce a relatively large handover delay, which seriously affects the seamless multicast session. 3) The packet loss problem. Due to the high bit error rate of wireless link, the packet loss rate is bigger when MN (or mobile subnet) moves fast; and so a lower multicast session rate will occur. 4) The QoS problem. These algorithms do not consider QoS constraints, for example, delay, delay jitter. 5) The distributed policy. These belong to the centralized routing algorithm, and rarely consider the distributed implementation.

The idea of bone node set is introduced in the paper, and a distributed routing algorithm based on bone node set is designed for mobile IPv6, which is called the bone node set-based multicast routing algorithm (BNSBMR). By BNSBMR algorithm, we can optimize the tree cost for bandwidth management, reduce the handover delay and lessen the packet loss rate. Distributed algorithm is more fitted to employ in large scale of network relative to centralized algorithm and is easier to realize the scalability. Especially in mobile IP environment, the distributed strategy has a unique advantage. The operation of the routing process only relies on the MN, without all the other nodes involved, and which will simplify the operation of mobile IP, optimize the handover, and improve network management performance.

The remainders of this paper are organized by the following way. In the second section, the concept of bone node set is put forward. In the third section, the multicast routing algorithm BNSBMR for mobile IP is described, and the distributed version is implemented in section 4. The distributed data structure, message complexity, and time complexity are analysed in section 5. In section 6 simulation experiments are done to verify the algorithm performance. Finally, in the seventh section we summary the paper and draw some conclusions.

2 Basic conception and model for MIP

A communication network can be modelled by an weighted graph $G=(V, E, W)$, where V is a set of host or router nodes, E is the set of communication links and W is the weight parameter belonged to a specific link. To any links $e \in E$, we can define the cost function as $Cost(e): E \rightarrow R^+$ and the delay function as $Delay(e): E \rightarrow R^+$. Given source node s , destination node set D , the destination node number is $m=|D|$, and the network node number is $n=|V|$.

Definition 1 (path)

Given $G(V, E, W)$, if there exists a node sequence $(s, v_1, v_2, \dots, v_n, t)$, such that $(s, v_1), (v_1, v_2), \dots, (v_n, t) \in E$, so we call those edges as a path, and write the path from s to t as $P(s, t)$.

Definition 2 (Least cost path)

We call the path from u to v a least cost path if the total weighed cost from u to v is the least one and we write the least cost path $p_{lc}(u, v)$.

Definition 3 (Least cost tree)

Given network $G(V, E, W)$, the source node s and a destination set D , if a tree T spans $s \cup D$ and its total cost satisfies the equation:

$$Cost(T) = \min\{Cost(T) = \sum_{v \in D} \sum_{e \in P(s,v)} Cost(e)\}$$

$$(\forall v \in D, P(s,v) \in T),$$

we call the tree T a least cost tree for $s \cup D$.

Definition 4 (Bone node set)

To mobile IP, the bone node set is a dynamic set of AR (access router) in some selected subnets and written

as N_b . The elements in N_b include these subnet's AR where existing static multicast node and these subnet's AR where existing mobile multicast node i with the continuous stay time T_i meeting the following condition, $\forall i \in D, T_i > \Delta$, where i denotes the mobile node i , and Δ is the time boundary. Using $Router(i)$ to denote the AR of subnet i , the bone node set N_b can be expressed as follow,

$$N_b = \{Router(i) | i \text{ includes a static member node}\} \cup \{Router(i) | i \text{ includes a mobile member node and } T_i > \Delta\}.$$

In MIPv6 architecture, two entities are mainly introduced: mobile node MN and home agent HA [8]. The MIP multicast model can be described as Fig. 1. Server source is the multicast source. Subnet1 is the home subnet of mobile node MN1, and HA is its home agent. MN1 is visiting subnet2 currently, and F is its AR. When MN1 moves from subnet2 to subnet3, handover happens. As MN1 moves into subnet3, it will obtain a care of address (CoA) from router G (that is, subnet3's AR) so that it can re-join multicast tree in subnet3 or accept continuously multicast data from its HA. It can be drawn from the model that the multicast routing problem for mobile IP mainly includes two aspects. The first is how to construct the optimization multicast routing tree. The second is how to deal with handover. The paper will mainly solve the first problem based on bone node set, and also consider handover in some degree.

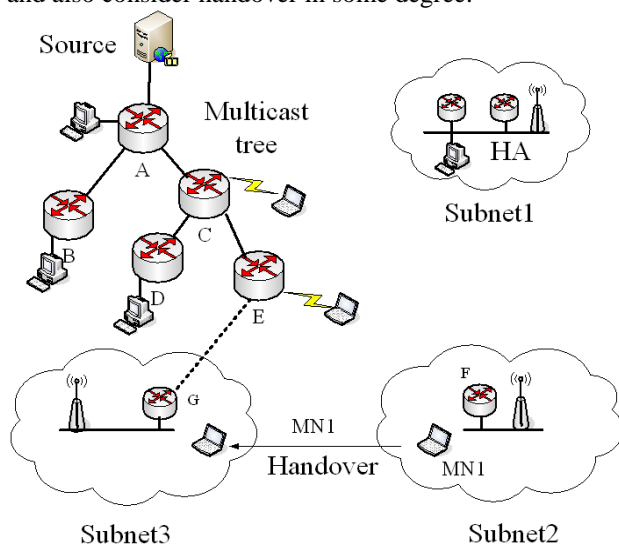


FIGURE 1 Multicast model for mobile IP.

3 Algorithm description

3.1 BASIC IDEA

The basic idea of BNSBMR algorithm is as follows. When source s sponsors a multicast session, BNSBMR starts to compute the multicast tree T for all destination nodes $\forall i \in D$ and the bone node set N_b , preparatively. The multicast routing tree T is used for forwarding data packet, and bone node set N_b is the basis and foundation

to optimize the routing path and reconstruct the multicast tree, when mobile node moves from a subnet into other subnet.

When handover occurs, BNSBMR algorithm will search the bone node set N_b and find a node $j (j \in N_b)$ which makes the $Cost(i, j)$ is the least cost from i to bone node set N_b . By bone node set, the new algorithm not only can reduce the multicast tree cost and optimize routing path but also can find a shortest path and lower the join delay.

At the same time, bone node set is not a stationary sharing tree. To mobile node i , If $T_i > \Delta$ is meet, the maintenance process for bone node set will be triggered and change the element of N_b . Therefore, the set of N_b is valid and superior.

Bone node set is the core idea of the algorithm. Therefore, how to maintain the bone nodes set is also a key technology of the algorithm. We can illustrate the generating and maintenance process of N_b combined with Figure 1. Supposed there exist seven subnets A, B, C, D, E, F, and G, and accordingly their ARs are also denoted as A, B, C, D, E, F, and G. Simply, supposed there is one and only one multicast node in each subnet. Among them, the multicast nodes in A, B, D are static and the multicast nodes in C, E, F are mobile nodes. 1) When preparatively computing, because there are static/fixed multicast nodes in subnet A, B, D, so $N_b = \{A, B, D\}$. 2) When MN1 moves from subnet F to subnet G, MN1 scan N_b and may select AR D to re-join the multicast tree. 3) After a period of time Δ , if the mobile node doesn't move out subnet C, E, so access router C, E can be regarded as bone node and added to set N_b . that is, $N_b = \{A, B, C, D, E\}$. If at this time MN1 moves from subnet F to subnet G, MN1 scan the set N_b and may select AR E to rejoin the multicast tree. 4) Supposed that MN1 moves into G and stay in G for a period of time $T > \Delta$ and the mobile node in subnet C moves out C, AR G will become an element of N_b and AR C will be delete from N_b , so $N_b = \{A, B, G, D, E\}$.

The specific algorithm and pseudo code are analysed in the following sub-section.

3.2 PREPARATION COMPUTING

At the beginning of the multicast session, source node s scans all multicast destination nodes, generates bone node set N_b and initializes routing tree T . The procedure is described as shown below.

Input: G, D, s

Output: T, N_b

Pre_Cal_Tree(G, D, s)

1. $N_b \leftarrow s, T \leftarrow s$
2. $Q \leftarrow D // Q$ is a list table to record the destination node set D
3. While (Q is not Null) Do
4. Select the destination i and a node $j \in N_b$ so that $P_{ic}(i, j)$ is minimum among $P_{ic}(i, \forall j \in N_b)$

5. **If** i is a static node **then**
6. Mark $Router(i)$ as a bone node
7. $N_b \leftarrow N_b \cup i$
8. $T \leftarrow T \cup P_{lc}(i, j)$
9. $Q \leftarrow Q - i$
10. **End if**
11. **If** i is a mobile node **then**
12. $T \leftarrow T \cup Path(i, j)$
13. $Q \leftarrow Q - i$
14. **End if**
15. **End while**
16. **Return** T, N_b

3.3 PROCESS OF MAINTENANCE FOR BONE NODE SET

Below we consider the process of bone node set maintenance.

On one hand, if a mobile node i stays in one subnet for a period of time $T_i > \Delta$, we can regard it as a stable node and take this subnet's AR as an element of N_b . On the contrary, it isn't regarded as an element of N_b . There is a timer in each mobile node, when handover occurs the timer begins to count the time. If $T_i > \Delta$, the below process Add() will be triggered, and this subnet's AR is labelled as an element of N_b .

Add()

1. **If** $T_i > \Delta$ **then**
2. Mark $Router(i)$ as a node of bone tree
3. $N_b \leftarrow N_b \cup Router(i)$
4. **End if**

On the other hand, the below process Delete() will be run when mobile node i moves out its visiting subnet.

Delete()

1. **If** there doesn't exist other multicast node connecting to $Router(i)$ **then**
2. Select a parent node for its son nodes
3. Mark $Router(i)$ as a non-bone node
4. $N_b \leftarrow N_b - Router(i)$
5. **End if**
6. **If** there exists a loop **then**
7. Remove the loop by changing the node's parent
8. **End if**

3.4 PROCESS OF ROUTING FOR MOBILE NODE

When mobile node i switches to a new subnet from the current subnet, it firstly inquires whether there is multicast members in the new subnet. If there exists, it receives multicast information directly from new subnet's AR and do nothing in reconstructing multicast routing tree. Vice versa, if no multicast members, it starts to run the below Cal_Tree process, and scans bone node set N_b to find a node $j \in N_b$, which has the shortest path to itself. After that it grafts into the multicast tree by node j and receives multicast data from node j .

Input: T, N_b, i

Output: T

Cal_Tree(T, N_b, i)

1. **For** each node in N_b **do**
2. Find node $j \in N_b$ which has the shortest path to i
3. $T \leftarrow T \cup P_{lc}(i, j)$
4. If there exists a loop then
5. Remove the loop by changing the parent node
6. **End if**
7. **Return** T .

It is clear that the new algorithm optimizes the routing path originating from RS method and does not rely on mobile nodes HA. Compared with BT, there are smaller join delay and information transmission delay in the new algorithm. Because BNSBMR optimizes the routing tree cost by bone node set, so compared to other RS method algorithms it has better cost performance and transmission delay.

4 Distributed implementation

According to the scalability and mobility, distributed methods are more suitable for implementing the routing algorithm. In this section BNSBMR algorithm is realized in distributed way. The algorithm idea and the variable symbols are the same as the previous centralized description. Distributed policies mainly include two parts: the distributed maintenance for bone node to renew N_b and distributed routing for mobile node to join the multicast tree.

4.1 TOPOLOGY INFORMATION

Supposed that each node i (including MN and router) stores a vector table $Routable(j)$ and uses the table to record routing information from node i to node j . In $Routable(j)$ there are $|V|-1$ entries and each entry is corresponding to one node j , including the node index j , the shortest path $P_{lc}(i, j)$ and its cost value $Cost(P_{lc}(i, j))$, the next hop node nh . Generally, this information can be drawn from the network topology graph.

4.2 ROUTING MESSAGES

The below three kinds of messages is essential to implement the distributed method for the algorithm: Message_BNS_modify(), Message_join(), and Message_setup().

Message_BNS_modify() is used to transmit message among element of N_b to maintain the bone node set, and its parameters include type (Add or Delete), the index of bone node i that is generating the message, the index of multicast session ID .

Message_join() is originated by the mobile node to select a bone node to join the multicast tree when handover occurs. Its parameters include the index of mobile node i , the selected bone node j , multicast session index ID .

Message_setup() is used to establish a path hop by hop between MN i and the selected bone node j with the

path $P(i, j)$ being the least cost path $P_{lc}(i, j)$. Its parameters include multicast session index ID , node k which is on the path from i to j , and k 's next hop $next_hop$.

4.3 MAINTAINING BONE NODE SET

Each network node $i \in N_b$ runs the following two processes to add into or exit from bone node set N_b .

Add()

1. Mark $Router(i)$ as a bone node
2. Add the router i to bone node list: $N_b \leftarrow N_b \cup i$

Delete()

1. Mark $Router(i)$ as a non-bone node
2. Delete the router i from bone node list: $N_b \leftarrow N_b - i$

4.4 SELECTING ROUTING PATH

Each member node i runs the following process to select a bone node to join multicast tree.

Join()

1. Scan the bone node list to find node $j \in N_b$ which has a shortest path to i
2. $next_hop = Routable(j).nh$
3. $Router(i)$ joins the $path(i, j)$
4. Return $j, next_hop$

Each network node i runs the following process to choose the next-hop route and establish routing path hop by hop.

Setup()

1. $Router(i)$ joins the $path(i, j)$
2. $next_hop = Routable(j).nh$
3. If there exists a loop then
4. Remove the loop by changing the parent node
5. End if
6. **Return** $next_hop$

4.5 DISTRIBUTED ROUTING PROCEDURES

Step 1. Preparation computing. Running preparation computing process in source node s and obtaining the original multicast tree T and bone node set N_b .

Step 2. Adding an element into N_b . To an In-tree node, if its $T_i > \Delta$, it runs the Add() process to mark itself as a bone node, initializes the message $Message_BNS_modify(add, i, ID)$ and sends the message to each node in bone node list. For each bone node which receives $Message_BNS_modify(add, i, ID)$ starts its own Add() process to deal with the message.

Step 3. Deleting an element from N_b . If a node belongs to N_b , it runs the Delete() process to delete itself from the bone node list, initializes the message $Message_BNS_modify(delete, i, ID)$, and sends the message to each node in bone node set. For each bone node which receives the message starts its own Delete() process to deal with the message.

Step 4. Starting to establish routing path. When a MN i switches to a new subnet and handover occurs, it runs

the Join() process to find a node $j \in N_b$ which has the shortest path to i , generates the message $Message_join(i, j, ID)$, scans the $Routable()$ to find the next hop nh in $Routable(j)$ and sent the message $Message_join(i, j, ID)$ to nh .

Step 5. Continuing to establish routing path. If one node k receives the message $Message_join(i, j, ID)$, it runs the Setup() process, generates the message $Message_setup(k, j, ID)$, and sent the message to next hop nh in $Routable(j)$. Looping procedure 5, the routing path can be set up hop by hop.

Step 6. Finishing the setup process. When bone node j receives the message $Message_setup(k, j, ID)$, it records the node index i , and sends an acknowledgement to i . If MN i receives the acknowledgement correctly, the routing path is finished to establish completely and data can be transmitted along the path bi-directionally.

In the routing procedures, there are two possible states for each router: the In-tree state and the Not-in-tree state. In detail, an AR belonging to bone node set may change state in three sub-states: In-tree and bone node state, In-tree and Non-bone node state, and Not-in-tree state. Bone node set maintenance process is operated in each AR. According to the receiving different messages and its current state, each AR may change its states among the three states.

5 Algorithm analysis

5.1 DATA STRUCTURE

It is clear from the above analysis that only the process of preparation computing is run in the multicast source s , and the other main parts (including process of maintenance for bone node set and process of routing for mobile node) are run in the mobile entities MN and AR. So we define the data structures for MN and AR as the following Fig. 2. Fig. 2(a) is the data structure for the subnet's AR to deal with mobile multicast routing, where *Group* is used to save the name of multicast group, *G* is a multicast session, *Member list* is used to record *G*'s member nodes, and *Bone node* is boolean and responsible for marking AR whether to be an element in bone node set N_b . Fig. 2(b) is the data structure for MN, where *Group* is used to save the name of multicast group of a mobile node, *Bone node list* records the element of N_b for each multicast group, *Distance* is used to save the shortest distance value/hop from MN to each bone node, and *Timer* is the variable to count the time.

The characteristics of the above data structure are as follows. It maintains a list of bone node in each MN to record the element in N_b for the current multicast. When MN switches from a subnet to other subnet, it scans the bone node list to find the most optimized bone node, which has the shortest distance/least hop to itself to re-join the multicast group. By this method, it can optimize the multicast tree cost and reduce the join delay. In the worst case, its space complexity is $O(n)$, and

computational complexity of scanning operations is also $O(n)$, which n is the network scales (that is, the number

of network nodes). See the next sub-section time complexity.

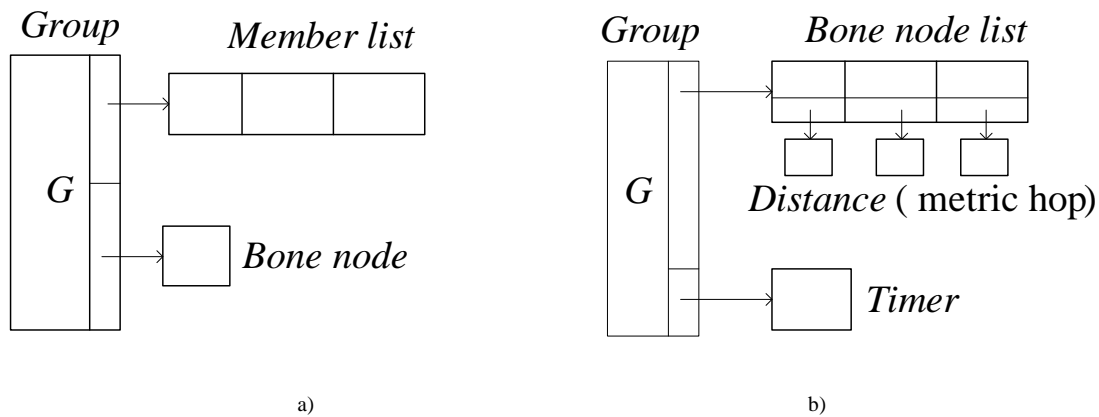


FIGURE 2 Data structures: a) Data structures for access router AR, b) Data structures for mobile node MN.

5.2 TIME COMPLEXITY

Supposed that the time complexity of generating a message is a constant 1 and the time complexity of transmitting a message is also a constant 1.

In step 4 the routing algorithm scans the bone node list and in the worst case the time complexity is $O(n)$; at the same time the time complexity of generating a Message_join(i, j, ID) is $O(1)$. Step 5 finds the next hop and in the worst case the time complexity is $O(n)$ because the network scale is n . In step 6, bone node returns an acknowledgement message and the time complexity is $O(1)$. So, the time complexity of distributed computing a routing path is $O(n+1+n+1)=O(n)$.

In steps 2 and 3 when a node i is added to or deleted from N_b , it generates a message Message_BNS_modify() with time complexity $O(1)$ and sends the message to each element in N_b . Because the number of the element in N_b is not more than $n-1$, so the time complexity is $O(1+n-1)=O(n)$.

So, the total time complexity of the distribute algorithm is $O(n)$.

5.3 MESSAGE COMPLEXITY

Supposed that the number of network is n , in the worst case there may be n nodes and $n-1$ edges on the path from source s to any one destination node. In this case, the message complexity of new algorithm in step 4 and 5 is $O(n)$. In step 6 message complexity is $O(1)$. Therefore, in the worst case the message complexity is $O(n+1)=O(n)$.

When a node i is added to or deleted from N_b (steps 2 and 3), it generates a message Message_BNS_modify() and sends the message to each element in N_b . It is clear that the message complexity is also $O(n)$.

So, the total message complexity of the distribute algorithm is $O(n)$.

6 Simulations

We assume that there are 49 subnets (a 7×7 mesh) in the simulation experiment [9, 10]. Each subnet includes a HA, a base station and an AR which can also perform as a multicast router if necessary. Assume the multicast source lies in the centre subnet of the topology and the multicast group members are randomly placed among the 49 subnets. Suppose that the distance count between the adjacent subnets is 1 hop.

In order to simulate mobility of MN, we define the mobility model as a 2-tuple (D, T), where D is the direction to move and T is the sojourn time to stay in the newly visiting subnet [11]. We assume that the mobile member node may randomly roam to each adjacent subnet. In the simulation the sojourn time T is exponentially distributed with the mean value 10 minutes. Besides, we also assume that there is only a multicast group with a single multicast source in the simulation. Initially the multicast tree includes only the source node and then some mobile nodes are randomly selected to add or remove from the multicast tree.

The main parameters and their values see Table 1.

TABLE 1 Simulation parameters

Parameters	Description	Value
N	Number of LANs	7×7
M	Number of multicast group	1
n	Number of group members	5-40
s	Sources per multicast group	1
D	The direction to move	Random
T	The mean sojourn time	Exponential distribution
Δ	The lower boundary for a bone node	20 minutes

The objective of the simulation experiments is to test the performance of BNSBMR and compare to three excellent same types of algorithm. Tested parameters

include tree cost and the average transmission delay. The two parameters can be defined as follows.

$$\text{Tree cost: } Cost = \sum_{e \in T} Cost(e).$$

Average transmission delay:

$$Trans_delay_average = (\sum_{i \in D} \sum_{e \in P(i,s)} Delay(e)) / n.$$

Parameter performances are tested by the following simulation experiments. Because RS and BT are the two basic multicast routing algorithms for mobile IPv6, and some other algorithms are originated or improved from them, we test and compare BNSBMR with the two algorithms. At the same time, due to RBMoM being a relatively successful same type algorithm in mobile IP area for multicast routing, we also compare BNSBMR's performances with RBMoM. When simulations are done, the number of destination/member nodes in equilibrium varies from 5 to 40, each time increasing 5, and all the member nodes are mobile nodes. We run a set of 1000 randomly generated handovers for each data point and calculated the average value over all runs as the final experiment data.

1) The tree cost. Firstly, the tree cost is tested and the simulation results are shown in Figure 3. The cost performance BNSBMR is the most optimized one in the four algorithms. RS algorithm also has a good cost performance. BT algorithm has the worst cost performance. The cost performance of RBMoM is between RS's and BT's. With the increasing of mobile nodes number from 5 to 40, the cost gap among the four algorithms becomes more and more serious. The reason is that BT relies on home agent HA and has a long transmission tunnel, which makes no any routing optimization. RS re-joins multicast tree from visiting subnet directly, as makes the cost performance is better. Because BNSBMR uses backbone node set to optimize the cost performance and bases on RS mode to reconstruct multicast tree, so it achieves the best cost performance. RBMoM is a compromise of RS and BT. Simulation results also verify the four algorithms' design idea well.

2) The average transmission delay. Average transmission delay is an important aspect for routing algorithm to provide QoS guarantee ability, which is mainly related to the path from the source node to the destination node. Experimental results are shown in Figure 4 according to the relation of the average transmission delay and the number of members. RS has the most optimal performance in transmission delay because a mobile node in RS method reconstructs routing tree directly by external AR, which cause the shortest path and the least transmission delay to the multicast source.

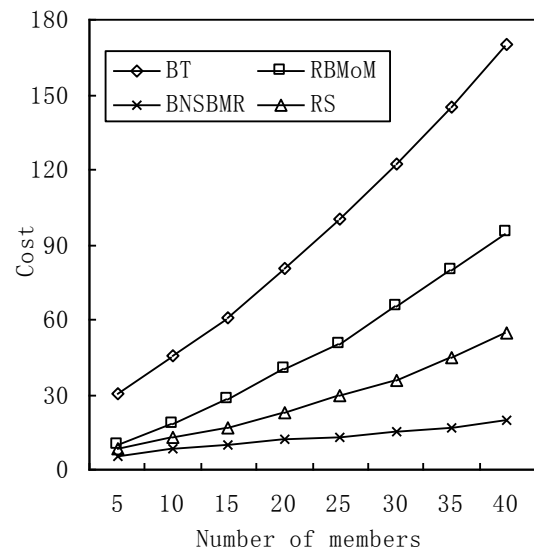


FIGURE 3 Relation of cost versus member number

The transmission delay performance of BNSBMR is only a little worse than RS but much better than RBMoM and BT, thanks to the bone node set used to share the routing path. The transmission delay performance of RBMoM is between BNSBMR's and BT's (supposed that the radius of parameters in RBMoM is $R=2$). BT has the worst transmission delay value also because it introduces a long delay by bi-direction tunnel in HA.

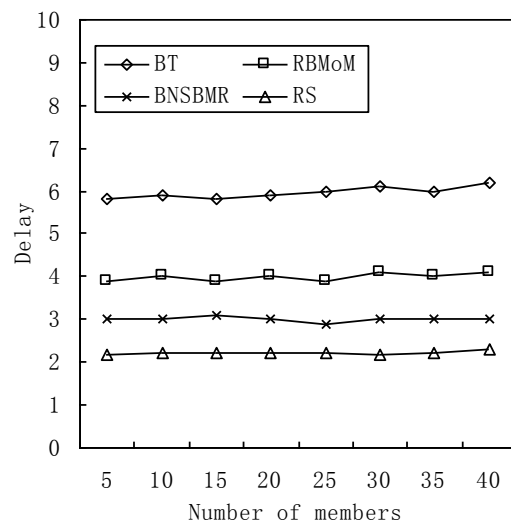


FIGURE 4 Relation of transmission delay versus member number.

7 Conclusions

In this paper, a distributed multicast routing algorithms BNSBMR is proposed based on bone node set for mobile IP, which can be efficiently used in large-scale mobile IP network and mainly has the following advantages.

- 1) The distributed algorithm is based entirely on MIPv6 protocol, and there is no extension or change in MIPv6 architecture. Also BNSBMR doesn't introduce any new protocol and entity and only relies on mobile

node MN and access router AR. In the process of running it needs only multicast source, MN, AR without any other irrelevant nodes.

- 2) BNSBMR optimizes the cost performance of multicast routing tree by bone node set, which is beneficial to the wire/wireless network bandwidth management.
- 3) The new algorithm optimizes the transmission path of multicast routing tree by bone node set, which reduces the average join delay when handover occurs and lessens the average transmission delay when data is

transmitted. It is clear that this is also helpful to realize the smooth handover.

- 4) BNSBMR is implemented based on distributed method which only depends on the bone node set with computing complexity $O(n)$. It makes the algorithm a good scalability, which is beneficial to realize multicast in large-scale mobile network.

In the future, we will do more experiments and test more parameters to verify the designed algorithm.

References

[1] Gossain H, Cordeiro C M, Agrawal D P 2002 *IEEE Commun. Mag.* **40**(6) 116-23
 [2] Kuntz M R, Noel J T 2013 *IEEE Commun. Mag.* **51**(1) 128-35
 [3] Lin C R, Wang K M 2002 *ACM Wirel. Netw.*, **8** 27-36
 [4] Wu J, Magnire G Q 2002 *IFIP Proc. PWC 2000*, Gdansk.
 [5] Suh Y, Shin H, Kwon D 2001 *ACM Wirel. Netw.* **7**(5) 443-53
 [6] Tan C L, Pink S Mobicast 2000 *ACM Mobile. Netw. Appl.*, **5** 259-

72
 [7] Seok J K, Gohar M 2010 *IEEE Commun. Lett.*, **14**(7) 676-78
 [8] Zhou L, Sun Y M 2008 *J. Comput. Res. Dev.t.* **45**(7) 1126-32 (in Chinese)
 [9] Madden S, Levis P 2008 *IEEE Internet Comput.* **12**(4) 9-11
 [10] Zhou L, Sun Y M 2006 *IEEE Proc. Int. Conf. Wirel. Commun. Netw. Mobile Comput.* 22-24
 [11] Suh Y, Shin H, Kwon D 2001 *ACM Wirel. Netw.* **7**(5) 443-53

Authors	
	<p>Ling Zhou, born on September, 1972, Yueyang, Hunan, China</p> <p>Current position, grades: Associate Professor; University of Foshan University study: Hunan Normal University, Changsha, 1998 Ph.D. degree in computer science and technology, Nanjing University of Science and Technology, Nanjing, 2007. Research activities: Computer Networks, Wireless Sensor Networks, Data Communication, and Routing Algorithm and Protocol. Publication: more than 10 papers in journals and international conferences including International Journal of Computer Networks & Communications, Journal of Computer Research and Development, Journal of System Simulation, IEEE proceedings of WiCom'2006, IEEE proceedings of CDCIEM'2012, etc. Experience: From July 2001 to June 2007, he has been a Lecturer in the Department of Communication Engineering, University of Changsha, Hunan, China. Since July 2007, he has been with Foshan University in the Department of Computer Science and Technology, and is currently an Associate Professor of computer science. Dr. Zhou is a member of CCF.</p>
	<p>Defeng Zhang, born on September, 1963, Dalian, Liaoning, China</p> <p>Current position, grades: Professor, University of Foshan, University study: Harbin Institute of Technology, Harbin, 1990. Research activities: Mathematics, Pattern Recognition, Statistical Decision Theory, Adaptive Control, and Computer Networks. Publication: more than 20 papers in journals and international conferences since 1993, including Journal of System Simulation, Acta of Sunyatseni University, Computer Engineering, etc. Experience: Since September 1993, he has worked in Foshan University in the Department of Computer Science and Technology, Guangdong, and is currently a Professor of computer science and automatic control. He is also a senior member of CCF.</p>

The study of campus network traffic monitoring platform

Bing Xu*

Chongqing Three Gorges University, Wanzhou Chongqing

Received 1 January 2014, www.tsi.lv

Abstract

From the view of practical campus network traffic monitoring platform, one kind of solution based on the model of SNMP and NETFLOW network management frame was put forward to elaborate the designed overall structure of campus network traffic monitoring platform, data acquisition, traffic plotting and so on. Using Visual C++6 to design this platform, not only the key technology and methods for realizing the campus network traffic monitoring platform could be achieved, but also the network traffic monitoring and management should be completed. The implementation of this platform can efficiently monitor the network traffic.

Keywords: Campus traffic, Network traffic, Traffic monitoring, management platform, VC++6

1 Introduction

With the development of computer network and expansion of communication scale, the network management as an important technology has become essential factor in the construction of digital campus. Due to the constant expansion of network scale, increasing complexity and enhanced isomerism, a campus network system usually includes a number of subsystems, integrates a variety of network operating systems, and collects network application equipment from different manufactures. In addition, the campus network system needs to be supported by a lot of network application service software. Therefore, we need an efficient campus network monitoring and management platform to for effective management [1]. In addition, the implementation of campus network traffic monitoring and management platform has important significance in network traffic billing, network security, etc.

2 Design concept

Network traffic monitoring and statistical analysis are important constitutional parts of network management and maintenance process. In recent years, with the constant expansion of the network scale, the increasing complexity of network and constantly emerging of new network services, the issues that understand and accurately describe the characteristics of internet network traffic and network behaviour model have become increasingly outstanding [2]. Moreover, the good management contributed by traffic monitoring and statistical analysis has become more and more important. Mainly based on the characteristics of network traffic and foundation of network, the construction of traffic monitoring platform should be established.

3 Solution

In order to timely view the network traffic, net resources and the application program, and improve the network performance, the solution based on the model of SNMP and NETFLOW network management frame was developed. In addition, the information acquisition system based on the Windows platform was explored and realized. According to the needs from network management, the provided information service could determine a certain self-interested domain, router and its port, traffic of fixed IP address.

4 Design and implementation

4.1 OVERALL STRUCTURAL DESIGN

Based on the collected and analysed user requirements from network traffic monitoring software, the features were summarized below:

- 1) Capture the network interface data packet as much as possible, set the card to promiscuous mode and then capture the data packet;
- 2) Resolve the content of data package, and analyse its protocol type, source address, acquisition time and so on;
- 3) User-defined monitor according to different requirements from users through given address range, related package had specific agreements, etc.;
- 4) Yield the monitoring results, which should contain real-time traffic charts, lists and so on;
- 5) Record the logbook for later analysis;
- 6) Detect and analyse some common attacks;

The platform written by VC++6.0 mainly had three modules which could be divided into data acquisition and display module, traffic information statistical module,

* *Corresponding author* e-mail: cqwx888@163.com

and traffic drawing module. The schematic diagram of platform was shown in Figure.1.

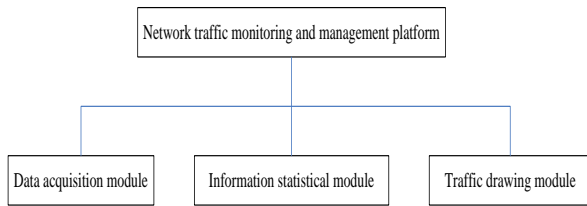


FIGURE 1 The component modules of the platform

Data acquisition module (*capture, analyse and display the network interface data*) is able to capture data based on user defined conditions, such as only monitor the data packets using TCP or UDP protocol, also monitor the data packets that related the IP addresses users want to pay close attention to, complete the data packets logging, enhance the flexibility of the system; meanwhile, judge some common attacked characteristics in the analysing process of data packets and send out warnings.

Information statistical module completes statistical functions, such as accounting the received number of data packets which could help IP finishing the statistics, the number of mistaken protocols among the received data, the number of messages regarding applying the transfers, the number of available routers in the list of routers, the number of discarded routers, the quantity of requiring organization/successful organization and so on, accounting ICMP required the number of messages including sending/receiving, exceeding the TTL, redirecting, timestamp request/reply and so on (the IP assistant function is used)

Traffic drawing module displays the total network traffic, input traffic, output traffic, instantaneous traffic value and maximum traffic value; acquires the related data through accessing the network performance data in the registry, displays these data by traffic diagram.

4.2 DESIGN OF SPECIES IN THE COURSE OF ACQUIRING DATA PACKET

Both overall function and work process of the system should be obtained based on the analysis and design as above. In addition, a series of operations including establishing, binding, setting work mode, creating thread, receiving data and so on, would take place in the process from editing socket to finally acquiring data. In order to solve every step in the process, the relationships of species regarding acquiring data were designed. The schematic diagram was shown in Figure.2.

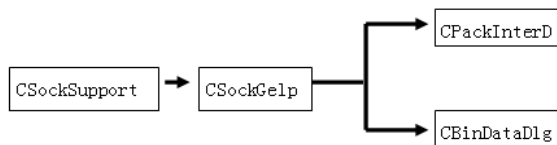


FIGURE 2 Relationships of species in the course of acquiring data packet

As shown above, like CSockSupport, CSockHelper, CPackInterDIg, CBinDataDIg and other species, different species were made up of different processing parts. In addition, these species contained relevant operational approaches.

4.3 MODULE OF DATA PACKET ACQUISITION AND ANALYSIS

4.3.1 Function specification

This functional module was mainly composed of CSockSupport, CSockHelper, CPackInterDIg and CBinDataDIg. These species would be described in detail as follows.

CSockSupport: mainly responsible for checking whether Socket was suitable for Version 2.0, embedded WSASocket was able to start Socket;

CSockHelper: mainly achieved these activities such as acquiring the information structure from local computer, creating Socket, binding, establishing, creating thread, and all of the methods from data reception to protocol analysis. Detailed processes were shown in Figure.3.

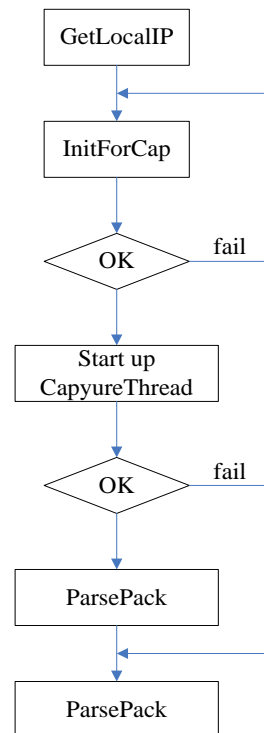


FIGURE 3 Flow chart of data packet acquisition

GetLocalIP implemented the method of acquiring the local address. The information structure of host computer was defined by LPHOSTENT lphp. The acquired process was completed by gethostname (szLocname, MAX_HOSTNAME_LAN) and gethostbyname (szLocname). The first parameter was used to deposit the buffer of local computer's name, and the second one was the length of buffer [3]. Finally, IP address was converted into "." address with inet_ntoa.

4.3.2 Creating socket, binding, setting up operational mode and creating thread

StartCapture could finish the creating socket, binding, setting up operational mode and creating thread. The detailed processes were shown as below:

```
m_sockCap = socket (AF_INET, SOCK_RAW,
IPPROTO_IP);//create socket
```

```
bind (m_sockCap, (PSOCKADDR)&sa, size of
(sa));//bind
```

```
setsockopt (m_sockCap, SOL_SOCKET,
SO_REUSEADDR, (char*)&bopt, sizeof (bopt));//set
operations
```

```
setsockopt(m_sockCap, IPPROTO_IP, IP_HDRINCL,
(char*)&bopt, sizeof(bopt));// set operation
```

WSAIoctl

```
(m_sockCap,SIO_RCVALL,&dwBufferInLen,size of
(dwBufferInLen), dwBufferLen, size of (dwBufferLen),
&dwBytesReturned, NULL,NULL);//promiscuous mode
```

```
m_hCapThread = CreateThread (NULL, 0,
CaptureThread, this, 0, NULL);//start thread
```

The data reception was finished by a thread function named **CaptureThread**. After data reception, the resolution process took place unless the data in buffer zone had been converted into IP data [4]. The detailed process of acquiring protocol names was shown as follows.

```
for(int i=0; i<MAX_PROTO_NUM; i++)
    if(ProtoMap[i].ProtoNum==iProtocol)
        return ProtoMap[i].ProtoText;
    return "";
```

ParIPPack was able to resolve the data packet.

```
int iIphLen = size of(unsigned long) * (pIphheader-
>h_lenver & 0xf) //acquire the length of data packet
```

Resolution protocol was shown as below.

```
switch(iProtocol)
{
case IPPROTO_TCP :
.....
case IPPROTO_UDP :
.....
case IPPROTO_ICMP :
.....
default :..... }
```

The thread and socket could be closed by **StopCapture** as follows.

```
if(m_hCapThread)
{ TerminateThread(m_hCapThread, 0);
//terminate process
CloseHandle(m_hCapThread); //close
handle
m_hCapThread = NULL;}
if(m_sockCap)
closesocket(m_sockCap); //close
socket
```

CbinDataDIg was mainly responsible for depositing and displaying the acquired data. The targets of **CbinDtaDIg** and **CsockHeliper** created by **CpackInterDIg** were used to acquire, analyse, display and deposit data. Moreover, **CpackInterDIg** could edit these controls including setup conditions for acquiring and record daily work.

4.4 TRAFFIC DRAWING MODULE

4.4.1 Design specification

In the practical process of programming, one kind of interfaces supplied by windows system was able to access relevant data of network performance, such as flux. This module was divided into three sub functions that were the sub module that can access to the performance data (take charge of accessing registry and acquiring data), display sub module (draw the data in the window), and frame sub module (reflect and deal with message), respectively.

In this module, one function called **RegQueryValueEX** was used to access the registry [5]. This function would search relevant styles and data through the open registry keys and names. The prototype of this function was shown as follows.

```
LONG RegQueryValueEx(HKEY hKey, LPCTSTR
lpValueName, LPDWORD lpReserved, LPDWORD
lpType, LPBYTE lpData, LPDWORD lpcbData);
```

Function parameters: **hKey** was the default key assignment of registry; **lpValueName** was the name of key assignment which was needed to be searched for; **lpReserved** should be reserved unless it was **NULL**; **lpType** was the style of key assignment; **lpData** was the data that input/output received key assignment; **lpcbData** was the mark of buffer size used for input/output received key assignment [7].

In Windows 2003, when **RegQueryValueEx** was called, the returned data was not directly displayed as requested data if the **hKey** was set as **HKEY_PERFORMANCE_DATA**. Therefore, programs had to traverse the entire data block [6]. The logic structure in the data block was shown in Figure.4.

The searching process for ensuring performance data block was shown in Figure 4. It started from the performance data structure named **PERF_DATA_BLOCK** to **PERF_OBJECT_TYPE**. Then, the **PERF_COUNTER_DEFINITION** could find the site deviation through **HeaderByteLength**, which belonged to **PERF_OBJECT_TYPE**. Each unit of

PERF_OBJECT_TYPE called DefinitionLength had the ability to ascertain the structure of corresponding PERF_INSTANCE_DEFINITION. The structure of PERF_COUNTER_BLOCK was controlled by the structure of corresponding

PERF_INSTANCE_DEFINITION [8]. The traffic diagram of this module in the course of running was shown in Figure 5.

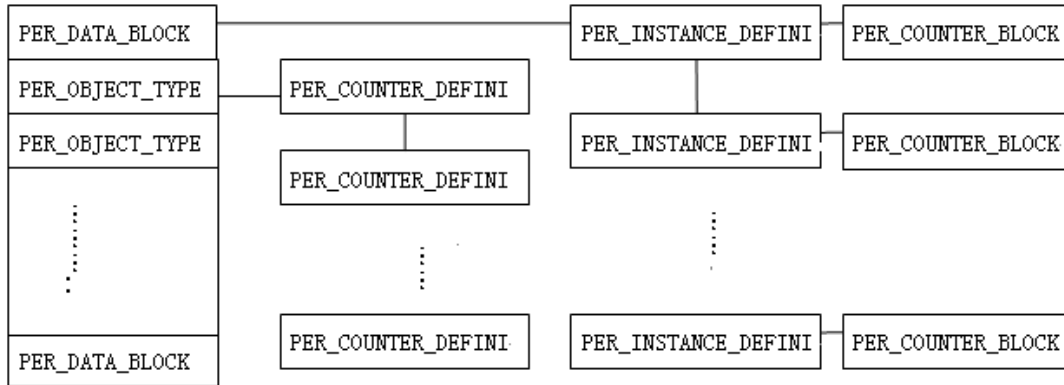


FIGURE 4 The logic structure of network performance data block in registry

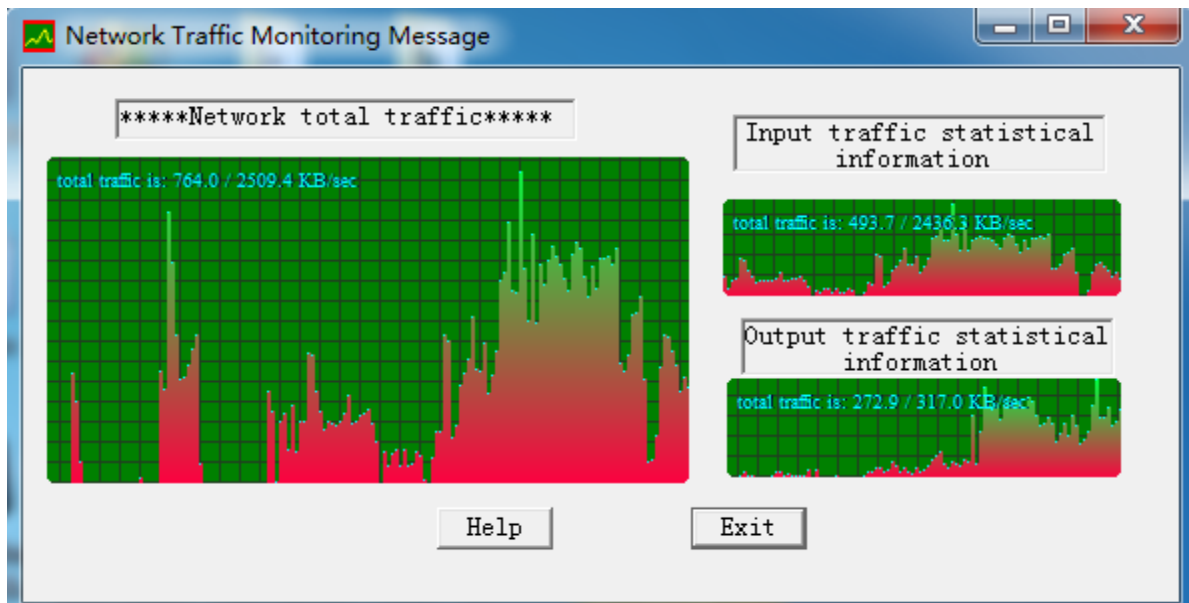


FIGURE 5 Traffic drawing diagram

4.4.2 The main codes in the traffic drawing module

//The function codes, which were defined by users were displayed by traffic when the network was in the connective state.

```

void
MFTrafficButton::OnTimer(UINT nIDEvent)
{
    // TODO: Add your message handler code here
    and/or call default
    if(nIDEvent == NETTIMER)
    {
        #ifdef
        _I_HAVE_PLATFORM_SDK_INSTALLED_
        DWORD flag, reserved;
        BOOL erg;
        flag =
        0;//INTERNET_CONNECTION_OFFLINE ;
    
```

```

        reserved = 0;
        TCHAR connectionname[1024];
        erg = InternetGetConnectedStateEx(
            &flag, //OUT LPDWORD
            lpdwFlags,
            (LPTSTR)&connectionname,//OUT LPTSTR
            lpszConnectionName,
            1024,//IN DWORD
            dwNameLen,
            0//IN DWORD dwReserved
        );
        #else
        isOnline = erg;
        #endif
        // the current traffic was obtained
        double traffic =
        m_cTrafficClass.GetTraffic(SelectedInterface);
    
```

```

        DWORD totaltraffic =
m_cTrafficClass.GetInterfaceTotalTraffic(SelectedInterfa
ce);
        double delta1;
        double divisor =
(1000.0/(double)NETUPDATESPEED);
        delta1 = (double)(traffic * divisor) /
1024.0;
        CurrentTraffic.Format("current
traffic: %.1f KB/sec",delta1);

        // Should we recalculate the local
maximum per session or per display?
        if(useAdaptiveScale==TRUE)
        {
            MaxTrafficAmount = 0.0;
        }
        // Shift whole array 1 step to left and
calculate local maximum
        for(DWORD x=0; x<TrafficEntries;
x++)
        {
            TrafficStats[x].connected =
TrafficStats[x+1].connected;
            TrafficStats[x].value =
TrafficStats[x+1].value;
            MaxTrafficAmount)
            if(TrafficStats[x].value >
MaxTrafficAmount)
                MaxTrafficAmount =
TrafficStats[x].value;
        }
        if(isOnline == TRUE)
        {
            TrafficStats[TrafficEntries].connected = TRUE;
            TrafficStats[TrafficEntries].value = traffic;
            if(TrafficStats[TrafficEntries].value >
MaxTrafficAmount)
                MaxTrafficAmount =
TrafficStats[TrafficEntries].value;
        }
    }
    else
    {
        TrafficStats[TrafficEntries].connected = FALSE;
        TrafficStats[TrafficEntries].value = traffic;
        if(TrafficStats[TrafficEntries].value >
MaxTrafficAmount)
            MaxTrafficAmount =
TrafficStats[TrafficEntries].value;
    }
    double delta2;
    divisor =
(1000.0/(double)NETUPDATESPEED);
    delta2 = (double)(MaxTrafficAmount *
divisor) / 1024.0;
    MaximalTraffic.Format("maximal
traffic: %.1f KB/sec",delta2);
    AllTraffic.Format("total traffic is: %.1f
/ %.1f KB/sec",delta1, delta2);
    }
    // draw
    Invalidate(FALSE);
    CButton::OnTimer(nIDEvent);
}

```

5 Conclusions

In the paper, from the view of campus network traffic monitoring platform, one kind of solution was put forward to elaborate the designed overall structure of campus network traffic monitoring platform, data acquisition, traffic plotting and so on. The database management technology and effective flexible warning system that could improve efficiency were applied into this platform. These technologies were not only able to record the time and site of fault, but also provide reliable information to analyse and solve the fault in time. Besides, the cost of platform development and operation was not high and the product was cost-effective. This platform was suitable for network of college and network with medium size.

References

- [1] Phaal P, Panchen S, McKee N *Mon Corporation's sFlow: A Method for Monitoring Traffic in Switched and Routed Networks, RFC3176* September 2011
- [2] *PPT 'sFlow & Benefits'* www.sflow.org Oct 2013
- [3] Zseby T, Molina M, Duffield N, Niccolini S, Raspall F 2010 Sampling and Filtering Techniques for IP Packet Selection *draft-ietf-psamp-sample-tech-06.txt* February 2010
- [4] Dietz T, Claise B 2012 Definitions of Managed Objects for Packet Sampling *draft-ietf-psamp-mib-04.txt* February 18 2012
- [5] Xu Peng, Qiong Liu, Sen Lin 2011 Internet traffic classification using support vector machine *Journal of Computer Research and Development* **46**(3) 407-14
- [6] Xu Peng, Sen Lin 2011 Internet traffic classification using C4.5 decision tree *Journal of software* **10**(20) 2692-704
- [7] Schroder C 2009 *Linux networking cookbook* (Translation Fen Liang) Nan Jing: Southeast University press 426-487(in Chinese)
- [8] Kundu Dinangkur, Lavlu S M Ibrahim *Cacti 0.8 network monitoring* UK: Packt Publishing Ltd 5-110

Authors



Xu Bing, born in February 10, 1976, Chongqing, China

Current position, grades: School of Computer Science and Engineering, Associate Professor and master supervisor in Chongqing Three Gorges University

University studies: Computer Science and Engineering in Chongqing University

Scientific interest: computer network and application

Publications: 3 Patents, 40 Papers

Research direction: computer network and application

The gait analysis on the sloping walking of goat

Fu Zhang^{1,2*}, Yakun Zhang¹, Binbin Yue¹, Guoying Zhang¹

¹ College of Agricultural Engineering, Henan University of Science and Technology, Luoyang, 471003, P R China

² Henan Haofeng Machinery Manufacturing Co, LTD, Xuchang, 461103, P R China

Received 15 January 2014, www.tsi.lv

Abstract

The 18° sloping walking state and movement rules of the goat was researched by the high-speed video camera system. The movement process and the imaging results of goat in 18° sloping fields were recorded in the computer. The experiment imaging results of goat movement process were analysed by SigmaScan software and Matlab software, the results showed that gait parameters and angle change curve of each leg on 18° slope was obtained. The research will provide the basis of the experimental data for bionic design of agricultural machinery of goats sloping walking mechanism.

Keywords: Goat, High-speed camera, Sloping fields, Gait

1 Introduction

The walking gait is a useful biometric that is the most commonly used of the nature quadruped, is the base of study on ground walking mechanism [1-3]. In recent years, the researches on walking state and movement rules has been reported a lot at home and abroad, which concerned on all kinds of quadruped, while less on the goat which has small body, firm limb bones, the ability of walking gait agile brisk and walking on the uneven surface of the ground, slopes, steep and mountainous freely and neatly, and can endure long-distance walking. The goat can be used as the base research of the gait walking mechanism [4-11]. In this study, it was shot by the high-speed camera that the movement state of goat in 18° sloping fields, and the obtained images were analysed by SigmaScan software and Matlab software.

2 Experiment equipment and methods

The walking process of a goat in the slope has been shot by a high-speed camera (up to 10,000 per second) in the experiment. The high-speed camera was adjusted to ensure that the shooting speed was adapted to the movement of the goat in different sloping fields. The moving sequence images of the goat in slope fields were recorded and saved to the computer. Subsequently, the data of the goat key point body were collected and analysed by SigmaScan software. The characteristic parameters and curves of the corresponding series, which were on the movement of goat in different sloping fields, were obtained.

2.1 EXPERIMENT EQUIPMENT AND OBJECTS

2.1.1 Experiment equipment

It was consisted of a high speed camera (FASTCAM - Super 10KC), a lamp (MODEL10000), a mechanical scale (XSJ 2 x 1300 w - 20), a microcomputer (P4—3.0 G), a tape and a inclinorator, and so on (Fig 1) that were used in this article. The high-speed camera equipment was mainly composed of the host processor, laptop, monitors, lenses, lighting device and data lines, etc. [4, 5].



FIGURE 1 High-speed camera equipment

2.1.2 Experiment sites and objects

The sloping fields, which were measured by inclinorator, were 18 degrees in the experiment. The data parameters of spatial movement of the key points, which were on the movement of the goat in sloping fields at any time, were obtained by high-speed camera.

Goats (Goat) belong to artiodactyla, were mainly composed of the head, neck, torso and limbs. They have

*Corresponding author e-mail: zhangfu30@126.com

solid head, good combination of the neck, body and forelimb, broad and deep chest, open ribs, generous and straight back waist, wide, long and plump hindquarters, strong limbs, well-balanced structure, moderate height. a female goat (1.5 years old) was analysed in the test (Fig. 2). The main posture characteristic was showed (Table 1). The test was operated on the outskirts of Luoyang City in Henan province, and the temperature was 32~35°C.



FIGURE 2 The photos of the goat

TABLE 1 The photos of the goat

Body length (cm)	Body height (cm)	Taller than	Weight (kg)
138±0.5	90±0.5	1.53	42±0.5
The length of first calf (cm)	The length of first thigh (cm)	The length of latter calf(cm)	The length of latter thigh (cm)
16±0.5	30±0.5	23±0.5	23±0.5

2.2 EXPERIMENT PROCEDURE

In order to reduce the photography error and make the processing of image datum easily, the high-speed camera

was used to shoot the whole moving process of the goat, which was towed to ensure it always walked in a straight line on the sloping fields.

During the experiment, the frequency of high-speed camera was adjusted to 10000 frames per second. The high-speed camera equipment was set on the basis of the test purpose. The camera was fixed on a tripod and the lens was set at the forward side of the goat motion. The hardware conditions of the camera, which including space position, point range, shooting range, aperture, focal length and so on, were adjusted to ensure that the main lens optical axis focused on the range centre of the goat movement. The main lens optical axis was adjusted to the goat moving plane as close as possible in order that the goat was clearly imaged. Then the whole sequence images that the goat walked along the straight line in the sloping fields were recorded and saved into laptop through image capture card. The trajectory of the goat walking process was shot expediently and completely under the condition that the goat walked normally.

2.3 EXPERIMENT DATA PROCESSING METHODS

2.3.1 The use of SigmaScan software

SigmaScan software was installed and the icon of the software was clicked directly. A new worksheet was created under the SigmaScan software environment. Then the picture was opened (Figure 3).In the image measurement, the lateral displacement values of the key marked points in the goat body were defined as the X-axis value (along a particular path direction), and the longitudinal displacement values were defined as the Y-axis.

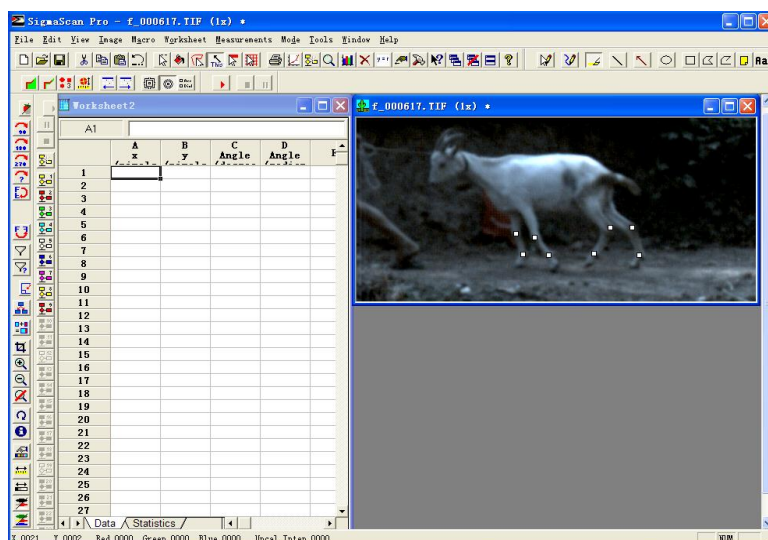


FIGURE 3 The joint marked point position of the goat

The size of the pixel value was equal to the coordinate size. The quadrant maximum values were limited by the resolution of high speed camera (i.e. the pixel size of image). The resolution values of high speed camera in this test were 512 (horizontal direction) ×240 (vertical

direction), so the quadrant maximum value was 512 (X axes) ×240 (the Y axis) in the experiment. Therefore the data values ranged from 0 to 512[6-8]. The relative coordinate values of the key points of the goat movement in the slope fields were obtained when the key points of

the goat limbs of in the gait cycle image sequences were clicked. The data were preserved in the Excel document format.

2.3.2 Data processing

The Excel data documents (Fig 4) which were processed by the SigmaScan software were arranged in chronological (if the shooting speed was 200 frame per second, then the time series was 0ms, 20ms, 40ms and so on).

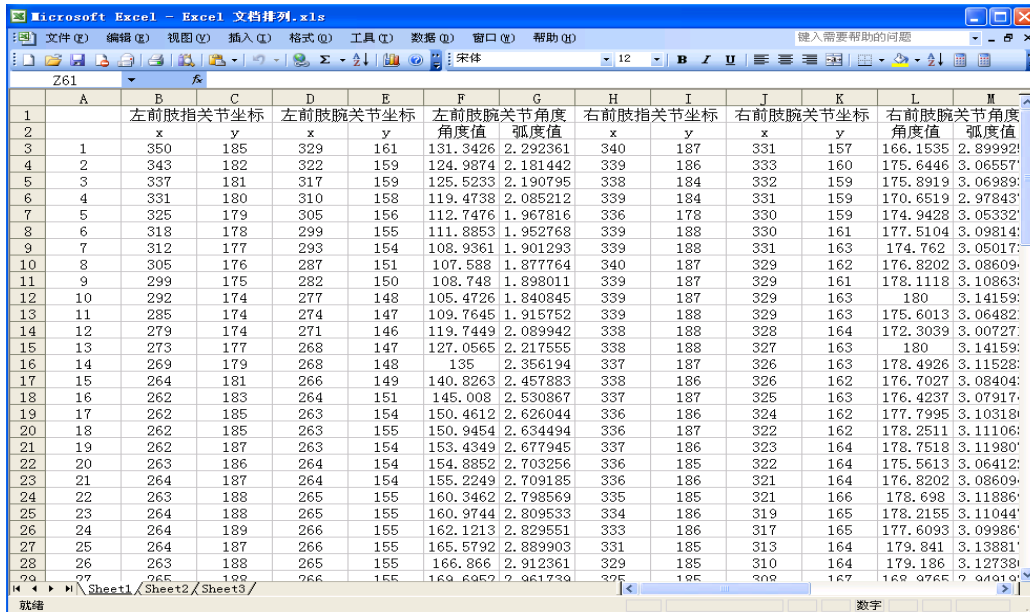


FIGURE 4 Excel document data

The obtained data sequence can be analysed and calculated by Excel directly, and it also can be processed by Matlab software when the data were imported. The corresponding series of movement characteristic curve were obtained, such as the joints point motion curve of goat leg, the range of motion and so on.

3 Test results and analysis

3.1 THE ANALYSIS OF GOAT GAIT CYCLE

A complete gait cycle goat figure (Figure 5) was selected at the time interval between every two pictures 240ms (12 frame). Besides a complete gait cycle that the goat walked on this sloping fields, was 904ms (452 frame).

In a complete gait cycle, the walking speed of the goat was 31.4cm per second and the footstep was 66.0cm on 18° sloping fields. It was discovered that the walking speed and walking step were decreased and then increased with the slope gradually increased.

In a complete gait walking cycle, the corresponding single leg span, single leg step and load factor of the goat in 18° sloping fields were shown in table 2. It is discovered that the single leg span and step of the goat left hind leg were the smallest, which were smaller than the other three legs, while the load factor of each leg was the same. It was the reason why the goat can walk on sloping fields stably.

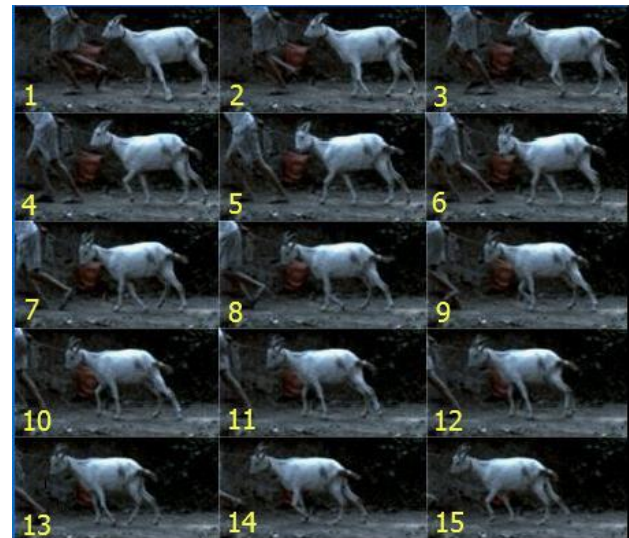


FIGURE 5 The gait cycle walking figure of the goat

TABLE 2 The single leg span distance of the goat

	single leg span distance (cm)	single leg step distance (cm)	load factor
The left front leg	61.0	17.0	0.75
The left hind leg	45.5	9.0	0.75
The right front leg	54.5	14.0	0.75
The right hind leg	58.5	13.5	0.75

The state diagram of the goat legs were shown in Fig 6 when the goat was climbing in 18° sloping fields. The period of three-legged support had emerged a total of four times during a complete transformation gait cycle of the goat walking. 520ms had been taken by the period when

the left hind leg was hanging and other three legs were at the support condition. 496ms had been cost by the period when the left hind leg was on the ground, the left front leg lifted off the ground, and other three legs were at the support condition. 512ms had been cost by the period when the goat body gradually moved forward, the left front leg took a step and was on the ground, the right hind leg lifted off the ground and was at the hanging condition, and the other three legs were at the support condition. 544ms had been spent by the period when the right hind leg took a step and was on the ground, the right front leg lifted off the ground and was at the hanging condition, and the other three legs were at the support condition. 512ms had been spent by the period when the body gradually moved forward, the right front leg took a step and was on the ground, the left hind leg lifted off the ground and was at the hanging condition, and other three legs were at the support condition. So far, the whole transformation gait cycle of goat had been completed.

Then, 416ms had been spent by the period when the left front leg was at the hanging condition and the other three legs were at the support condition. 400ms had been cost by the period when the right hind leg was at the hanging condition and the other three legs were at the support condition. 500ms had been taken by the period when the right front leg was at the hanging condition and the other three legs were at the support condition. Walking repeatedly as mentioned above, the goat's body was supported by the cyclical alternation of the front and hind leg. The body centre of gravity was projected on the triangle stent which was supported by the three legs, thus the balance and stability of the goat has been guaranteed during goat walking.

In a complete gait cycle, the proportion of each goat leg's support period of the whole gait cycle was shown in Table 3. It is discovered that the condition of three goat legs common support always existed during the walking in 18° sloping fields.

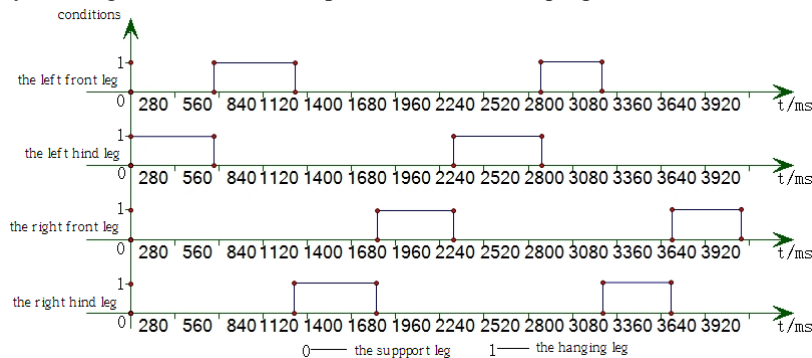


FIGURE 6 The state chart of each leg of the goat

TABLE 3 The proportion of the leg support period of the whole gait cycle

Type of Support	The number of occurrences of a single-cycle
Common support of the diagonal legs	0
Common support of the ipsilateral legs	0
Common support of the three legs	1
Common support of the four legs	0

The reason why there were a total of four times period of three legs common support were that the hind leg lag behind the front leg which was diagonal to the hind leg when the legs were on up and down during the goat walking. The hanging time of hind leg was longer than the front leg when the goat started moving at the base of slope. However, the hanging time of the front and hind legs had a sudden reduction compared with the initial movement at the base, and the hanging time of hind legs was lower than the front.

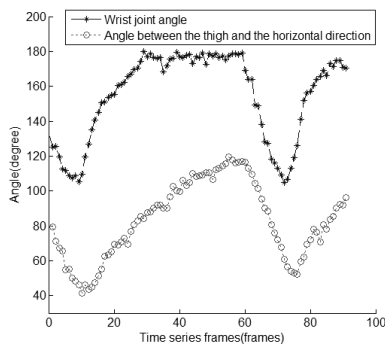
3.2 THE ANGLE ANALYSIS OF THE GOAT LEGS

The angle change range value of each leg measured was shown Tab.4 during the walking in 18° sloping fields.

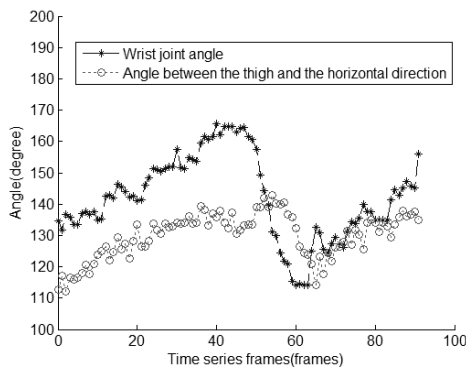
TABLE 4 The angle change range value of each leg

Slow walking in sloping fields	Angle range
α angle (front legs)	100°~185°
β angle (hind legs)	46°~139°
α angle (hind legs)	103°~169°
β angle (front legs)	111°~158°

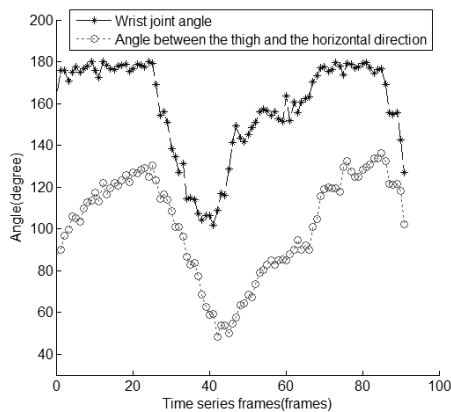
Processing with the experimental data of the whole gait cycle, and the angle change curves of goat four legs were obtained during the walking in sloping fields. (the frame number between angle and time series of the goat left forelimb, the frame number between angle and time series of the goat left hind limb, the frame number between angle and time series of the goat right forelimb, the frame number between angle and time series of the goat right hind limb were shown in a, b, c and d of figure 7 correspondingly). Among them, the angle between the things and the calfs was always presented by the α angle when the goat walking, namely the wrist angle. The angle between the things and the forward direction was always presented by β angle, namely the angle of thigh and horizontal direction.



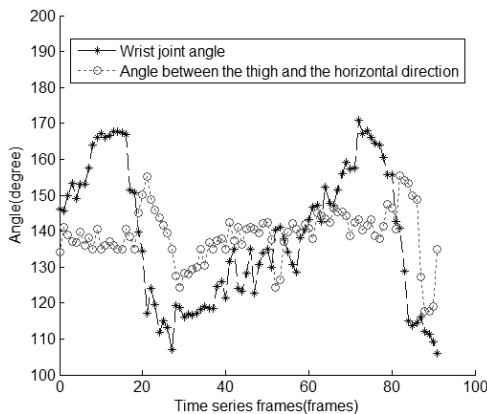
a) The frame number between angle and time series of the goat left forelimb



b) The frame number between angle and time series of the goat left hind limb



c) The frame number between angle and time series of the goat right forelimb



d) The frame number between angle and time series of the goat right hind limb

FIGURE 7 The frame number curve between angle and time series of the goat in 18° slope fields walking.

The legs angle changes of the goat walking in 18° slope fields were as follows.

The left forelimb: the α angle reached the minimum firstly when the leg was bending and lifting, and achieved the maximum after a forward stretching out and stride, then attained the minimum after falling off again and again. The β angle reached the minimum when the α angle was the minimum firstly (the leg was bending), and achieved maximum when the α angle was the minimum secondly (the leg was on a forward stretching out and stride).

The left hind limb: the α angle reached the maximum firstly when the leg was on a forward stretching out and stride, and achieved the minimum when the leg was shrinking and bending, then the leg was on a forward stretching out and stride again. The β angle was little when the α angle was the maximum firstly. However, it increased along with the decreased of α angle gradually, and was the maximum when the leg was ready to fall on the ground.

The right forelimb: the α angle reached the maximum firstly when the leg was on a forward stretching out and stride. After a forward stride, the leg fell on the ground, the body gradually was leaning forward and the gravity centre was moving forward, the α angle reached the minimum when the feet pedalled firmly, and the leg was bending at the same time. The β angle had linearity change with the α angle changes when the leg was on bending and falling on the ground. Namely, the β angle increased with the increased of the α angle, decreased with the decreased.

The right hind limb: the α angle reached the maximum firstly when the leg was lifting, after a forward stride, the leg fell on the ground, the body was gradually leaning forward and the gravity centre was moving forward, the α angle reached the minimum when the feet pedalled firmly, and the leg was bending at the same time. The β angle was on the intermediate state when the α angle was the maximum firstly (the leg was lifting), and was the maximum when the α angle reached the minimum (the leg was bending).

Discoveries were found from Table 6 and graph 11 that, the initial values of β angle of front legs were the minimum, while the final values of α angle of front legs were the maximum, the initial and final values of α and β angle of hind leg were larger than the front legs. The range of α and β angle of hind leg, however, were smaller than the front legs. Some characteristics of leg angular variation, which the goat was walking in slope fields were as follows: The leg angle changing curves between the left front leg and the right front leg had the similarity and regularity, and the leg angle changing curves between the left hind leg and the right hind leg also.

4 Conclusion

- 1) The test process and the test methods of the goat walking in sloping fields were studied, and the method of SigmaScan software applications in the high-speed camera of was key analysed, and the kinematic data of the key point in goat body were collected and classified at the same time. A more complete kinematic data parameters had been obtained, which provided the research basis for gait analysis.
- 2) The experimental sequences of image that had been saved in computer were processed and analysed by SigmaScan software, the gait feature parameters of

gait and the leg angle changing curves and regulations of the legs were founded when the goat was walking in 18° sloping fields, and the basic experiment data for goat-like design of slope walking mechanism were provided.

Acknowledgements

This work was supported by 2012 Post Doctoral Research Project of Henan Province, Scientific Research Foundation for Ph. Doctor, and Innovation Ability Foundation of Natural Science (Grant No. 2013ZCX002) of Henan University of Science and Technology.

References

- [1] Zhang Fu, Zhang Guoying, Qiu Zhaomei et al 2011 The Status and Progress of Gait Research of Bionic Running Mechanism on the Ground *Journal of Agricultural Mechanization Research* **33**(1) 240-3 (in Chinese)
- [2] Chen Donghui, Tong Jin, Li Chonghuan et al 2003 A Review of man and animal gait and walking robot” *Journal of Jilin University (Engineering and Technology Edition)* **33**(4) 121-125(in Chinese)
- [3] Autumn K, Hsieh S T, Dudek D M, Chen J, Chitaphan C, Full R 2006 *Journal of Experimental Biology* **20**(9) 260-72
- [4] Sun Xiangyi, Hu Jian, Wang Kunpeng, et al 2010 Technique and Application of 3-D Image Analysis for High-speed Video *Journal of Astronautic Metrology and Measurement* **30**(6) 30-4(in Chinese)
- [5] Liu Chaolao, Ruan Ping, Xiong Rensheng et al 2001 The Computer Aided Method of High Frequency Photographing Measurement *Acta Photonica Sinica* **30**(1) 113-6(in Chinese)
- [6] Jia Yunqi, Cao Zhizhong, Wang Yunwen et al 2009 Dynamic analysis of autumn color changes based on digital imaging technique in zoysiagrass *Acta Photonica Sinica* **18**(3) 94-102
- [7] Wang Lidong, Wang Liangyi, Yang Haitao et al 2007The application of SigmaScan Pro5 software in spine Imaging measurement and analysis *Chinese Journal of Postgraduates of Medicine* **30**(z1) 90-1 (in Chinese)
- [8] Chen Shenglin, Wu Zhihua, Ma Shengjian 2006 Measuring Methods of Eucalypt Leaf Area with Digital Image Processing Technology *Eucalypt Science & Technology* **23**(1) 6-10 (in Chinese)
- [9] Pa P A Design of a modular assembly of four-footed robots with multiple functions 2009 *Robotics and Computer-Integrated Manufacturing* **25**(4/5) 97-105 (in Chinese)
- [10] Xu Yiqun, Wan Changjun 2001 Gait stability analysis of quadruped walking robot *Manufacturing Automation* **23**(8) 5-7 (in Chinese)
- [11] Tan Xiaoqun, Li Jun, Zhao Guobin et al 2008 Application of Center Gravity Adjusting Device on Static Walking of Quadruped Robot *Manufacturing Information Engineering of China* **37**(23) 25-8 (in Chinese)

Authors

	<p>Fu Zhang, born on June 25, 1978, Xingtai, China</p> <p>Current position, grades: Doctor of Agricultural Engineering, associate professor and the Deputy Dean in Henan university of Science and technology.</p> <p>University studies: bachelor degree and master degree from the college of mechanical and electronic Engineering of Hebei Agricultural University in 2001 and 2004 respectively, and received Ph. D. degree from the College of biological and Agricultural Engineering of Jilin University in 2007</p> <p>Scientific interest: agricultural equipment technology and automation, biomimetic technology of the terrain machinery.</p> <p>Publications: 40 Patents, 40 Papers, 5 books.</p> <p>Experience: a supervisor of master program in agricultural mechanization engineering, the Deputy Dean of agricultural engineering department, member of China agricultural engineering society and youth work committee member of china society for agricultural machinery, 19 research projects including 2 state projects, 4 provincial projects, 9 departmental projects and 4 horizontal subjects.</p> <p>Research interests: agricultural equipment technology and automation and biomimetic technology of the terrain machinery.</p>
	<p>Yakun Zhang, born on February 3, 1991, luoyang, China</p> <p>Current position, grades: College of agricultural engineering in Henan university of Science and technology.</p> <p>University studies: Agricultural Engineering in Henan university of Science and technology.</p> <p>Scientific interest: agricultural equipment technology and automation, biomimetic technology of the terrain machinery.</p> <p>Publications: 12 Patents, 3 Papers</p>
	<p>Binbin Yue, born on September 27, 1989, Jiaozuo , China</p> <p>Current position, grades: College of agricultural engineering in Henan university of Science and technology.</p> <p>University studies: Agricultural Engineering in Henan university of Science and technology.</p> <p>Scientific interest: Agricultural Electrification and Automation.</p> <p>Publications: 1 Patents, 1 Papers</p>
	<p>Guoying Zhang, born on September 27, 1982, Nanyang Henan, China</p> <p>Current position, grades: Engineer, mainly engaged in the tracto quality supervision and inspection and measurement , as well as the measurement test/calibration work in Luoyang xiyuan vehicles and power institute co., LTD (National Tractor Quality Supervision And Inspection Centre).</p> <p>University studies: Master's degree in agricultural mechanization engineering, Henan university of science and technology.</p> <p>Scientific interest: The bionic robot application in mechanical engineering, and the application of the metrology in scientific research.</p> <p>Publications : 1 Patents, 5 Papers</p>

RSSI Enhanced indoor LBS platform design

Shang Zhang^{1,2}, Tingyan Xing^{1*}

¹ Faculty of Information Engineering, China University of Geosciences Beijing, China

² College of Computer and Information Technology, China Three Gorges University, Yichang, China,

Received 1 March 2014, www.tsi.lv

Abstract

The fast development of WSN (Wireless Sensor Network) provide the solution to indoor localization application. To which, the position accuracy become the essential problem need to solve. This paper introduces in detail the composition of the whole system and the design of localization algorithm based on RSSI. Through computing the relevance of adjacent nodes of target tag, the enhanced localization method is introduced and reasonable system design prove the possibility to build a high accuracy indoor localization system based on WSN.

Keywords: WSN, RSSI, indoor localization, LBS

1 Introduction

With the rapid advances in wireless telecommunication and portable devices technologies, the need for smart applications that could offer personalized services to the mobile users has attracted a lot of research interest in the past few years. This interest has led to the development of a range of services called "Location Based Services". Location Based Services (LBSs) provide personalized services to the subscribers based on their current position using Geographic Information System (GIS), spatial positioning such as Global Navigation Satellite System (GNSS) and Wireless Communication (WC) technologies. LBS offers modern world the tool for efficient management and continuous control. More and more people involve LBS in their industry and day to day life to better achieve their goals.

Nevertheless, most of the previous implementations of location based services focused on outdoor localization and outdoor services due to the widespread use of GPS systems.

Only a few indoor localization systems could locate users indoors relied on installing sensor networks and other options that increased the cost of system deployment. A WSN (Wireless Sensor Networks) is a network consisting of a large number of wireless radio nodes equipped with sensing devices, which have transceiver to communicate with another node within its communication radio range [1], and are densely distributed for specific localization applications.

Alternatively, LBS can also be grouped due to their thematic properties, e.g., mobility applications, entertainment applications, e-commerce applications, or emergency applications, these were the objective of the navigation application is to guide users (pedestrians or

drivers) towards a selected location. Accuracy is a key factor as the user must be able to select an address down to the number of the building within a street. Hence, a hybrid satellite-WN (wireless network) system is necessary to locate the terminal to the required precision, including vertical accuracy (user can be underneath a bridge, indoor, or in any high floor in a building) [2].

In this paper, we introduce an enhanced RSSI based algorithm to improve the position accuracy of WSN nodes, which implicated in a middle-ware based localization platform [3].

2 WSN indoor localization Technology

Wireless Sensor Network (WSN) is a newly emerging wireless network technology with short distance and low velocity, which is mainly applied in wireless connection with short distance and low velocity [4]. In WSN, randomly distributed minimal nodes that integrates sensors, data processing units and communication modules compose a network in the way of self-organization. Real-time sensing and sampling interesting phenomena cooperatively in peripheral environment are carried out and related information is processed with the aid of all types of sensors built in the nodes to obtain more detailed and precise information [5]. WSN can provide solutions for not only accurate indoor location but also higher precision than traditional satellite positioning. WSN senses and monitors real-time information for various environments and objects with considerable low-cost WSN nodes.

Currently, there are two major types of localization algorithms, namely, the localization algorithms based on distance measurement, and the localization algorithms not based on distance measurement. The localization

* Corresponding author e-mail: xtygis@163.com

algorithms based on distance measurement need to position the node according to the measured distance and angle between reference nodes. The least squares principle, the triangulation method and the maximum likelihood method are used to compute the best estimated position of non-reference nodes by this type of algorithms, e.g. AOA algorithm, TDOA algorithm and the RIPS algorithm. The localization algorithms not based on distance measurement estimate the position according to network connectivity and other information as well as the relative position between itself and the reference nodes, having no need to measure the distance. This type of algorithms includes centroid method, DV-HOP algorithm, Amorphous algorithm, APIT algorithm, etc. In recent years, a lot of experimental work about localization algorithms has been conducted by domestic and overseas experts and scholars, at the same time, some achievements are accomplished. Savvides et al proposed a theoretical model for node geometrical arrangement [5], which enhanced the accuracy of node localization and was able to get the corresponding results, but showed great errors in localization the nodes on the edge of the network. This is mainly caused by the condition that the reflection angles between nodes to each anchor node are small. The convex optimization algorithm proposed by Niculescu was an algorithm with no requirement of distance measurement [6]. This algorithm determines then position of the node through the convex programming of the anchor node. It has high position accuracy in determining node position, but with such disadvantages as high algorithm complexity, a large amount of computation and considerable network power consumption. The GFF algorithm proposed by Farid applied the idea similar to the distance-vector routing in DV-HOP algorithm to estimate the distance between nodes through the number of hops [6]. The disadvantages of this algorithm lie on that it requires high density of nodes, being prone to produce redundant nodes, and with significant measuring errors. Based on the abovementioned algorithm research ideology and the shortcomings presented, a node position control algorithm based on RSSI technology was proposed in this study. This algorithm used target parameter measurement quantification and relating geometric knowledge to compute the distance. Then iteration refinement was employed to get the precise solution of node Localization.

3 Problem defining and network model

3.1 INTRODUCTION

Physically, the Time of Signal (Acoustic or RF) Arrival (TOA) calculates the distance by use of signal propagation velocity and propagation time, Angle of Arrival (AOA) is measured by getting the signal direction send by the adjacent node through the combination of array antenna and multiple receivers, while Received Signal Strength Indicator (RSSI) measures received

power by receiving node, calculates propagation loss and transform propagation loss to distance by theoretical or empirical signal path loss model [1, 7].

In range-based schemes, the RSS-based techniques have been widely used, because the techniques are less expensive and simple to implement in hardware, and there is no need for additional hardware. RSS is defined as the voltage measured by a receiver's received signal strength indicator (RSSI) circuit. In fact, most of RF transceiver chips have a built-in RSS indicator, which provides RSS measurement without any additional cost. The strength of RSSI at a given transmission distance can be described as:

$$RSS(d) = PL(d_0) - 10\beta \log_{10} \left(\frac{d}{d_0} \right) + \omega, \quad (1)$$

where d is the transmission distance, d_0 is the reference distance, β is the rate at which the signal decays, ω is a zero-mean normally distributed random variable, $PL(d_0)$ is the received signal strength, $RSSI(d)$ is the received signal strength from the sender. The $RSSI$ can be estimated when the unknown nodes receive the RF signal from the anchor nodes, and the distance can be calculated via (1). Unfortunately, some studies showed large variability in RSS, because RSS is easily influenced by indoor environments [6].

Based on the advantages and disadvantages of the two types of localization methods, this paper proposed a highly precise RSSI-based indoor localization method using a transmission power adjustment strategy to acquire RSSI after reducing the effects of indoor environments. A variety of RSS patterns in the real indoor environment are also developed to increase the accuracy of the estimated distance between two nodes [6].

3.2 SYSTEM STRUCTURE

The whole system mainly consists of three parts, namely, card reader, tag, host and server. The Server based on Ubuntu Server 10.04 system and PostgreSQL/PostGIS as database to store the map of room and location coordinates. The calculation station is set up based on Ubuntu Desktop 12.04 with the compiler gcc and gdb. The geography programming is based on GDAL with GRASS (Geographic Resources Analysis Support System), which is a free Geographic Information System (GIS) software used for geospatial data management and analysis, image processing, graphics/maps production, spatial modelling, and visualization. GRASS GIS is currently used in academic and commercial settings around the world, as well as by many governmental agencies and environmental consulting companies [7].

When the system is working, each reader get to collect the RSSI values of the reference tags and pending tags, and then through the positioning algorithm to

calculate the undetermined tag position and display on the terminal in real-time [9].

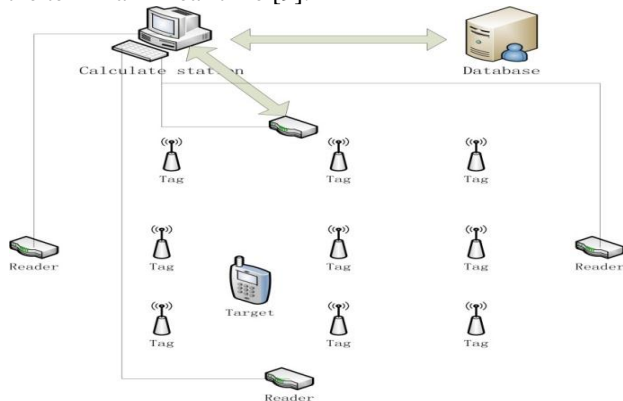


FIGURE 1 Indoor localization system

The whole positioning process can be divided into two stages, as shown in Figure 2 and Figure 3. The first stage is offline phase. Several reference tags are set up in room area, the position of reference tags can use equal grid to determine. On the other way, the distance among the reference tags should be carefully selected; too much interval will reduce the accuracy and too small will increase the load of database. 1 or 2 meters interval would be suitable.

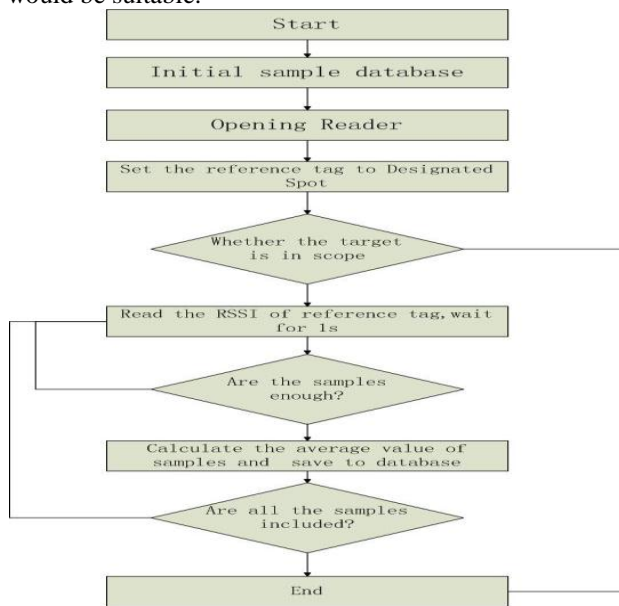


FIGURE 2 Offline phase

Phase two is the stage to accomplish the target localization. In this stage, the position tags contrast with samples in system, then choose the reference tags which has highest degree of association with the target tags to calculate the accurate distance.

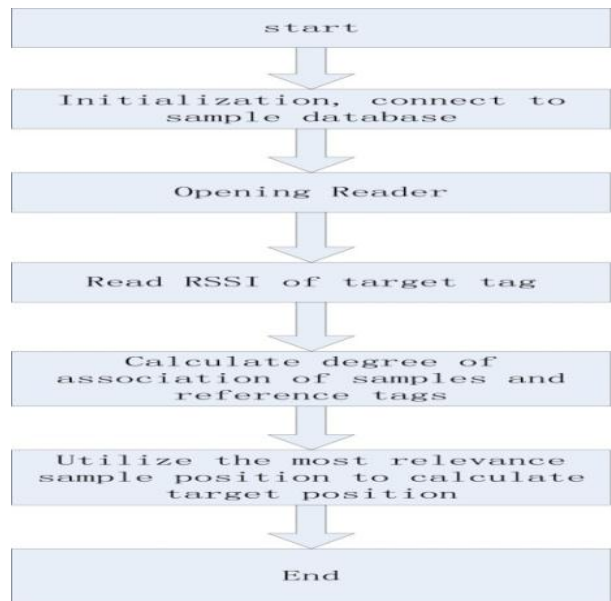


FIGURE 3 Online phase

Assume there are N pending tags, L reference tags, M readers. Then the Signal intensity matrix $S_{ij}(i=1,2,\dots,N; j=1,2,\dots,M)$ represent the j reader get the RSSI value of i reference tag, $H(i=1,2,\dots,L; j=1,2,\dots,M)$ represent j readers access the RSSI value of i pending tag. The Correlation matrices of reference tags and pending tags

$$[10]: E = \sqrt{\sum_{k=1}^M (H_{ik} - S_{jk})^2}, \quad (i=1,2,\dots,L; j=1,2,\dots,N)$$

represents the degree of association of the j reference tags and i pending tag, namely the Euclidean distance between them, If they are, the smaller the value of the correlation means that the closer the distance between the two labels.

Knowing the signal strength of the relevant reference tags of the k pending tags, the unknowing coordinates could be calculated:

$$(x, y) = \sum_{i=1}^k W_i(x_i, y_i) \tag{1}$$

W_i , represent the weight of some reference tag, the closer the tag is, the tag weights greater, k represent the number of the nearest reference tags, to some pending tag p, which is the smallest collection of value E which in $(E_{p_1}, E_{p_2}, \dots, E_{p_n})$ of k reference tags, in this system let k=4, through equation:

$$W_i = \frac{1/E_i^2}{\sum_{j=1}^k (1/E_j^2)} \tag{2}$$

calculate value of W_i , the position error $e = \sqrt{(x-x_0)^2 + (y-y_0)^2}$, among which (x_0, y_0) is the actual coordinates, (x, y) is the calculating coordinates. So, whether reduce the amount of reference tags or readers,

both the Maximum error and the mean error will increase [8].

4 Conclusions

Based on the analysis of advantages and disadvantages of various frequency bands and positioning algorithms, according to the indoor positioning existing NLOS phenomenon and the multipath propagation

characteristics of electromagnetic wave, combined with signal strength (RSS) and the experience of the corresponding relation between geometric space distance, using the existing hardware device by calculating the most adjacent reference label location to accomplish target localization. In practice, the system greatly reduces the development cost, accuracy and reliable, portability is strong, can be well applied to various actual situations.

References

[1] Zhang Weile, Yin Qinye, Wang Wenjie 2013 *Wireless Communications, IEEE Transactions* 12(2) 527-537

[2] . Garbin D A, Noblis J L F 2010 *IT Pro*, Open Source for Enterprise Geographic Information Systems 38-45

[3] Adewumi O G, Djouani K, Kurien A M 2013 RSSI based indoor and outdoor distance estimation for localization in WSN *In Industrial Technology (ICIT) 2013 IEEE International Conference* 1534-9



[4] Wang C M, Wang Z M, Zhu L Z 2007 *Geomatics & spacial information technology* 32 124-7

[5] Jiing-Yi W, et al 2012 High-Precision RSSI-based Indoor Localization Using a Transmission Power Adjustment Strategy for Wireless Sensor Networks *High Performance Computing and Communication & 2012 IEEE 9th International Conference on Embedded Software and Systems (HPCC-ICESSE), 2012 IEEE 14th International Conference* 487-94

[6] Xu Y, et al 2012 WSN Node Localization Algorithm Design Based on RSSI Technology *Intelligent Computation Technology and Automation (ICICTA2012 Fifth International Conference on)* 175 - 8.

[7] Zhang Shang, Xing Tingyan, Tao Liufeng, Zhang Fan 2013 *Proceedings of the 2013 International Conference on Software Engineering and Computer Science OSGIS & FOSS in LBS Design* 261 <http://dx.doi.org/10.2991/icsecs-13.2013.57>

[8] Zhang Shang, Xing Tingyan, Tao Liufeng 2014 *The 2014 International Conference on Information GIS and Resource Management(ICGRM2014)*, FOSS-based WSN indoor localization platform design 401-8

Authors	
	<p>Shang Zhang, born on June 8, 1979, HuBei, China</p> <p>Current position, grades: PhD candidate of Cartography and GIS, Lecture of China University of Three Gorges</p> <p>University studies: China University of Geoscience (Beijing)</p> <p>Scientific interest: GIS, Embedded system, WSN</p> <p>Publications: 8</p> <p>Experience: 2002-2011 Teaching in China University of Three Gorges in YiChang, China; 2004 - visiting scholars at University of British Columbia, Canada; 2011 - studying in China University of Geosciences(Beijing) for PhD</p>
	<p>Tingyan Xing, born in December, 1971, HuBei China</p> <p>Current position, grades: Professor ,Dean of of Information Engineering, China University of Geosciences Beijing, China</p> <p>University studies: China University of Geoscience (Beijing)</p> <p>Scientific interest: GIS, software engineering,</p> <p>Publications :90</p> <p>Experience: 1996- 2003 Teaching in Wuhan University</p> <p>2000-2003 studying in China university of geosciences (wuhan) and received a doctor's degree from the department of engineering</p> <p>2006-2007, visiting scholars and postdoctoral research at the university of Milan, Italy</p>

Study on semi-global matching algorithm extended for multi baseline matching and parallel processing method based on GPU

Zhaohua Liu*, Yuxia Yang

School of Architectural and surveying & mapping engineering, Jiangxi University of Science and Technology, Ganzhou 341000 China

Received 22 January 2014, www.tsi.lv

Abstract

This paper extended semi-global matching algorithm into multi baseline matching to improve matching reliability, especially studies kernel function optimization strategies and GPU threads' executing scheme of matching cost cube computing and aggregating, and realized its fine granularity parallel processing based on GPU. The experiment results using three UCD aerial images based on Tesla C2050 GPU showed that MVLL's semi-global optimize algorithm can improve matching effectiveness and productiveness.

Keywords: semi-global matching, multi baseline matching, dynamic programming algorithm, GPU, parallel processing

1 Introduction

Image matching is one key technology of generating DSM or DEM in photogrammetry. According to the matching strategy, it can be divided into local matching algorithm and global matching algorithm. The local matching algorithm is defined in a local window of a fixed or adaptive size, and the parallax of each pixel depends only on the pixels within the window. So its computation strategy is relatively simple, but its reliability is poor. Global matching algorithm takes full account of the compatibility, consistency and overall coordination of matching point and the points around, so it can get a more accurate parallax map through choosing a global energy function and minimizing the energy function. This strategy takes full account of the compatibility, consistency and overall coordination of matching point and the points around.

Common used global matching algorithm includes dynamic programming method, graph cut method and belief propagation method. Wherein the dynamic programming method is only carried out in the scan line direction, and can quickly achieve global optimization of the search with poor result. Graph cutting method and confidence propagation rules take full advantage of a two-dimensional constraint (horizontal direction, vertical direction) to obtain accurate dense parallax map, but the calculation is inefficient. Semi-Global Matching is an image-matching algorithm put forward by the German scholar Hirschmuller in 2005. This algorithm applies one-dimensional smoothing constraint in more than one direction to approximate a two-dimensional smoothing constraints, which cannot only get comparable results with graph cuts method and confidence spread, but also

achieve higher efficiency than these algorithms. The regular structure of the process is easily mapped to the GPU, DSP and SIMD parallel processing platform [1, 2].

2 Semi-global matching algorithm and multi-baseline extended application

2.1 THE BASIC PRINCIPLES OF SEMI-GLOBAL MATCHING

As other global matching algorithm, semi-global matching algorithm generally include: matching cost calculation, matching consideration polymerization, depth assignment and optimization [3-8].

Matching cost refers to difference of similarity between two corresponding points, and is usually measured by correlation coefficient, mutual information, etc. Under normal circumstances, disparity space image is used to describe the matching cost of corresponding points in scanning line. Thus a plurality of scanning lines constitute a matching cost cubes. There are two existent space parallax image-generating programs: one is constituted by the right and left scanning lines; another is constituted by the left scanning line and the parallax range.

As previously mentioned, the matching cost aggregation is generally achieved by solving the minimum energy function. The semi-global matching algorithm reinforces the reliability of matching through the imposition of additional smoothness constraint on the global energy function, as shown in Equation (1) below, where the first representation of all the matching pixels cost; the second and third represent the punishment of the pixel and its neighbourhood pixels difference is the

* *Corresponding author* e-mail: lhzhufang@163.com

presence of small change and a large change in both cases, namely the smoothness constraint, apparently $P_1 < P_2$. $T(\dots)$ refers to analysing function, if and only if its argument is true, the function value is 1, otherwise 0.

$$E(D) = \sum_p e(p, d) = \sum_p \left\{ \begin{array}{l} c(p, d) + \sum_{q \in N_p} P_1 T(|d - d_q| = 1) \\ + \sum_{q \in N_p} P_2 T(|d - d_q| > 1) \end{array} \right\}, \quad (1)$$

For the two-dimensional image, the global minimum for the formula (1) has been proven to be NP problem (Nondeterministic Polynomial), dynamic programming algorithm is introduced to efficiently achieve the one-dimensional path minimization of the energy. Therefore in semi-global matching algorithm, Hirschmuller utilizes 8 (or 16) one-dimensional smoothness constraint to approximately fit a two-dimensional smoothing constraint:

Firstly based on the idea of dynamic programming, compute on each path in accordance with the formulas (2) and (3):

$$L_r(p_0, d) = c(p_0, d).$$

$$L_r(p, d) = c(p, d) + \min \left\{ \begin{array}{l} L_r(p-r, d), \\ L_r(p-r, d \pm 1) + P_1 \\ \min_{i=d_{\min} \dots d_{\max}} L_r(p-r, i) + P_2 \end{array} \right\}, \quad (2)$$

$$- \min_{i=d_{\min} \dots d_{\max}} L_r(p-r, i),$$

$$L_r(p_0, d) = c(p_0, d). \quad (3)$$

The first item in (2) represents a matching cost of d endowed with p ; the second is a point on this path containing the minimum penalty coefficient matching cost; the third item of the doesn't influence the optimal path, solely for the purpose of preventing L excessively large, to make $L \leq C_{\max} + P_2$.

Then, the cost of the matching in each direction is formed by adding the total number of matching cost, such as in (4):

$$S(p, d) = \sum_r L_r(p, d). \quad (4)$$

After computing matching cost of all pixels, the depth assignment or disparity map is a simple process: each point on the reference image parallax corresponding matching cost minimum parallax value $d_p = \min_d S(p, d)$; each pixel in the reference image corresponding to the parallax $d_m = \min_d S(e_{mb}(q, d), d)$.

Although optimization is not focus of matching algorithm, but in some cases can significantly improve the quality of matches, commonly used optimization

method include: sub-pixel, consistency detection and median filtering.

2.2 MULTI-BASELINE EXPANSION METHOD

In multi-baseline matching, exterior orientation elements of the image are used as a link between the multiple images to undertake the matching process. But the parallax is no longer is the only parameter of the point of the same name in the search process, and the intersection point in the Object Space is only one [10]. So we must use elevation instead of parallax, namely (x, y, Z) instead of (x, y, d) . The corresponding match measure is rewritten as a function of elevation value. For a given image point in the reference image and its approximate elevation value, you can define all single stereo pair matching measure function into a single framework, such as: traditional dual-like the matching correlation coefficient match measure in the multi-baseline matching pattern is rewritten as SNCC, where the NCC match measure the correlation coefficient of the two images based on elevation variable [11]:

$$SNCC(p_0, Z) = \frac{1}{n} \sum_{i=1}^n NCC_i(p_0, Z), \quad (5)$$

$$NCC_i(p_0, Z) = \frac{\sum_{s \in W} (I_0(s) - \bar{I}_0) \times (I_i(s_i(Z)) - \bar{I}_i)}{\sqrt{\sum_{s \in W} (I_0(s) - \bar{I}_0)^2} \sqrt{\sum_{s \in W} (I_i(s_i(Z)) - \bar{I}_i)^2}}, \quad (6)$$

$$\bar{I}_0 = \frac{1}{M \times N} \sum_{s \in W} I_0(s), \bar{I}_i = \frac{1}{M \times N} \sum_{s \in W} I_i(s_i(Z)).$$

$s_i(Z)$ refers to the corresponding point in the i search image computed according to the exterior parameter and DSM.

To integrate the multi-baseline matching algorithm into the overall framework of the semi-global optimization, we should construct a matching cost cube according to the elevation search range, the length and width of the matching area. As shown in Figure 1, wherein each cube element indicates that, the matching cost of corresponding point on the multiple images.

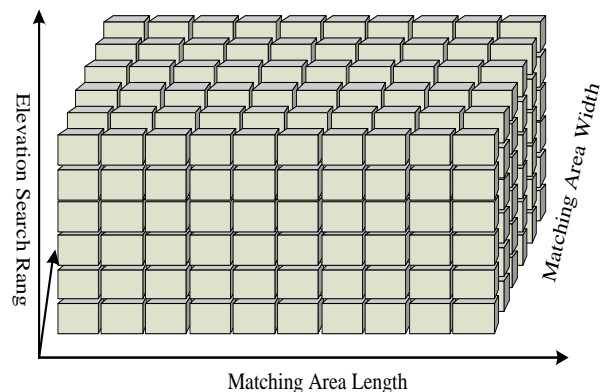


FIGURE 1 The matching area

After generating a matching cost cube, undertake multi-path optimization in the matching region of the X-coordinate direction and Y coordinate directions respectively based on the idea of semi-global optimization. The above method does not change the semi-global matching algorithm overall framework, so can take advantage of the GPU parallel processing.

3 GPU parallel processing technology

So-called GPU parallel processing, in essence, is the co-processing of the GPU-CPU, namely, to integrate GPU and CPU to form the synergistic mode on the hardware; In accordance with the Compute Unified Device Architecture defined (CUDA) programming model to achieve the synergy of GPU and CPU. The CPU is responsible for executing sequential code, while the GPU is responsible for intensive parallel computing, so that the CPU and GPU perform their duties to improve processing efficiency [9-10]. In the process of programming, CUDA allows the programmer to define a C-like language core (kernel) function to achieve the GPU parallel processing.

GPU parallel processing based on Semi-global matching algorithm focuses on compiling the "kernel" function of matching cost and the rational organization of GPU threads to run "kernel" function.

3.1 THE GPU PARALLEL COMPUTING OF MATCHING COST CUBE GENERATION

In the process of generating matching cost cube, each cube element calculation process is completely independent, with a high degree of parallelism, which is suitable for GPU grained parallel processing. The corresponding GPU threads organizational scheme: each thread is responsible for calculating points with the same plane coordinates and elevation coordinates in the search range; thread block and thread grid are divided in accordance with the two-dimensional plane coordinates of the matching region.

Taking into account that the calculation frequently reads the image data and the exterior orientation data, and if you do not optimize but directly read and write from the GPU global memory, will inevitably affect the GPU's performance due to the memory reading and writing latency. We consider these two data access characteristics and the memory structure of the GPU and use the shared memory and texture memory optimization respectively.

Each thread reads the same exterior orientation elements data in the calculation process, and the data is relatively small, so use shared memory for optimization, that is, a one-time elements of exterior orientation data read from the global memory to the high-speed shared memory all the threads within the entire thread block, to improve data access efficiency.

In the calculation process, the image data read by each thread is different. There will be some re-read, but after the projection transformation, the data's regular

structure is destroyed so it is difficult to meet the combined access conditions. Taking into account that the array structure of the original image is easily mapped into texture, use texture memory to save the data of the original image. The GPU threads use texture fetch function to read the original image data.

3.2 GPU PARALLEL COMPUTING OF MATCHING COST CUBE'S AGGREGATION

In the calculation process of Polymerization match cost cube, the calculation for each pixel is no longer independent, but need to use this match cost of a point, which makes the traditional block mode in accordance with the image matrix no longer applicable. However, can according to scan row (or column), as is shown in Figure 2.

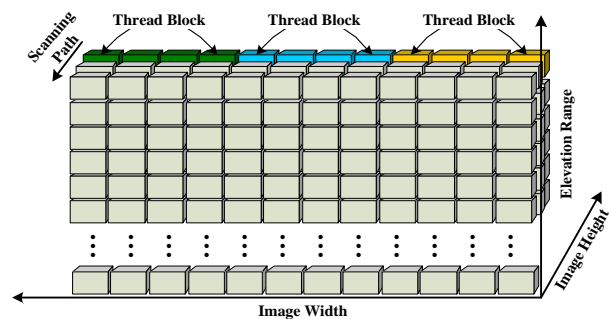


FIGURE 2 The Polymerization match

In the calculation of the three directions, each row of the image is divided into N segments, and $N = \text{ImageWidth}/\text{PixelAmount}$. Each segment consists of a $\text{PixelAmount} * \text{DisparityRange}$ thread blocks to be calculated. In the calculation process for each thread loop is executed ImageLength times to scan the entire image. ImageWidth represents the image width and ImageLength represents image height. PixelAmount represents the number of pixels for each thread block. DisparityRange represents elevation search range.

In addition, in order to further speed up the process, in a scanning process, simultaneously calculate matching cost polymerization of a plurality of directions.

4 Experiment and analysis

4.1 EXPERIMENT

In order to verify the advantages of GPU parallel processing, use large degree of overlap UCD area array image of Hubei Baoying to do multi-baseline MVLL match and GPU parallel processing experiments. The experimental images are shown in Figure 3, wherein (B) is the reference image, (A) and (C) as the search image, the image size is $7500 * 11500$ pixels, a pixel size of 9.0 microns, and a focal length of 101.4000 mm.



(A) search image 1 (B) reference image 1 (C) search image 2
FIGURE 3 Experiment image

Experimental platform is a personal desktop supercomputer developed by the United States AMAX PSC-2N, the specific configuration is as follows: two Intel E5620 2.40GHz CPU, $2 \times 8 = 16$ cores; $6 \times 4 = 24$ GB system memory; 2TB hard drive and three NVIDIA Tesla C2050 GPU parallel accelerators. The Tesla C2050 parallel Accelerator configuration parameters: memory capacity of 3.0GB, Memory Interface 384bit GPU processing unit number 448, the frequency of the GPU processing is 1.15GHz.

4.2 RESULTS AND ANALYSIS

First carry out an adjustment processing of POS data to obtain a precise exterior orientation elements of the image; then select the central portion of the reference image, and set the elevation of the search range from -5.0 m to 25.0 m, the search step 0.5 m, DSM grid interconnection interval of 0.5 meters, matching window of 7×7 pixels to carry out the half global optimization MVLL matching experiment. As shown in Figure 4, 800x800 size DSM data generated for the match, the building outline basically preserved. The experiment proves the effectiveness of the semi-global optimization of multi-baseline MVLL match algorithm.

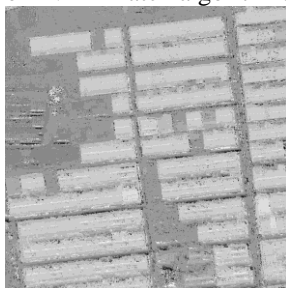


FIGURE 4 Matching result

References

- [1] Boykow Y, Veksler O, Zabih R 2001 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23(11) 1222-39
- [2] Sun J, Zheng N N, Shum H Y 2003 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25(8) 787-800
- [3] Hirschmuller H 2008 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30 328-341
- [4] Hirschmuller H, Scholten F, Hirzing G 2005 Stereo Vision Based Reconstruction of Huge Urban Areas from an Airborne Pushbroom Camera (HRSC) *The 27th DAGM Symposium, Vienna, Austria* 58-66
- [5] Hirschmuller H 2008 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30 328-41
- [6] Ernst I, Hirschmuller H 2008 *The 4th International Symposium on Visual Computing (ISVC08) USA* 228-39
- [7] Zhu K, Butenuth M, d'Angelo P 2010 Computational Optimized 3D Reconstruction System For Airborne Image Sequences *The international society for photogrammetry and remote sensing (ISPRS)* 2010), CANADA <http://www.isprs.org/proceedings/XXXVIII/part1/>
- [8] Heinrichs M, Rodehorst V, Hellwich O 2007 Efficient Semi-Global Matching for Trinocular Stereo *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences (PIA07)* 185-190
- [9] NVIDIA. *CUDA 2.0 for WINDOWS CUDA 2.0 Programming guide [EB/OL]* http://developer.download.nvidia.com/compute/cuda/2.0/doc/NVIDIA_CUDA_Programming_guide_2.0.pdf 2008-06-07
- [10] NVIDIA. *CUDA 2.0 for WINDOWS CUDA 2.0 Reference Manual [EB/OL]* http://developer.download.nvidia.com/compute/cuda/2.0/docs/CudaReferenceManual_2.0.pdf 2008-06-12

Make a statistic of CPU and GPU processing time of different matching points (unit: ms), and calculates the acceleration efficiency, as shown in Table 1. It can be seen: through GPU parallel processing computational efficiency can be greatly improved, which is beneficial to the popularization and application of the algorithm.

TABLE 1 CPU and GPU processing time

Matching points	CPU processing time	GPU processing time	Acceleration
200x200	305292	9139	33.4
400x400	1223372	19274	63.5
800x800	4896547	50397	97.2

5 Conclusions and prospect

This paper firstly analysed the basic principles of the semi-global matching algorithm, and then expanded it to the multi-baseline matching of the remote sensing image. Further improved the matching reliability, at the same time, retained the regular structure of the algorithm; then studied the GPU fine-grained parallel processing technology of the expanded algorithm. Focus on the nuclear function optimization strategies of the polymerization procedure and thread organization scheme. Finally, take advantage of Tesla C2050 GPU parallel accelerator card to process three UCD Aerial Images in semi-global optimization MVLL-multi-baseline-matching, the result proves the effectiveness and efficiency of the algorithm.

This paper is only a research of a single GPU, while the experimental platform PSC-2N personal supercomputer equipped with three Tesla C2050-GPU parallel accelerator cards. In the follow-up study, build the GPU cluster, and further develop the coarse-grained multi-block GPU paralleling between the accelerator cards.

Acknowledgement

The paper is funded by Science and Technology Plan Project of Education Department of Jiangxi Province (NO.Gjj11474).

Authors	
	<p>Zhaohua Liu, born on August 7, 1977, Shanxi, China</p> <p>Current position, grades: Master of Photogrammetry and Remote Sensing, Associate Professor and postgraduate supervisor in Jiangxi University of Science and Technology</p> <p>University studies: Jiangxi University of Science and Technology</p> <p>Scientific interests: Remote Sensing Image Processing</p> <p>Publications : 16 Papers</p> <p>Experience: an expert in Remote sensing image processing. In recent years, the study of parallel processing method based on GPU and Semi-Global matching algorithm extended for Multi Baseline matching.</p>
	<p>Yuxia Yang, born on September 3, 1986, Hebei, China</p> <p>Current position, grades: Post-graduate students of School of Architectural and survey & mapping engineering, Jiangxi University of Science and Technology</p> <p>University studies: Jiangxi University of Science and Technology</p> <p>Scientific interests: Remote Sensing Image Processing</p> <p>Publications: 3 papers</p> <p>Experience: She is a parallel algorithm for lovers. Her works are devoted to parallel processing based on GPU.</p>

Industrial product innovative design of toilet sensor timer

Rui-Lin Lin*

Department of Commercial Design, Chienkuo Technology University, No. 1, Chieh Shou N. Rd., Changhua City 500, Taiwan

Received 1 March 2014, www.tsi.lv

Abstract

This study came up with an innovative product design for toilet safety timer, which can be installed in public lavatories or toilets at home. When an elderly person is in a toilet for a period of time longer than the time set in the device, a warning will be sent to the outside. The unique feature is the safety device of the sensor timer with the innovative structure containing an infrared sensor, a CPU, a timer, and a warning device. Users can set up the time on the timer. In case of an emergency such as users falling by accident or being in a coma, a warning will be sent out in time to get help to prevent tragedy from happening.

Keywords: Toilet Safety, Timer, Industrial Product

1 Introduction

Toilets are essential in people's daily lives. With the advancement of time, life quality of humans has been improving. Therefore, the needs of the elderly in toilets have gradually been considered. For example, there are stools with armrests to help users to get up [1]. Moreover, there are alarms in some toilets for calling for help in case of an accident. Although designers have already made some improvements based on users' needs to make toilet environments better, according to newspapers and news reports, there have been still a lot elderly people stumbling or falling unconscious due to infirmity or illness and not found in time, leading to tragedies, which could have been avoided.

However, privacy of public lavatories and toilets at home is very important in order to maintain users' dignity. However, when elderly users use toilets, they are isolated after locking the doors. If they are tripped by accident or fall unconscious and cannot trigger the alarms, they may not be rescued in time. Furthermore, their families outside have no idea what happens inside the toilets and may worry about the elderly users having some accidents inside and not being able to be rescued in time without warning signals. These are all issues related to toilet environments, which need to be improved. Therefore, this study thought about how to develop an innovative structure, which can send warning to the outside of toilets to ensure users' safety while being inside.

2 Literature Review

Since 1993, the Taiwanese society has become an aging society defined to the UN. As age increases, problems caused by ageing would come one after another, and thus auxiliary tools would play an important role in the lives

of the elderly. People go to toilets every day. However, sitting on stools is a burden, which the elderly have to face every day. Therefore, Huang considered heights of stools and users' habits and simulated people's movements of standing up from sitting positions, for the analysis and development of an auxiliary mechanism. With the features of the slider crank mechanism, Huang designed an auxiliary device to help users stand up from toilet stools using the changes of sliding blocks and rods. According to the results of the study, if the height of the stool is increased by 81mm with the angle of 30°, the elderly can get up a lot more easily [9].

Some scholar improved toilet stools for elderly female people, in hopes of solving their problems with movements limited due to their knee osteoarthritis [2]. Cheng considered that degeneration of lower limb functions might make daily activities difficult and even increase the possibility of stumbling and falling due to their moving difficulties. Thus, Cheng objectively assessed the abilities of sitting and squatting of the elderly to learn more about their risks of tripping. After the investigation, it was found that the explosive force from the lower limbs of the elderly is rather weak. Their lower limb movements are slow but stable. The result is that when they perform a functional active, they do it slowly with small strength. However, to prevent one from tripping and falling, his lower limbs must be able to make efficient and quick reactions in a very short time. Therefore, those who may trip and fall are those with small strength and those who need a longer time to develop strength in lower limbs. The study found that explosive force trainings are good for the elderly to reduce their chances of tripping and falling [3]. Furthermore, the elderly with back pains may have problems with dynamic balance while walking and overstriding an obstacle [5].

*Corresponding author e-mail: linrl2002@gmail.com

Lee discussed the public lavatories in the train stations in Taiwan and found that they needed to be improved for the lack of sitting stools, unqualified equipment or barrier-free space, and insufficient software or greening [4]. In addition, Chen used a research population of public lavatories, which were categorized into railway station lavatories, bus station lavatories, pit stop lavatories, theatre lavatories, and university lavatories, and high school lavatories. A random sample of 662 lavatories from 114 places was used for investigation. The result shows that lavatories with toilets for the disabled still had the defect of having only squatting toilets. In addition, the commodity shelves, auto hand dryers, alarms, and broadcasting equipment were all not ideal. Chen suggested that there toilet papers should be provided and there should be armrests for the elderly and the weak [8, 11].

3 Innovative product design

3.1 PRODUCT DEVELOPMENT

The procedure of product development includes 6 phases: (1) planning: performing technology development and market target evaluation through the opportunities confirmed by the enterprise; (2) concept development: confirming the demands of the target market; (3) system-level design: defining product structure and dividing the product into systems and components; (4) detail design: confirming the geometric shapes, materials, and common differences of all the parts of the products; (5) testing and refinement: making several samples for testing, making improvements, and assessments; and (6) production ramp-up: training related staffs to solve the unsolved problems during the manufacturing process.

Product development is a cross-field activity. Almost all departments of an enterprise must participate in the enterprise's product development and cooperate. The key departments are marketing department, design department, and manufacturing department. An enterprise's marketing department is the bridge of interactions between the enterprise and customers. It helps to confirm product opportunities, define marketing segmentation, and confirm customers' demands. The design department is in charge of defining physical forms of products to achieve the highest customer satisfaction. The main job of the manufacturing department is to operate or coordinate production systems to manufacture products [6].

Product development is an enterprise's activity beginning with finding market opportunities, manufacturing, marketing, and ending with delivering products to customers. The indicators of whether product development is successful include: (1) product quality: reflecting whether a product satisfies customers' demands and whether the product is reliable; (2) product cost: cost of manufacturing a product; (3) development time: how much time the development team needs to complete the

development project; (4) development cost: the cost from the development procedure; and (5) development capability: whether the enterprise can develop a product in a more economic and efficient way [7].

3.2 MANUFACTURING

A manufacturing activity is a composition of inputs, transformation, and outputs. A manufacturing process includes all the activities of developing and manufacturing new products or providing services to customers. When developing and designing a new product, encountering difficulties is very common. If there is a problem with a product, the reasons behind it must be analysed with the person in charge and supporting staffs assigned, to come up with a solution and date of completion of improvements, in order to control the schedule and results. Before mass production of a product, units related to R&D, design, and engineering all must be very strict with reviewing and quality control.

The design quality evaluations for a product before mass production include: (1) Engineering Verification Test (EVT): evaluating the appearance, mechanical structure functions, part specifications, and applicability of the product; (2) Design Verification Test (DVT): evaluating the electric property specifications, mechanical functions, and environment reliability, and marginal test; (3) Manufacturing Verification Test (MVT): performing small-quantity pre-production to evaluate the suitability of product functions and production ability; and (4) Reliability Test: performing tests impacting or dropping the product, temperature cycling test, high/low temperature cycling test, and life test. In case of a defect or product failure during the evaluation, the Quality Department must notify the Engineering Department to analyse the reasons of the defect and provide a solution. After corresponding improvements are made, these tests must be performed again to confirm that the issue has been resolved [10].

4 Design results

The technical feature of this innovative product in solving the issue is to install the infrared sensor in a toilet stall and connected with a timer and a warning device which are placed near the door on the outside through the CPU. The CPU is connected to the infrared sensor and the timer. When the CPU receives a signal from the infrared sensor when a user goes into the toilet, the timer starts. Moreover, the warning device is connected to the timer. Once the time is up, the alarm is triggered. The alarm can be a warning flashing light or a buzzer. The connections between the timer, the warning device, the infrared sensor, and the CPU can be wireless or wired.

The procedure of how this product works is as follows: After the infrared sensor detects a user's body, it sends a signal to the CPU. Then the CPU transforms the signal from the infrared sensor and sends the transformed signal

to the timer. Moreover, the timer starts. If the user takes more time than the default time set in the timer, the timer sends a signal back to the CPU. The CPU transforms the signal from the timer and sends the transformed signal to the warning device. The warning device is activated. It rings or flashes to warn others. This product is composed of an infrared sensor, a CPU, a timer, and a warning device. It can send out a warning when a toilet user trips and falls by accident or falls into a coma, so that other people can rescue the user immediately and handle the situation. Figure 1 is a flow chart, which shows how signals travel between devices. Figure 2 & 3 shows the case that the user does spend more time than the set time so the safety device is not triggered by the timer and the warning device does not send out any warning. Figure 4 & 5 shows the case that the user spends more time than the set time so the safety device is triggered by the timer and the warning device sends out a warning. The utility model patent for this innovative product has been approved by the TIPO, with the patent number of M 445232 and the patent duration of 2013/1/11-2022/8/27.

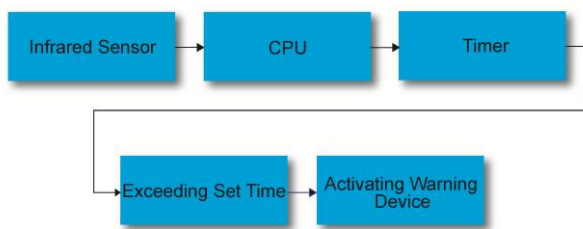


FIGURE 1 Flow Chart of How Signals Travel Between Devices

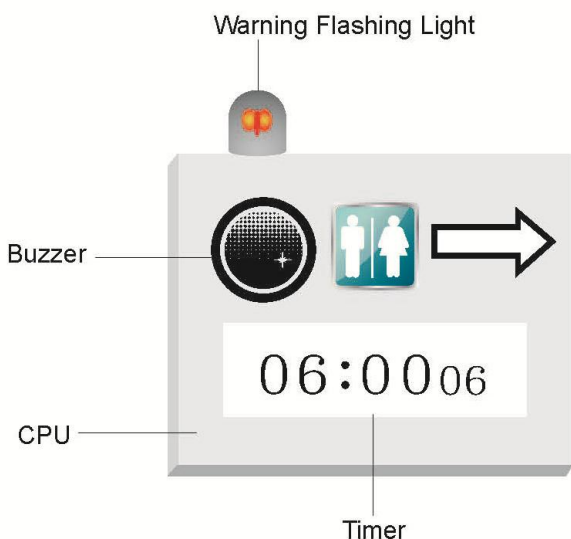


FIGURE 2 Normal Situation

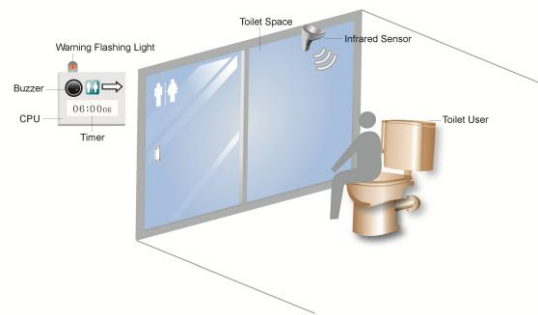


FIGURE 3 Normal Situation

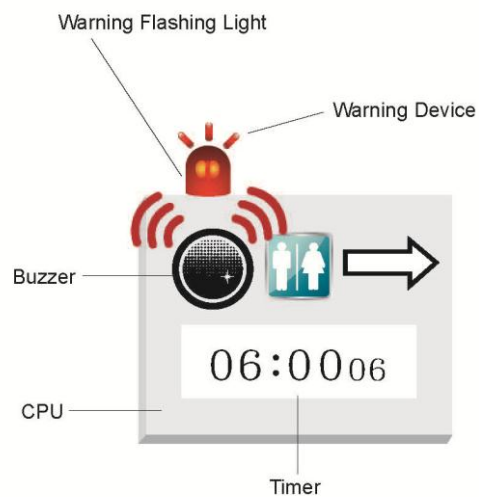


FIGURE 4 Emergency Situation

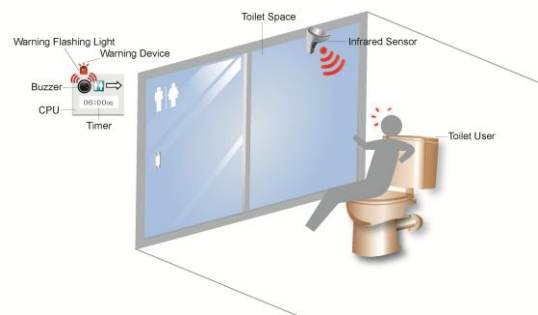


FIGURE 5 Emergency Situation

5 Conclusions

The first consideration of developing this innovative product, a toilet sensor time safety device, is that the elderly may trip and fall or fall into a coma due to illness or accident. In addition, the purpose is to provide the elderly a safer toilet environment.

5.1 DISCUSSIONS

- 1) *Designing with care for lives:* The innovative product of this study was developed based on the care for lives. After having considered users' needs, necessary safety designs were proposed to build a safe toilet

environment, show the care and respect for human lives.

- 2) *Confirming customers' demands*: This study observed and recorded the toilet behaviours of the elderly, found users' needs, made plans of important layers, and then proposed a strategy, which can meet customers' demands regarding toilets.
- 3) *Choosing target market*: In the face of the aging society, this study chose the needs of the elderly as the target market orientation to design an innovative product, in hopes of helping the elderly to use toilets more safely.
- 4) *Safe toilet environment*: It is hoped that, through the innovative design of this product, a safer toilet environment can be created, indirectly releasing caretakers' mental pressure.

5.2 SUGGESTIONS

- 1) *Innovative product design*: It is essential to carefully observe various people, events, and objects in daily lives and keep records. After proposing a solution or a strategy, it must be implemented for the purpose of accumulating new knowledge of technology and it

can be used as a reference for manufacturing and practical applications.

- 2) *Patent application*: When a good idea is generated, it is suggested to file a patent application in time to protect the patent right later when the product is developed and manufactured.
- 3) *International invention competition*: In the global innovative invention competition, Taiwan is only second to the US in competition results, showing the Taiwanese people's deep interest in innovative design and their potential. It is suggested to, after the patent application is approved, actively participate in international invention competitions to get more opportunities to promote the product and increase the possibility of mass production.
- 4) *Technology transfer*: During the model-testing phase of the innovative product, it is suggested to participate in more exhibitions for industrial exchanges to strive for more opportunities of technology transfer, in order to implement the good idea of the innovative product in daily lives, making people's lives safer, more convenient, and more comfortable.

References

- [1] Dekker D, Buzink S N, Molenbroek J F M, Bruin R D 2007 *Applied Ergonomics* 38(1) 109-18
- [2] Chou E C L, Yang P Y, Hsueh W H, Chang C H, Meng N H 2011 *Journal of Urology* 186(3) 949-53
- [3] Cheng I C 2007 Biomechanical Comparisons of Lower Limb Muscle Power among Healthy Elderly, Falling Elderly and Young Adults during Performing Sit-to-stand and Squatting movement *Master Thesis - Physical Therapy and Assistive Technology* National Yang-Ming University 1-2 (in Chinese)
- [4] Lee J Y 2012 A Study of Users' Reused Intention with the Public Toilet in Taiwan Railway Station-A Case Study of Taichung and Kaoshiung Transportation Services Segment *Master Thesis* Institute of Architecture: Feng Chia University 2-5 (in Chinese)
- [5] Chiang P T 2012 Gait Characteristics and Dynamic Balance during Walking and Obstacle Crossing in Late Middle-aged Persons with Chronic Low Back Pain *Master Thesis* Institute of Physical Therapy: National Cheng Kung University 2-4 (in Chinese)
- [6] Ramani S V, Ghazi S S, Duysters G 2012 *Technological Forecasting and Social Change* 79(4) 676-87
- [7] Zhang S W 2012 *Product Design and Development* McGraw-Hill: Taipei 2-44
- [8] Mamee W, Sahachaisaeree N 2010 *Procedia-Social and Behavioral Sciences* 5 1246-50
- [9] Huang W Y 2012 Research and Design of Toilet Assistive Device for Elders *Master Thesis* Industrial Technology R&D Master Program on Product Design for Excellence - Chaoyang University of Technology 1-10 (in Chinese)
- [10] Guo Y L 2012 *Manufacturing Practices*, Chwa: Taipei 7-56 (in Chinese)
- [11] Chen Y R 1996 Public Toilet Equipment Investigation and Improvement in Taiwan *Master Thesis*, Institute of Public Health: Kaohsiung Medical University 1-5 (in Chinese)

Author



Rui-Lin Lin, born on December 8, 1971, Changhua, Taiwan

Current position, grades: Associate Professor, Department of Commercial Design, Chienkuo Technology University

University studies: Doctor of Philosophy, Department of Industrial Education and Technology, National Changhua Normal University.

Scientific interests: Creative Thinking, Product Innovation Design

Publications: 21 Patents, 23 Papers

Experience: Served as an assistant professor at Department of Commercial Design, Chienkuo Technology University; Specialized in the related spheres and relative domains of designing education, hued scheme, products' layout, and designing market, etc.

Mechanism design and flow estimation method of a hydraulic actuated robot

Hongkai Li*, Zhendong Dai

*Institute of Bio-inspired Structure and Surface Engineering, Nanjing University of Aeronautics and Astronautics,
29 YuDao St., Nanjing 210016, China*

Received 1 March 2014, www.tsi.lv

Abstract

With the extension of robot applications, robot with high adaptability and high load capacity become a new focus in the recent years. Wheeled robots have the advantages of high load and speed, but this is limited in specified substrate. Legged robots inspired by the legged animals could move on rough terrain, so it was selected as a robot prototype for the high adaptability and high load capacity robot. In this paper, the structure of a hydraulic actuated quadruped robot was proposed. And then the kinematics of single leg was analysed. To estimate the required flow, a trot gait with 50% duty cycle is schemed and the trajectories of feet were planned. Then the total flow of the system required was calculated with the planned motion. The result can be taken as a reference to optimize the robot mechanism and select the hydraulic system.

Keywords: hydraulic actuated, quadruped robot, mechanism design, flow estimation

1 Introduction

The wheel is a remarkable invention as an important transportation in human history. Vehicles with wheels have the advantages in speed, efficiency and payload on flat road, but the application was greatly restricted due to the requirement of road [1], though it could be designed to achieve an arbitrarily small turning radius [2] or to negotiate extreme terrain [3] such as the planetary rover. Most of the terrain on earth is rough and irregular, and there are also many special circumstances that human cannot or difficult to achieve, such as the field of natural disasters and space exploration. So an effective and reliable mechanism is desired for search and rescue in the extreme environments. Moreover, legs offer an alternative for such application.

Legged mechanism is prior to the wheeled in terrain adaptability, obstacle negotiation and achievement of variable locomotion mode by using intermittent footholds [4]. Adopted by most of the animals on the terrain after millions of years' evolution, legged mechanism must have its natural superiority, and animals support excellent bionic prototypes for the design and control of robots. In the process of natural selection, Strong adaptability and effective motion rhythm mode of animals were established, which is available for the study of improve the performances of robots.

Quadruped robot needs the least feet for the static stability, and has the potential to realize variable motion type, such as trot, pace, jump and so on.

Quadruped robot greatly developed in the past several decades. Many of the existing robots are actuated by motors, and some unconventional actuators are also used for the study. Piezoelectric material have been used in the designs of small motion mechanisms [5, 6]. Energy efficiency was mainly concerned on such functional material actuator, and more attentions were paid to the performance of the materials. With the extension of applications, robots, which have the performance of high power density, high motivation and high payload, is desired to help people to do dangerous and boring work in real life. Hydraulic actuator has the characters of high power density and high power-to-weight ratio. Hydraulic system has been adopted to be the actuator for robots [7].

Hydraulically actuated quadruped robots have become another research focus due to its distinctive character [7, 8]. LS3 (Legged Squad Support Systems, LS3) is a large, dynamic quadruped, produced by Boston Dynamics. It is by far the most impressive robot designed to go anywhere Marines and Soldiers go on foot, helping carry their load. As the previous product of LS3, BigDog is their another amazing product [9, 10]. They are highly dynamic and have demonstrated recovery ability from significant perturbations, including being kicked or unexpectedly slipping on ice, using deliberate foot placement to control body attitude, although most of the details on their design and control remains unpublished. HyQ is another hydraulically actuated quadruped robot developed by Italy researchers [11, 12]. It was reported on the cyber

*Corresponding author e-mail: hkleee@nuaa.edu.cn

that HyQ has robust motion on the treadmill with desultory wood block on the surface.

2 Mechanism design of robot

The objective of this project is to develop a robot that can walk, trot and jump over rough terrain. The robot can balance dynamically and recover balance from unexpected disturbances such as rough terrain or external forces acting on robot. The weight is about 70kg including power and control system, and the payload capability is 50kg. The designed velocity is about 1m/s.

2.1 OVERVIEW OF THE ROBOT

Legs of robot are driven by hydraulic cylinders and controlled by servo valves. Motor and pump are needed as the power system to actuate the cylinders. The robot would be capable of highly dynamic locomotion as hydraulic actuation allows the handling of large payloads and impact forces, high bandwidth control, high power-to-weight ratio and superior robustness.

The CAD model rendering of the robot with a description of key parts and components is shown in Figure 1.

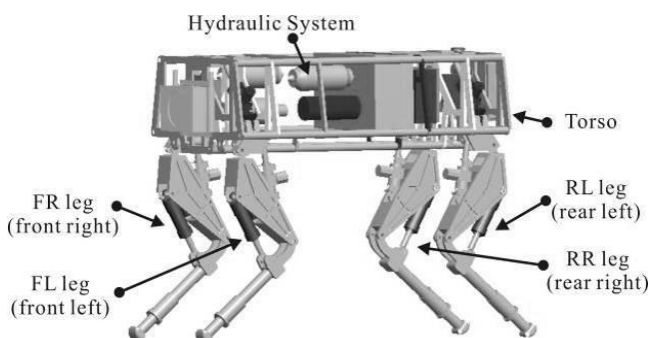


FIGURE 1 CAD model of hydraulically actuated robot

It is built up of a robot torso and four identical legs generally. Four legs are arranged on the torso symmetrically in the forward/backward configuration, where the front and hind knees point to each other. The torso is around 1.05m long and 0.56m wide to satisfy the space requirement of motion range of each leg and all the power support components. The total weight of robot including power system is about 72kg. Single leg excluding the parts embedded in the torso (the cylinder that controls the abduction/adduction of leg and the corresponding hip joint) is around 5.97kg.

The material of all mechanisms of robot is hard aluminium alloy, which have low density and high hardness. The torso is composed of aluminium tube and sheet metal. Moreover, the structure of leg mechanism is hollow. Leg and torso mechanisms are analysed by Finite Element Analysis software to determine the minimum thickness of tube and sheet metal that provide enough force and torsion.

2.2 LEG DESIGN

Three active revolution degrees of freedom (DoFs) are assigned to each leg to satisfy the requirement of Omni bearing motion ability. One controls the abduction and adduction of leg, and the others two control the flexion and extension. Three hydraulic cylinders are used to actuate the joints, which is the least requirement to allow a foot positioning in the three dimensional workspace round the hip. Leg mechanism is shown in Figure 2.

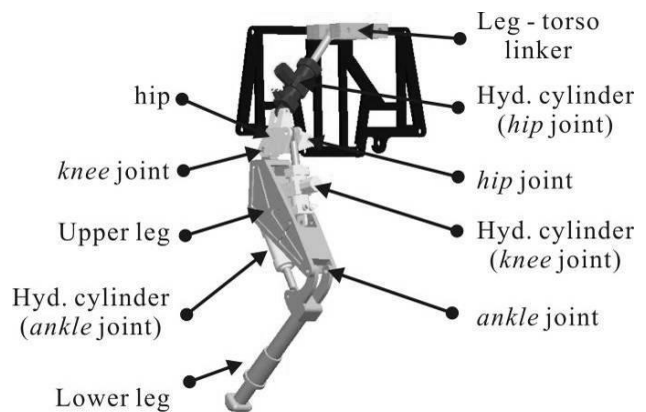


FIGURE 2 Mechanism of single leg.

Leg abducts or adducts around the hip joint axis parallel to the longitudinal axis of the body, responsible for the lateral motion of leg. Knee joint, located between hip and upper leg, is responsible for the upper leg motion that brings the knee closer/further from the torso. Ankle joint, located between upper and lower leg, is responsible for motion of lower leg that brings the foot upper/lower from the ground. Knee joint and ankle joint control the leg segments by flexion and extension along the rotation axis, and the axes of the two joint axes are parallel.

3 Kinematic analysis of single leg

The rotation joint of leg is actuated by the revolution of each hydraulic cylinder. Therefore, the relationship between the position of foot in workspace and the displacements of cylinders is necessary for the motion control of legs, which is also used for the further flow estimation.

According to the modified D-H method, the motion coordinate frame of single leg is built as shown in Figure.3. Take J_1 , J_2 and J_3 represent the hip, knee and ankle joints respectively. Σ_1 is built on the cross point of the axis of J_1 and the common perpendicular line between axes of J_1 and J_2 . In addition, the base coordinate system, fixed on the hip joint, is coincident with Σ_1 initially. Σ_2 and Σ_3 are placed on the rotation joints J_2 and J_3 . Σ_4 is placed at the centre of foot. Moreover, the specific position and orientation of each coordinate system are shown in Figure 3. Symbols c_1 , c_2 and c_3 represent the rotation angles of Z axis, and the arrows show the positive direction.

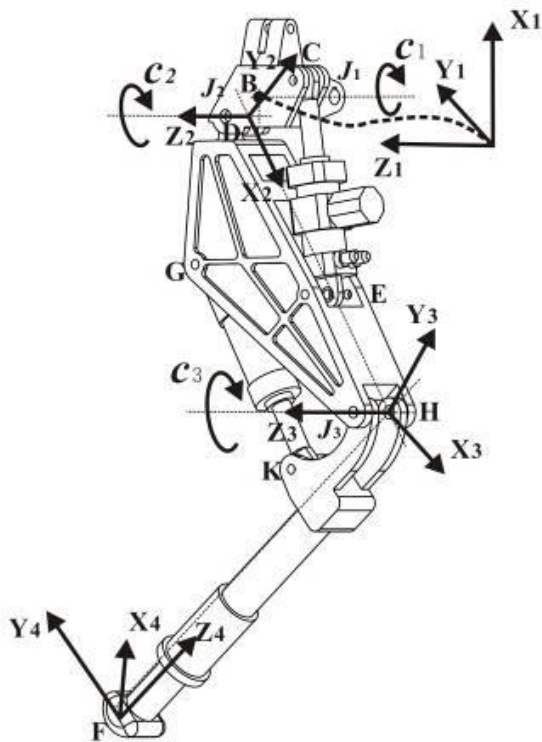


FIGURE 3 Kinematics coordinates frames.

The transform matrix between coordinates could be obtained according to Craig’s convention. So the homogeneous transform matrix from the base coordinates to the foot coordinates could be deduced.

According to the geometrical relation, the displacement of cylinder and joint angle could be represented as follow,

$$l_i(t) = f(c_i(t)), \tag{1}$$

where, f is the function between displacements of cylinder and joint angle. l and c are the displacements of cylinder and joint angle. i ($i = 1, 2, 3$) is the hip, knee and ankle joint, respectively.

Assumed that the p_x, p_y, p_z are the displacements of foot along axis direction in the base coordinate system. The relation between joint angles and foot displacements, represents as f could be deduced based on the kinematical analysis,

$$c_i(t) = g(p_x(t), p_y(t), p_z(t)). \tag{2}$$

So the displacements of each cylinder could be presented as the composite function of the foot displacement based on (1) and (2):

$$l_i(t) = f(g(p_x(t), p_y(t), p_z(t))). \tag{3}$$

The speed of cylinder can be deduced from (3),

$$\frac{dl_i}{dt} = \frac{\partial l_i}{\partial c_i} \left(\frac{\partial c_i}{\partial p_x} \frac{dp_x}{dt} + \frac{\partial c_i}{\partial p_y} \frac{dp_y}{dt} + \frac{\partial c_i}{\partial p_z} \frac{dp_z}{dt} \right), \tag{4}$$

Based on (4), the speed of each cylinder can be calculated if the trajectory of foot is given.

4 Trajectory plan of foot

The flow of hydraulic system is estimated when trotting in the air with the longest stride length in the workspace to obtained the minimum flow. The trajectories of robot foot are planned assumed that robot moves at constant speed without wagging at lateral and vertical directions. The trajectory along each direction is planned as below:

$$p_z(t) = \begin{cases} 45000 * (t - \frac{T}{20})^2 - 218 & 0 \leq t \leq \frac{T}{20} \\ -20 - 200 * \cos(0.4\pi T * (t - \frac{T}{20})) & \frac{T}{20} < t \leq \frac{T}{2.5} \\ -4500 * (t - 0.45T)^2 + 180 & \frac{T}{2.5} < t \leq \frac{T}{2} \\ 160 - 1800 * (t - \frac{T}{2}) & \frac{T}{2} < t \leq T \end{cases} \tag{5}$$

$$p_x(t) = \begin{cases} -702.5 - 42.5 * \cos(0.25\pi T) & 0 < t \leq \frac{T}{2} \\ -745 & \frac{T}{2} < t \leq T \end{cases}$$

$$p_y(t) = 0 \quad 0 \leq t \leq T$$

The trajectory of foot was planned relative to the hip joint. On swing phase ($0 - 0.5T, T=0.4s$) two quadratic functions and one trigonometric function were adopted to compose the trajectory along the moving direction (Z direction), and one trigonometric function was used along the up-down direction (X direction). On stance phase, foot moved at a constant speed along moving direction and kept the constant height at up-down direction. Shown in Figure 4 are the trajectories of front left foot. Feet could move with smooth velocity and acceleration based on the planned trajectory. In Figure 4, the step height is 85mm; step length of single leg is 400mm, the velocity is 1m/s.

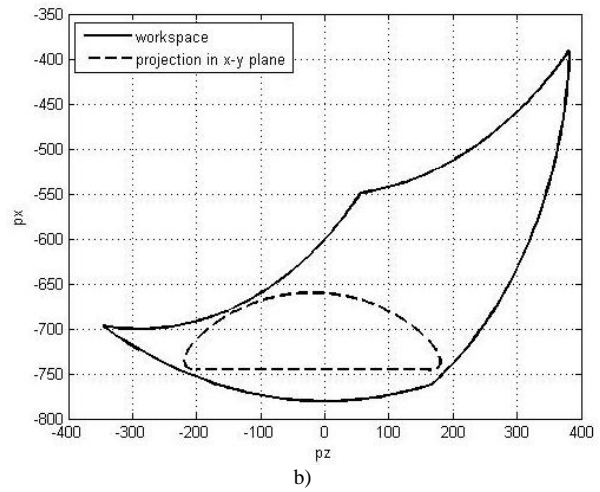
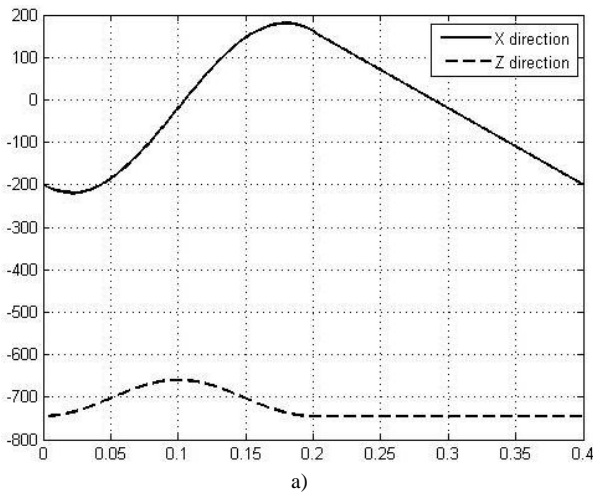


FIGURE 4 Planned trajectory of front left foot: a) is the trajectories at moving direction and up-down direction; b) is the project of workspace and trajectory in X-Z plane.

5 Flow estimation of hydraulic system

Estimated oil flow is an important parameter for the design of hydraulic system as it directly affects the size of the on-board pump, tank, cooler and pump engine, and therefore the total robot mass.

For an approximate analysis of the system, the leakage flow in the valve and cylinder and the compressibility flow in the lines and volumes can be neglected. The flow Q_{hdu} inside a hydraulic drive unit that consists of one cylinder can be expressed as follows:

$$Q_{hdu} = \begin{cases} v_{cyl} \frac{\pi D^2}{4\eta_v} & \text{if } v_{cyl} > 0 \\ |v_{cyl}| \frac{\pi(D^2 - d^2)}{4\eta_v} & \text{if } v_{cyl} < 0 \end{cases}, \quad (6)$$

where, v_{cyl} is the velocity of cylinder rod which equals to dl_i/dt . D is the diameter of cylinder barrel. And d is the diameter of cylinder rod. η_v is the volumetric efficiency, normally the value is 1.

The selected hydraulic is dissymmetrical. Oil flows into the cylinder barrel from the side without rod when $v_{cyl} > 0$, and from the other side with rod when $v_{cyl} < 0$.

For the estimation, it assumed that the hydraulic cylinder actuating the hip joint was fixed and the robot could move steady. So the total oil flow curve can be obtained by adding all the 8 cylinders at the same time

ticks. The curve in one cycle is shown in Figure 5. And the peak value of flow in about 76.4L/min.

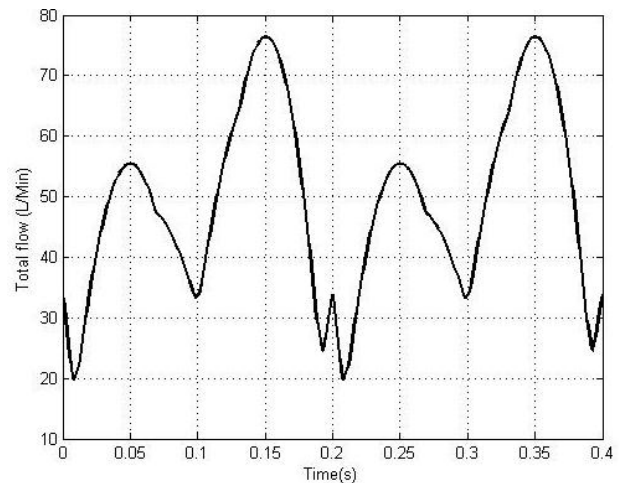


FIGURE 5 Total oil flow of robot in one cycle

6 Conclusions and discussion

The largest total oil flow is mainly decided by the absolute speed of all the motion cylinders at the same time if the cylinders are selected. Shown in Figure 6 are the speed curves of all the six motion cylinders.

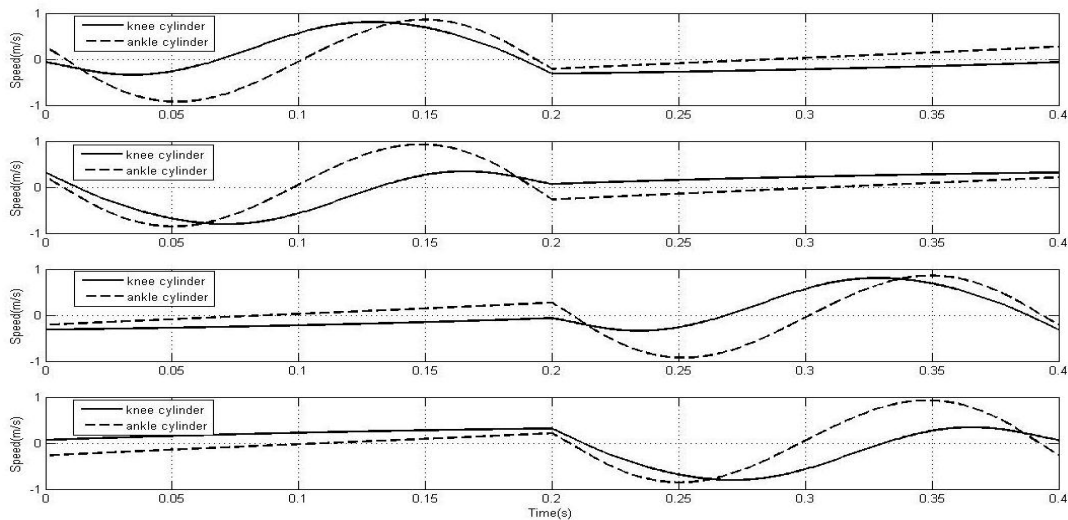


FIGURE 6 The speed curves of motion cylinders. The solid lines are the speeds of each knee cylinder, and the dashed are the speed of ankle cylinder.

Compared to cylinders speed curves with the total oil flow it could be found that the maximum value of flow is at 0.15s and 0.35s, which is corresponding with the time that the maximum sum of absolute cylinders speed. Therefore, the total flow will increase with the increase of cylinders speed. Moreover, the flow will increase if the motion speed robot increases by increase the step frequency. The total load of robot is improved at the cost of increase the total flow by keeping the cross points between cylinder and the relative leg.

In this paper, a hydraulic actuated quadruped robot prototype is designed. The method of estimating the total oil flow was deduced. At last, the oil flow was calculated, which offered a reference to the design of power system.

References

- [1] Michael L 1983 *Am Nat NLM* **121**(3) 395-408
- [2] Brett B, Jeremy S, Rybski P E, Veloso M 2005 *Ind Rob* **32**(2) 149-56
- [3] Karl I, Steven D 2004 *Int J Rob Res* **23**(10-11) 1029-40
- [4] Marc H R 1986 *Commun ACM* **29**(6) 499-514
- [5] Taesin H and Chong-Ho C 2007 *Rob Auton Syst.* **55** 795-810
- [6] Seok H, Tedy W, Hoon C P, Nam S G 2007 *J Bionic Eng* **4**(3) 151-58
- [7] Xuwen R, Yibin L, Jiahong R, Bin L 2012 *J Mech SCI Technol* **26**(4) 1171-7
- [8] Cai R B, Chen Y Z, Lang L, Wang J, Ma H X 2013 *Int J Adv Robot Syst* **10**(26) 1-8

Acknowledgements

We would like to acknowledge the suggestions of all colleagues during the design of robot and especially the technicians in Nanjing Industry of Mechanics-electronics-hydraulics in manufacturing, assembling and testing the hydraulic system.

This project is supported by National Natural Science Foundation of China (Grant No. 51205194, 61233014, 61170105) and the National Defence Pre-Research Foundation of China (Grant No. B2520110013).

- [9] Raibert M, Blankespoor K, Nelson G, Playter R 2008 *Proc. of the 17th World Congress the International Federation of Automatic Control* Seoul Korea 10822-25
- [10] Buehler M, Playter, R, Raibert M 2005 *Int Symp on Adaptive Motion of Animals and Machines* Ilmenau Germany 526
- [11] Boaventura T, Focchi M, Frigerio M, Buchli J, Semini C, Medrano-Cerda G A, Caldwell D G 2012 *IEEE Int. Conf on Intelligent Robots and Systems 2012 October 7-12* Vilamoura Algarve Portugal, 4066 -71
- [12] Shuang P, David T B, Emanuele G, Thiago B, Darwin G C 2012 *Proc of the 11th Biennial Conf. on Engineering Systems Design and Analysis 2012 July 2-4* Nantes, France 227-34

Authors



Hongkai Li, born on February 20, 1981, Henan, China

Current position, grades: Doctor of Mechanical Engineering, R.A., supervisor of supervisor of postgraduate in Nanjing University of Aeronautics and Astronautics.

University studies: College of Astronautics in Nanjing University of Aeronautics and Astronautics

Scientific interests: bio-inspired robotics, computer modelling, intelligent system, automatic control

Publications: 6 Patents, 12 Papers

Experience: B.S in Central South university in 2003, M.S. and Ph.D. in Nanjing University of Aeronautics and Astronautics in 2006 and 2010 respectively.



Zhendong Dai, born in 1962, Shanxi, China

Current position, grades: Doctor of Mechanical Engineering, professor, doctoral supervisor in Nanjing University of Aeronautics and Astronautics

University studies: College of Astronautics in Nanjing University of Aeronautics and Astronautics

Scientific interests: bio-inspired robotics, biomechanics, Tribology

Publications: 30 Patents, 160 papers

Experience: B.S., M.S. and Ph.D. in Nanjing University of Aeronautics and Astronautics in 1983, 1986 and 1999 respectively. Now he is the director of Institute for Bio-Inspired Structure and Surface Engineering.

Ground point filtering method of vehicle-borne laser point cloud in urban street

Maoyi Tian^{1,2}, Rufei Liu^{1*}, Xiushan Lu^{1,2}

¹ Geomatics College, Shandong University of Science and Technology, Qingdao, China

² Key Laboratory of Surveying and Mapping Technology on Island and Reef, National Administration of Surveying,

Mapping and Geoinformation, Qingdao, China

Received 1 January 2014, www.tsi.lv

Abstract

Through the analysis of the spatial characteristics of vehicle-borne laser point cloud data in urban street, a method to extract ground points accurately from point cloud data is proposed. Firstly, three-dimensional virtual grid is used to organize point cloud. Secondly, the initial low ground point in a grid is extracted by level plane constraint (LPC) method, and then a multi-scale neighbourhood analysis (MSNA) method is taken to optimize the low ground points further. Finally, the ground points from original point cloud data are filtered based on the local slope. The experiment shows that this method can effectively extract the ground points.

Keywords: Vehicle-borne Laser, Urban Street, Ground Point Filtering, MSNA

1 Introduction

With the development of city digitalization and informatization, obtaining the city three-dimensional spatial information by efficient and accurate ways has become a critical issue for the construction of city digitalization and informatization. As an advantaged method of measurement, vehicle-borne laser scanning is fast, non-contact, real-time, dynamic, proactive, high-density and high-precision etc. [1]. A vehicle-borne laser scanning system can collect a large area of high-precision and high-density surface information of objects in urban street, such as buildings, ground and vegetation etc. And high-precision ground elevation data can be extracted from the ground point cloud which provide basis data for further analysis of ground surface subsidence and damage. So that, the fast extraction of the ground point from vehicle-borne laser point cloud data has a vital significance.

In existent researching files, [2] showed that all the data points were projected in the grid, then the grid data were classified by the maximum height of each grid before and after projection, and some characteristic objects such as the ground and buildings were extracted etc. [3] proposed a vehicle-borne laser point cloud filtering method based on scanning line. In this method, the scanning lines were divided into different segments according to the slope difference, and then the different segments were classified by corresponding attribute. [4] put forward a method for point cloud classification based on characteristic of objects. The objects were classified on the basis of geometric features extraction of point cloud data in multiple streets and the summary of

characteristic objects knowledge. [5] proposed a filtering algorithm based on change of slope, the key of which was to select the appropriate slope threshold. He thought that slope thresholds should be chosen by prior knowledge of experimental area. The difficulty of the algorithm was increased because of the need for all the ground form samples. [6] introduced a region growing method into airborne point cloud filtering algorithm and obtained ideal classification by some certain conditions. [7] proposed a cluster analysis method based on point cloud spatial feature vector which classified sidewalk, pavement and curbstone by calculating the normal direction and characteristic value.

The current researches on vehicle-borne point cloud data are mostly for independent object extraction. When extracting the ground point, it is judged only on basis of simple elevation threshold without comprehensive consideration of terrain features in urban street. While there are a large amount of researches on the ground point filtering in airborne-borne point cloud data processing, the vehicle-borne point cloud data filtering cannot follow the airborne because of the different scanning range, targets, densities and precision. In this paper, through the study of features of urban street point cloud data, the author design a method for vehicle-borne laser point cloud data processing, which applies the ideas of three-dimensional virtual grid, MSNA and local slope filtering to extract ground points from original point cloud data.

*Corresponding author e-mail: liurufei_2007@126.com

2 Description of algorithm principle

2.1 THREE-DIMENSIONAL VIRTUAL GRID

The concept of three-dimensional virtual grid is introduced into point cloud data processing, which overcomes the low efficiency and information loss by other data organization operation [8]. Virtual grid diagram is shown in Figure 1, the dot represents the point cloud and rectangle block represents a virtual grid. While establishing a virtual grid, we need to set the appropriate scale of grid depending on the maximum slope and scanning point destiny, rather than according to the airborne scale. There are two ideal conditions:

- 1) The ground points are approximate horizontal within a single grid;
- 2) A single grid should have a certain amount of points.

The relationship formula between the step of urban street grid and the slope is as follows:

$$slope = \frac{\Delta h}{\sqrt{(\Delta x)^2 + (\Delta y)^2}} \tag{1}$$

In the formula, Δh is the ground elevation difference in grid(In the LPC model, the elevation difference between any two points should be less than the threshold value), Δx and Δy are respectively the step length of the direction of x and y . Assume that the maximum terrain slope is 5%, and the maximum elevation difference (Δh) is 0.1m when the grid is appropriate horizontal, then you can calculate the diagonal distance of a single grid is $\sqrt{2}$ m according to the formula, and determine the grid is 1m*1m. Therefore, the usual step length of urban street grid is always 0.5m—2m.

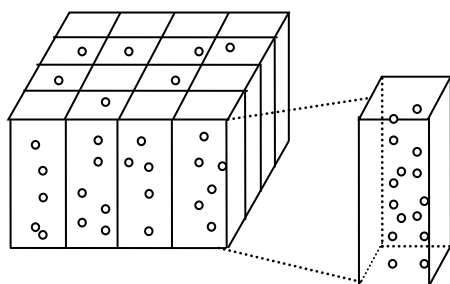


FIGURE 1 Three-dimensional virtual grid

2.2 NOISE POINTS REMOVAL BASED ON GRID STATISTICS

Due to the presence of random errors, in the processing of data acquired, noise points inevitably exist in original point cloud data, which need to be eliminated before filtering processing [9, 10]. The noise points' elevation is far higher or lower than the normal ones in a three-dimensional grid, which makes noise points distribute

discretely in space with others, as the red dot shown in Figure 2. Therefore, the number of points with different elevation interval layers in the grid could be analysed statistically, and the point is considered to be noise point when the number of points in the layer is less than a certain threshold value.

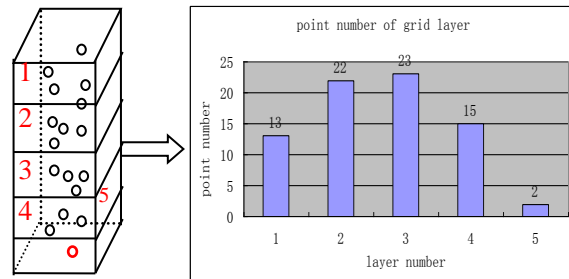


FIGURE 2 Principle diagram of grid statistical denoising

2.3 INITIAL GROUND POINT EXTRACTION OF ONE GRID

In the three-dimensional virtual grid, the lowest point is defined as the initial ground in each grid usually [11-12]. In this paper, the initial ground point is got by the method of LPC method, so as to improve the robustness of the lowest point in the grid. The principle is: The surface of the local area of urban street is usually approximate level. If a single grid contains ground point, there should be an approximate level surface plane, which requires at least three non-collinear points on ground. Three factors, which are point elevation, triangular plane shape and area need to be considered when selecting three points. All the points in the grid should be removed when there does not exist three points that meet the conditions. By this way, the probability of the low point as the ground point in a grid can be raised. As it is shown in Figure 3, the red dot is the lowest point of grid. If it is a ground point, the two blue ground points must exist and they can form an approximate level triangular plane.

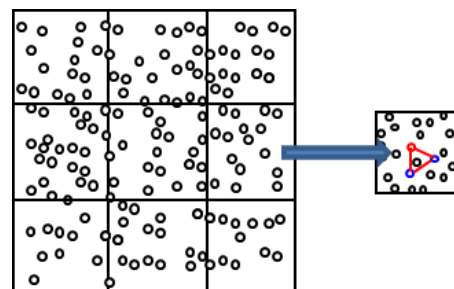


FIGURE 3 Low ground points by LPC model

2.4 MULTI-SCALE NEIGHBORHOOD ANALYSIS

The initial ground point with much more probability can be obtained by LPC. Due to the diversification of urban surface features, grid low point will be misinterpreted as ground point when the approximately parallel surface to the ground plane is scanned, such as the low and neat

vegetation surface. To solve this problem, this paper designs MSNA method to do the multiple filtering for target grid, further determining whether the low points within the grid are the ground points.

The main idea of the algorithm is: As it is shown in Figure 4a, firstly, comparing the lowest point of the unknown grid A with the lowest points in the 8-neighborhood grids as shown in Figure 5, calculating the slope by formula (1) in which Δh is elevation difference between two points, and Δx and Δy are the difference of x and y ; and then comparing the slope with the initial threshold. If the slope value between grid A and other grids are all less than the initial threshold value, the lowest point in grid A is considered to be a ground point, Otherwise, all points within grid A are non-ground points which should be removed, and that grid will not be calculated in the next step filtering. However, the grid A is difficult to be removed by the above method if the low points in A and other 8-neighborhood grids are all non-ground points and with similar elevation. So the important precondition of the 8-neighborhood analysis method is that at least one grid should contain ground points. From above analysis, we propose MSNA method. Firstly, using the same neighbourhood window to iterate the last results, and then expanding the neighbourhood analysis window size and repeating iterative processing until the window area is greater than the area of the largest non-ground approximate level plane. The algorithm is iterative processing until the difference between the numbers of reserved grid is zero. As it is shown in Figure 4b, the neighbourhood window is expanded from blue line to red line.

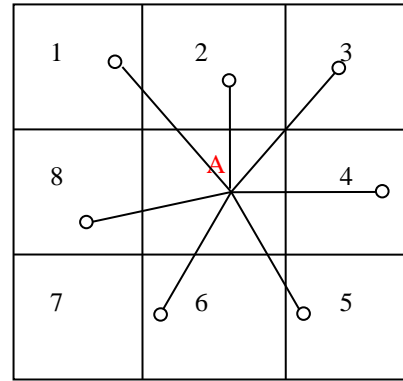


FIGURE 5 Principle of neighbourhood slope calculation

2.4 EXTRACTING GROUND POINTS BASED ON LOCAL SLOPE FILTERING

In order to improve the filtering accuracy of ground points cloud, range is limited within 8 neighbourhood window so that we can set lesser threshold during slope filtering. The method is described as follows. Firstly, the slope between unknown-point in grid A with low points in eight neighbourhoods is calculated. Due to the spatial relationship, the neighbourhood grid number of continuous topography cannot be less than two. If the grid number is less than two, the window size should be increased until it is less than 10m. If the number is also less two, the grid A should be judged as an isolated grid, which should be removed. Then taking the maximum slope value to compare with the slope threshold, those within the threshold value range are the surface points.

3 Experiment and analysis

In this paper, the above algorithm was implemented by using C++ programming language. In order to test the validity of the algorithm proposed in this paper, some typical area was selected for filtering experiment. The feature types include roads, building elevation, vegetation, cars and an uncompleted building frame in the test area. The parameters of original data are shown in Table 1.

Figure 6 shows a comparison before and after noise removal, in the red circle the noise points were successfully removed. Figure 7 shows the situation of inquiring the low point before and after MSNA, As can be seen from the figure that a large number of non-ground points are successfully removed after MSNA. Figure 8 shows the results of the ground point filtering; we can see that the point such as vegetation, traffic facilities and so on, are filtered exactly.

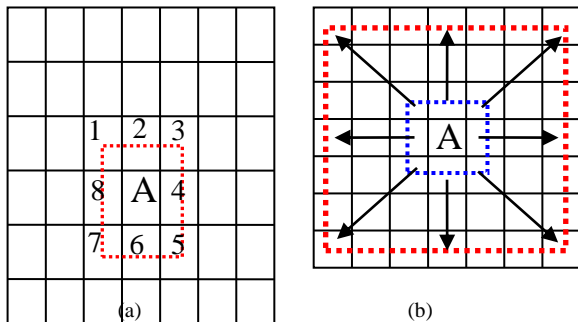


FIGURE 4 Principle of MSNA

TABLE 1 Experimental data parameters

Experimental Region	Point number	Area (m*m)	Average point distance(m)		Surface slope (%)
			Travel direction	Scanner direction	
Region 1	2892491	184*120	>0.11	>0.027	<4
Region 2	5017200	224*150	>0.06	>0.033	<5

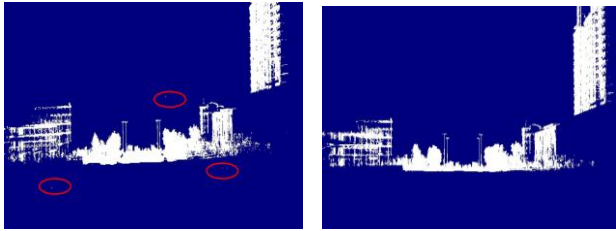


FIGURE 6 Comparison before and after noise removal

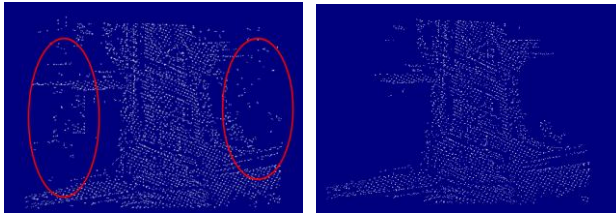


FIGURE 7 Comparison before and after MSNA

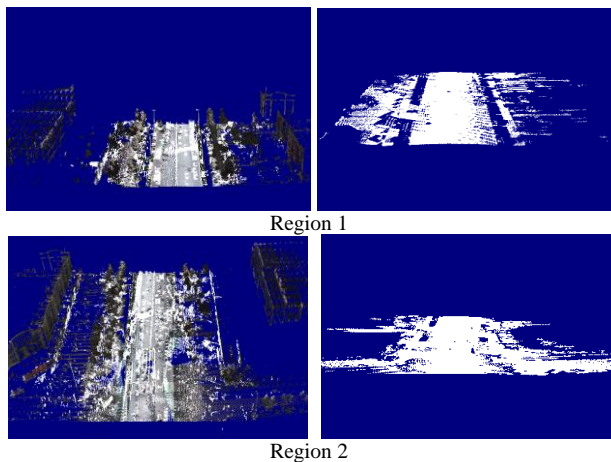


FIGURE 8 Ground filtering results of different region

The statistical results of algorithm processing are in Table 2. Grid horizontal step length and vertical direction are respectively 1m*1m*2m when removing noise points. According to the scanning point density and feature of space rod-shaped object, it is determined to be noise grid when the point's number in grid is less than 10 points, which need to be removed. In the experiment, 39582 noise points are removed in region 1 and 8685 noise points are removed in region 2. It can be known that the LPC and MSNA can effectively increase the probability of the initial low ground points and at the same time it can greatly reduce the grid number required for processing. 5690 non-ground grids of region 1 are removed by two step operation and 5882 non-ground grids of region 2 are removed.

4 Conclusions

A vehicle-borne laser scanning system can quickly acquire urban street ground point data with high accuracy. By analysing urban street spatial features of point cloud data, the filtering algorithm is proposed in several aspects, which include establishing three-dimensional virtual grid, removing point outliers and determining initial ground low points. Experimental results show that the filtering method of urban street ground points proposed by this paper is strong adaptability and high stability. In addition, the threshold parameter of algorithm requires some human experience and the adaptive threshold needs to be further researched.

TABLE 2 Statistics of data processing results

Experimental region	Noise removal (point number)		LPC (grid number)		MSNA (grid number)		Ground Point (point number)	Time (s)
	Before process	After process	Before process	After process	Iterative number	After process		
Region 1	2892491	2852909	7138	1620	3	1448	987174	45
Region 2	5017200	5015489	8090	2499	4	2208	1520717	82

Acknowledgments

This study is supported by National key scientific instrument and equipment development projects (2013YQ120343), Supported By the Key Laboratory of Surveying and Mapping Technology on Island and Reef, National Administration of Surveying, Mapping and Geoinformation, Special Project Fund of Taishan Scholars

References

[1] Shi Wen Zhong, Li Bi Jun, Li Qing-quan 2005 A Method for Segmentation of Range Image Captured by Vehicle-borne Laserscanning Based on the Density of Projected Points *Acta Geodaetica et Cartographica Sinica* 34(2) 95-100 (in Chinese)
 [2] Wu Fen Fang, Li Qing Quan, Xiong Qing 2007 On classification of vehicle-borne laser-scanning data *Science of Surveying And Mapping* 32(4) 75-77 (in Chinese)
 [3] Yang Yang, Zhang Yong-Sheng, Ma Y-i Wei, Yang Jing-Yu 2010 A Point Cloud Filtering Method of Vehicle-BorneLIDAR Based on

of Shandong Province, National Key Technology Research and Development Program of the Ministry of Science and Technology of China (2011BAK07B01),and Graduate Innovation Fund Project of Shandong University of Science and Technology(YC130103). The authors thank anonymous reviewers for their constructive comments and suggestions.

Scanning Beam *Journal of Geomatics Science and Technology* 27(3) 209-12 (in Chinese)
 [4] Li T, Zhan Q M, Yu Liang 2012 A Classification Method for Mobile Laser Scanning Data Based on Object Feature Extraction *Remote Sensing For Land & Resources* 117-21 (in Chinese)
 [5] Vosselman G 1999 Slope based filtering of laser altimetry data *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Amsterdam, XXXIII (B3)* 935-42
 [6] Abo A N, Zilberstein O, Doytsher Y 2003 Automatic DTM Extraction from Dense Raw Lidar Data in Urban Areas. *FIG wording week* Paris, France 13-7

- [7] Sherif El-Halawany, Adel M, Derek D L 2011 Detection of road curb from mobile terrestrial laser scannerpoint cloud *2011 ISPRS Calgary 2011 Workshop, Vol. XXXVIII-5/W12*, Calgary, Canada 29-31
- [8] Li Y Q, Sheng Y H, Liu H Y, Zhang K, Dai H Y 2008 3D road information extraction based on vehicle-borne laser scanning *Science of Surveying and Mapping* 33(4) 23-5 (in Chinese)
- [9] Zuo Z Q, Zhang Z X, Zhang J Q 2012 Noise removal algorithm of LIDAR point clouds based on three-dimensional finite-element analysis *Journal of Remote Sensing* 16(2) 303-9 (in Chinese)
- [10] Jaakkola A, Juha H, Hannu H 2008 *Sensors* 8 5238-49
- [11] Zhang H, Jia X M, Zhang Y S et al 2009 Filtering of Airborne LIDAR Data Based on Pseudo-Grid Concept and Modified Slope Filtering Algorithm *Journal of Geomatics Science and Technology*, 26(3) 224-7 (in Chinese)
- [12] Bisheng Y, Lina F, Jonathan L 2013 Semi-automated extraction and delineation of 3D roads of street scene from mobile laser scanning point clouds *ISPRS Journal of Photogrammetry and Remote Sensing* 79 80-93

Authors	
	<p>Maoyi Tian, born on October, 1976, Gaomi, China</p> <p>Current position, grades: Ph.D. and Associate Professor in Shandong University of Science and Technology University studies: Geodesy in Shandong University of Science and Technology Scientific interest: laser scanning data processing, image processing and 3S technology and application, digital city, digital mine and point cloud data processing Publications: 20 Papers Experience: He participated in or presided over a number of projects which are the national natural fund, the State Bureau of Surveying and mapping of basic surveying and mapping project, Doctoral Fund of Shandong province, the national science and technology support program key laboratory and other 8 vertical projects</p>
	<p>Rufei Liu, born on September, 1986, Nanjin, China</p> <p>Current position, grades: Ph.D candidate in Shandong University of Science and Technology University studies: Geodesy in Shandong University of Science and Technology Scientific interests: laser scanning data processing, aerial image processing and integration of 3S technology and application, digital city and point cloud data processing Publications: 10 Papers Experience: He participated in a number of projects, including National Natural Science Foundation of China, the State Bureau of Surveying and Mapping Surveying and mapping project, Doctoral Fund of Shandong Province and so on.</p>
	<p>Xiushan Lu, born on July, 1961, Suqian, China</p> <p>Current position, grades: Professor in Shandong University of Science and Technology University studies: Geodesy in Shandong University of Science and Technology, doctor of engineering degree in June 2000 and did Postdoctoral Research in May 2001 in Wuhan University Institute of Surveying and mapping Scientific interests: surveying data processing, GPS application and 3S integration Publications: 5 Patents, 60 Papers Experience: He is the first level discipline subject leaders of "Surveying and Mapping Science and technology"</p>

Development of automatic number plate recognition system

B Amirgaliyev^{1*}, M Kairanbay¹, Ch Kenshimov¹, D Yedilkhan²

¹ *ABY Applied Systems*

² *International Information Technologies University*

Received 1 March 2014, www.tsi.lv

Abstract

Today, the automatic number plate recognition (ANPR) system is a key aspect in traffic congestion. This will help minimizing the different kind of violations in the road. Advanced systems for tracking and fixing stolen, unauthorized vehicles are based on automated number plate recognition. This paper's main objectives is to review other methods and develop, at the same token evaluate our proposed approach. A very short review is performed on the various methods of number plate recognition systems. Further explanations of the proposed algorithm is illustrated in graphical forms to show how algorithm works. The paper is going to be concluded with test and evaluation results.

Keywords: ANPR, Plate area, Segmentation, OCR

1 Introduction

In the new global economy, traffic congestion has become a central issue for most of the developing countries. The number of cars are increasing rapidly; respectively the numbers of violations are increasing, too. Speeding, stealing the cars and other wide spectrum of violations in the road are general things in our daily life. Shortages of parking places are lead to entering of unauthorized cars to the private areas and spending a lot of time to find free places in the parking lot.

Automated number plate recognition system is a key aspect in resolving all the problems listed before. Adding other features to the system, we can identify and track the vehicle, fix the time and coordinates of appearance and disappearance of the cars.

Automated number plate recognition system consists of three main parts like number plate localization, number plate segmentation and optical character recognition.

2 Related work

There are various solutions of relevant problems. The main issues in number plate recognition are climate conditions, environmental interference, and accuracy of number plate localization. One of the methods of recognizing the number plate is utilizing the colour characteristics and probability distribution of the license plate between the two lights [4]. Another popular method of number plate recognition algorithm is template matching [2]. The License Plate Detection algorithm based on template matching was designed and written for managing the parking lot system by identifying the unregistered cars from off- campus. At the same time

vertical edges-based car license plate detection [3] are popular, too. However, others prefer to find the location number plate using horizontal and vertical projections of image [8]. The Genetic Algorithm [10] and Hough transform [6] can be applied to detect the license plate region.

Some of the methods above are very complex and requires too much computation time which is a bit difficult to use in real time applications. However, other approaches can be used only in specific countries with specific characteristics of number plate like background colour etc.

3 Research objectives

The following list gives the objectives of this research paper:

- 1) To solve and develop the automated number plate recognition system;
- 2) To evaluate, test developed system and presents the evaluation results.

4 Proposed approach

The whole problem consists of three parts:

- 1) Plate area detection;
- 2) Segmentation and extraction of characters from number plate;
- 3) Optical character recognition of extracted symbols.

4.1 PLATE AREA DETECTION

The image with number plate will be given as an input to the program and the number plate must be identified then

* *Corresponding author* e-mail: amirgaliyev@gmail.com

cropped as output image to the next stage. In order to determine the number plate from whole image, the following sequence of actions must be performed to image.

4.1.1 Grey scale image

In this stage, we need to read the image and convert it to grey scale format. Such conversion will not lead to loss of important data, at the same token it will be more

convenient to work with one channel in preference to three (red, green and blue).

4.1.2 Blur

The noise is a main issue in our problem. In order to reduce them it is better blur the image. There are different types of smoothing such as homogeneous, Gaussian, median as well as bilateral [5]. The following cumulative error distribution graph shows the comparison among each of them.

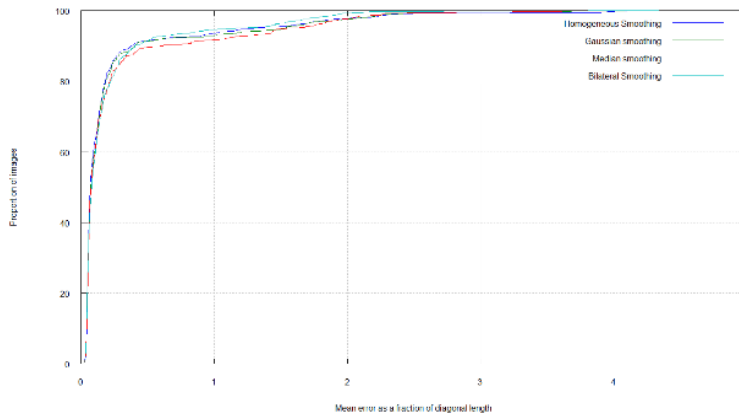


FIGURE 1 Cumulative error distribution for smoothing

According to the figure above, we can say that the homogeneous smoothing is the best one in comparison to others.

4.1.3 Vertical edge detection

Vertical edge detection - the number plate contains the characters. As we know, the characters contain mostly vertical edges in comparison to horizontal. That is why one of the best approaches is to find vertical edges that are too close to each other [1]. The edge detection is basic and fundamental operations in computer vision field. There are exist different kind of edge detectors like Prewitt, Sobel, Canny and etc. Each of them is used in different cases and problems. In our problem we use Prewitt, Sobel and modified version of Sobel [9]. However, after investigation and testing we came to conclusion to use the modified version of Sobel, because it correctly identifies vertical edges and reduce the most of horizontal edges that impede to find the number plate. Vertical edge detection can be implemented using convolution operation with specific matrix. For Prewitt and Sobel, the following matrices will be used.

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix} * A \quad G_y = \begin{bmatrix} +1 & +1 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} * A$$

FIGURE 2 Prewitt edge detector for vertical and horizontal edges

$$G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * A \quad G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A$$

FIGURE 3 Sobel edge detector for vertical and horizontal edges

where G_x , G_y derivative of image in X and Y directions respectively, A – is an input image. However, in order to use modified version of Sobel edge detection the gradient magnitude and gradient direction must be used. They are identified by the following formula:

$$|\nabla L| = \sqrt{L_x^2 + L_y^2} \quad \theta = \text{atan2}(L_y, L_x)$$

FIGURE 4 Gradient magnitude and direction

Using the value of θ it is possible to find only vertical edges. If the value of θ will be between 45 and 135 then we will get only vertical edges. The Figure 5 illustrate the result and difference of each edge detector.



FIGURE 5 The difference among vertical edge detectors

4.1.4 Binary Image

Binary image contains only two colours such as black and white. There are wide spectrum of methods for threshold like Otsu, Niblack, Souvola, Wolf, Feng and etc. [7]. Each of them used in special cases for different purposes. It is more convenient to work with binary image. After finding vertical edges, we will apply Otsu threshold to our current image.

4.1.5 Close morphology

Close morphology used mostly to combine close elements together. Such as our goal is to find the area of number plate, we do not need much information about characters. That is why we apply close morphology, where all letters and digits combined together. The Figure 6 illustrates the result of morphology operation.

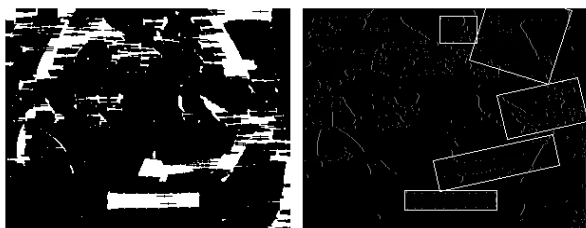


FIGURE 6 Image after morphology and finding contours

4.1.6 Find contours

Find contours. After applying the close morphology, we will find those contours that look like number plate, where the area and aspect ratio of contour must be taken into account. At the same time, the contours must be located horizontally like in the Figure 6.

4.1.7 Find correct candidate

Find correct candidate. In order to find the correct candidate among others, each number plate must be investigated. First of all, characters from number plate must be extracted. In order to do that, we need to take a threshold of image using nick method for determining the texture inside of each contour (Figure 7). After extraction of characters, each symbol will be recognized and appropriately, recognition probability will be returned for each character. Summing up all these probabilities for each symbol in number plate and taking into account the value that identifies how much the contours located in one line, we will get set of probabilities for each

candidate. When we already have set of probabilities for each candidate, we must to count the number of probabilities that are more than 0.9 (0.9 got from practical observation and testing). The candidate with maximum number of such probabilities will be considered as number plate.



FIGURE 7 Find contours that look like character

4.2 SEGMENTATION AND EXTRACTION OF CHARACTERS FROM NUMBER PLATE

From number plate characters must be extracted. There are two basis algorithms for segmentation, where first one is based on projection of image into X axis, however, the second one based on finding of contours that look like character. According to investigation and testing, we came to conclusion that second algorithms works better than first one. The Figure 8 shows the result of this algorithm.

4.3 OPTICAL CHARACTER RECOGNITION (OCR)

Extracted character must be recognized. For recognition, modified version of INN algorithm was used. The character was divided to small 49 subparts like in the Figure 9. For each subpart, the number of white pixels should be counted. The feature vector that contains 49 elements will identify each character. In the Figure 10 feature vectors for A, B and C classes were illustrated. For each class average element based on feature vectors should be calculated, then for unknown element the distance to all average elements of each class must be calculated. Unknown element will be joined to those neighbour classes that are closest to that element.

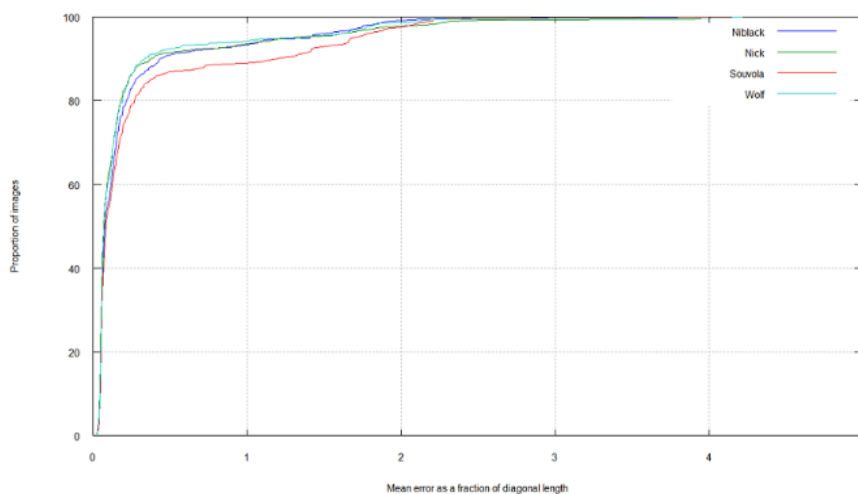


FIGURE 8 Threshold methods for identifying the texture in number plate

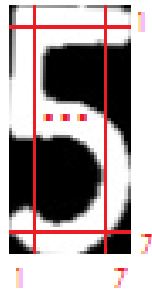


FIGURE 9 The character that divided to small (7 x 7) subparts

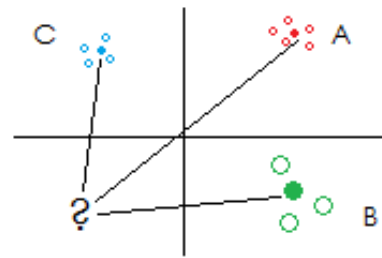


FIGURE 10 1NN for optical character recognition

The Figure 11 shows the result of whole algorithm.



FIGURE 11 The result of whole algorithm

$$(95.3\% + 93.15\% + 95.5\% + 94.46\% + 93.88\%) / 5 = 94.458\%$$

5.2 SEGMENTATION

The number plates were grouped based on their formats in order to segment. There were used different kinds of formats of number plate. The most popular are:

- (KZ) DDD LLL | DD (8 character);
- (KZ) DDD LL | DD (7 character);
- L DDD LLL (7 character);
- L DDD LL (6 character),

where D – is digit and L – is letter. The number plates above are new and old types of number plate in Kazakhstan, where (KZ) is a prefix and last two digits identify the region of Kazakhstan. As well as, there were used the number plates from other countries such as Russia, Kyrgyzstan etc. The following are their formats:

- L DDDD L (6 character)
- LLLL DDDD (8 character)
- LLL DDD (6 character)

5 Result and discussion

The program was tested with 1469 real car photos. The cars were taken from different sides and in different climate conditions. The Figure 12 illustrates the test cases that were used in testing stage.



FIGURE 12 Test cases

At the same token, we encounter with the some company cars that have their own number plate formats like L DDDDDD (7 character). Looking to the number plate formats above, we came to conclusion to decompose all number plates based on number of characters. In other words, we have three groups, where in first group - 6 characters (60 number plates), second group - 7 characters (976 number plates) and last group contains the number plates with 8 characters (410 number plates).

5.1 PLATE AREA DETECTION

The whole tests were divided to five subparts. Some subparts determines, from which side (front or rear) the photo was taken. The Table 1 shows the result of plate area detection using these test cases:

TABLE 1 The result of plate area detection Points sizes and types styles

Front	Rear	Subpart	Subpart 2	Subpart 3
95.3%	93.15%	95.5%	94.46%	93.88%

Taking the average value of results above, we will get the whole performance of plate area detection system:

TABLE 2 The result of segmentation

Types of number plate	6 character	7 character	8 character
Number of extracted characters	203 out of (60 * 6)	2553 out of (410 * 7)	5036 out of (976 * 8)
Percentage of extracted characters	56.38%	77.83%	64.49%

Taking into account the value for each group, we will find the total performance of segmentation algorithm $(56.38\% + 77.83\% + 64.49\%) / 3 = 66.23\%$. The Figure 13 shows the extracted letters after segmentation.



FIGURE 13 Extracted letters after segmentation

5.3 OPTICAL CHARACTER RECOGNITION

Some characters like “5 and S” and “O and 0” are looking similar. Taking into account this fact, we test our OCR solution. Our proposed algorithm correctly found 90 characters out of 100, or in other words, the performance of our solution is 90%.

6 Conclusion and future work

The main parts of number plate recognition system was successfully implemented. Our proposed solution works

References

- [1] Baggio D, Emami S, Escriva D, Ievgen K, Mahmood N, Saragih J, Shilkrot R 2012 *Mastering OpenCV with Practical Computer Vision Projects* Packt Publishing
- [2] Kroto H W, Fischer J E, Cox D E 1993 *The Fullerenes* Pergamon:Oxford
- [3] Benjapa R, Kittawee K, Paruhat P, Thaweesak Y 2012 License Plate Detection Based on Template Matching Algorithm *International Conference on Computer and Communication Technologies ICCCT'2012 May 26-27 Phuket*
- [4] Beverly S, Will H, Peter L, Patrico R 2012 Automatic Number Plate Recognition *Retrieved from CS 175 Fall '12*
- [5] Kuo-Ming H, Ching-Tang H 2010 A Real-Time Mobile Vehicle License Plate Detection and Recognition *Tamkang Journal of Science and Engineering* 13(4) 433-42
- [6] Bradski G, Kaehler A 2008 *Learning OpenCV O' Reilly Media Inc: 1005 Gravenstein Highway North, Sebastopol CA 9547*

for general cases, where there is no limit for the distance from camera to the vehicle and climate conditions. However, for specific problems, when the distance from camera to the vehicle will be constant the performance of our system will increase.

For the future work, we need to improve the segmentation part and gather more data for training. As well as optical character recognition can be improved using other popular algorithms like Artificial Neural Network and Markov chain.

Acknowledgments

This research study was funded by the World Bank, Ministry of Education and Science, Kazakhstan. This paper would not have been possible without the assistance, support and patience of our co-workers. Finally, I would like to thank our parents for giving birth to us and supporting us throughout our life.

- [7] Duan T., Hong Du T, Phuoc T, Hoang N 2005 Building an automatic vehicle license plate recognition system *Proc. Int. Conf. Comput. Sci. RIVF* 59–63
- [8] Khurram K, Imran S, Claudie F, Nicole V 2005 *Comparison of Niblack inspired Binarization methods for ancient documents* <http://www.ppgia.pucpr.br/~facon/Binarizacao/NiblackComparison.pdf>
- [9] Ondrej M 2007 *Algorithmic and Mathematical Principles of Automatic Number Plate Recognition Systems* BSc Thesis Bruno university of Technology
- [10] Wenjing, J, Xiangjian H, Huaifeng Z, Qiang W 2007 Combining Edge and Colour Information for Number Plate Detection *Proceeding of Image and Vision Computing New Zealand* 227-32
- [11] Yoshimori S, Mitsukura Y, Fukumi M, and Akamatsu N 2003 License plate detection using hereditary threshold determine method *Lecture Notes in Artificial Intelligence* 2773 ed V Palade, R J Howlett and L C Jain New York: Springer-Verlag 585–93

Authors	
	<p>Beibut Amirgaliyev, born in November 27, 1986, Kazakhstan, Almaty</p> <p>Current position, grades: Scientific director University studies: PhD Scientific interests: AI, Computer vision Publications: 17 Experience: 7 years</p>
	<p>Magzhan Kairanbay, born in October 26, 1991, Kazakhstan, Almaty</p> <p>Current position, grades: Junior research fellow in Computer Vision field University studies: Master, International IT University, Kazakhstan Scientific interest: Computer Vision, Machine Learning, Artificial Intelligence Publications: A Review And Evaluation of Shortest Path Algorithms, International Journal of Scientific & Technology Research Volume 2, Issue 6, June 2013 Experience: Junior research fellow in Computer Vision field in ABY Applied Systems, Development and analysis wide spectrum of projects related to Computer vision, Machine Learning.</p>
	<p>Chingiz Kenshimov, born in May 20, 1987, Kazakhstan, Taraz</p> <p>Current position, grades: Project manager University studies: Master Scientific interests: Machine Learning, AI, Computer Vision Experience: 7 years</p>
	<p>Didar Yedilkhan, born in May 30, 1990, Kazakhstan, Almaty</p> <p>Current position, grades: PhD student University studies: AI Scientific interests: AI, Machine Learning Publications: 3 Experience: 3 years</p>

Study on HDFS improvement scheme based on the GE code and dynamic replication strategy

Song Fei*, Cui Zhe

Chengdu Computer Institute of Chinese Academy of Science

Received 1 March 2014, www.tsi.lv

Abstract

There is a lot of valuable information in the massive amounts of data. Any loss of data may result in a great loss. Data security cannot be ignored. There are varieties of data disaster recovery technologies. However, most of these techniques depend on the hardware devices or data redundancy greatly. This paper presents a distributed data disaster recovery technology that minimum dependence on data redundancy and hardware system redundancy. In addition, this technology has nothing to do with the user equipment and application data structures. The test proved that this new data disaster recovery method can not only enhance disaster recovery capabilities and reduce the redundancy of the system greatly, but also suitable for large-scale distributed data disaster recovery.

Keywords: data disaster tolerance, HDFS, GE code, dynamic replication

1 Introduction

Now the data disaster recovery technology is divided into two categories: the first-class technology is closely related to the hardware devices and data structures. The first-class technology focused on the probability of damage to data equipment and the solution is based on a new type of memory devices and integrated device instead of the old data devices. The second type of study continued the system redundant technology roadmap (equipment redundancy and data redundancy). The second type of study is not very high to the quality of the storage devices, but the realization cost is closely related to the goal of the data disaster recovery system [1]. The lower target disaster recovery system only requires that the system can be able to provide the most basic data and services when the data was accidental damage. The higher goal of disaster recovery system requires no matter how much the price, the system can be able to recovery the lost data and interrupt service completely [2]. Because these two types of technical are exist the problems of higher redundancy for equipment and data, so, this paper presents a data disaster recovery method that based on encoding and dynamic replication strategy [1].

This new method based on computer reasoning [3] and data security storage [4] and it is a distributed technical method of disaster recovery storage [5]. This method is help to expand the scale of the disaster recovery system and enhance disaster recovery capabilities and greatly reduce system redundancy data. The disaster recovery capabilities of the new method mainly depend on the organization of data and the support of software method. It has a minimum dependence on data redundancy and hardware system redundancy and has nothing to do with the user

equipment and application data structures. This method is equally effective for the data disaster recovery that at the network distributed system-level and the level of memory storage units, and the larger the scale of the disaster recovery system, the better the effectiveness of disaster recovery. This technology is not a copy strategy of the data and it does not distinguish between work equipment and backup equipment, but with a minimum of redundancy cost to achieve the high efficiency of data disaster recovery and data recovery. Once a large area data devices disaster, this technology can be taken over the services and completed data recovery in a relatively short period of time and the cost of doing so just only software execution time of during data recovery. Therefore, this technology reduces the cost of storage space and equipment redundancy costs greatly [2].

2 GE Code

2.1 GE CODE TECHNOLOGY

Based on the low redundancy and high-performance data disaster recovery technology, this paper proposed a new type of array erasure coding technique that is E code family (including AE code and GE code). It expanded the data recovery performance and its range of applications for traditional array erasure codes [6-7]. The E code family has the significant advantages of powerful erasure and erasure parameters unlimited. Its operation is completely established bit arithmetic in finite field GF (2) and thus avoiding the difficulty of encoding is performed on a large-scale finite field [3].

In the E code, the information will be stored in an $n \times n$ data array. There are n blocks in a data array. Each data block is divided into n segments. Using segments $n-t, \dots,$

*Corresponding author e-mail: asfei@aliyun.com

n-2, n-1 as the check segment and the rest k=n-t segments as the information segments.

A specific n=16, k=12, t=4 data array instance shown in Figure 1.

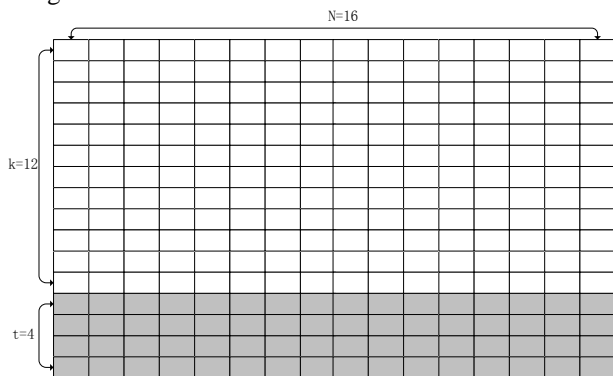


FIGURE 1 E code data array layout

As shown in Figure 1, E code is a parameter array code $[n, k(n-k)/2+1]$. This means that the length of E code is n and it can be able to accommodate k columns effective information and then we can use $t = n - k$ columns of parity information to get columns of fault tolerance ability.

Using U represents a storage unit and using n as the number of storage units. We agreed and damaged number of the storage units is here.

In order to clearly stated, we agreed that $t/n=q\%$. A typical case is $q=25$, then $t/n=q\%=1/4$.

GE code process is as follows.

- 1) There are n numbers of U and each U is divided into n blocks average. The number of y^{th} block in the x^{th} U is a_{yx} and $0 \leq x \leq n, 0 \leq y \leq n$. The before n-t blocks are blocks of information and that are used to store the effective information. After t blocks are parity blocks and which are used to store the parity information.
- 2) Construct a $[n-t, n-1]$ binary matrix and homogenizing the matrix so that the difference of the maximum line weight and the minimum row weight is less than 1, the difference of maximum column weight and minimum column is less than 1 also. Each line of the matrix has n 1 and each column has j 1. All the elements of the value of 1 in the coding matrix obtained is saved as a collection of tuples A.
- 3) Let a_{ij} as the element in the data array on j-th column i-th row, we can use matrix $A=[a_{ij}]$, $0 \leq i \leq n, 0 \leq j \leq n$ to express E code. In matrix A, the information elements with $[a_{ij}]_{0 \leq i \leq n-t, 0 \leq j \leq n}$ indicated and the parity elements indicated by $[a_{ij}]_{n-t \leq i \leq n, 0 \leq j \leq n}$.
- 4) These elements randomly divided into t collections, as follows:

$$D_0 : [d_{0,0}, d_{0,1}, \dots, d_{0,n-t-1}]$$

$$D_1 : [d_{1,0}, d_{1,1}, \dots, d_{1,n-t-1}]$$

.....

$$D_{t-1} : [d_{t-1,0}, d_{t-1,1}, \dots, d_{t-1,n-t-1}]$$

So, the parity elements can be generated by the formula

$$a_{i,j} = \bigoplus_{s=0}^{n-k-1} d_{i-(n-k),s} \quad n-k \leq i \leq n-1, 0 \leq j \leq n-1.$$

- 5) When a U goes wrong and need to restore the data stored within the U, we first need to do is marked the status of all the parity blocks in U that no error has occurred as "available".
- 6) Choose an "available" status parity block randomly and check whether the information block that verified by parity block is deleted. If there is no information block is removed, or only one information block is deleted, then the parity block is marked as "useless". The deleted information block can be restore according to the parity block and the XOR operation results of other information blocks which verified by the parity block. A recovery formula is:

$$d_{i',j'} = a_{i',j'} \bigoplus_{s=0}^{n-k-1} d_{i-(n-k),s} \quad i-(n-k) \neq i', s \neq j'$$

- 7) Repeat step 6, until all of the U, are no longer contains the wrong information blocks. At this time, all of the data stored in the broken U have restored.

2.2 DISASTER RECOVERY PROCESS

Suppose there are n storage devices, either t piece of equipment fails or is unable to obtain its data. Let $q\%=t/n$. If the system is able to fully recover data within the device failure t, the system can be considered a successful disaster recovery. At this time, the effective storage space of the system is not less than $(1-2*q\%)$ times of the total amount of n storage device storage space, usually $(1-q\%)$ times. If set $q=25$, then the effective storage space is n devices storage space total quantity 75% and the rest of the space for additional coding information.

Stored procedure as follow.

- 1) The application procedure will line up source files according to the FIFO order when the application receives the file storage request and then the source files will be deal with one by one. The application procedure will segmentation and coding each source file accordance with the GE coding scheme. Each source file is divided into n sub-file average and the bit lengths of each sub-file fragment are same.
- 2) The application procedure sent the n sub-file of the source file f separately to the different n servers on the network storage system accordance with the pre-set parameters n. The total amount of sub-files on n servers of the network storage system is less time than the amount of data in the source files (1.5 times the total amount of source data typically). At this time the stored procedure is completed.

Download and read process in the conventional case as follow.

- 1) When the application procedure receives a download request to read a source file f, it sends to a download request to the remaining n-1 servers in the network system and collects all the transferred sub-file fragmentations.

- 2) After all n fragments collected from the source file f (one local server fragment, $n-1$ offsite server fragments), local application directly recovery data of source file f from n debris and thereby completing the download/read of the source file f . Because the sub-file fragmentation is complete, so this conventional process read from the source file f without decoding time loss.

Download and reading process in exceptional circumstances (Disaster recovery mode) as follow.

Assuming that there are certain units (the number t) of data storage servers in the network system failure or unable to respond.

- 1) The application procedure will send a handshake information broadcast to other $n-1$ servers in the network system after it gotten the download/read request from the source file f . There are at least undamaged servers will return the handshake information.
- 2) Chooses servers stochastically from the undamaged servers as the partner servers of data recovery. In accordance with step 2, the source file f can be recovered and reconstructed.

Attention. There is no need to recover each sub-file fragmentation of source file f and not need to re-coding f too. We can complete the data stored profile of source file f just need to recovery the t error fragments and restore them to the t intact servers.

Different source files are independent of each other and the reconfigurable architecture recover decoding is independent of each other too. Therefore, if there are multiple files need to restoration and reconfiguration, we can assign these files to faultless servers and carry out restoration and reconfiguration using the mode of distributed and parallel processing. This method can speed up the speed of data recovery and improve the efficiency of the disaster recovery system [4].

3 HDFS improvement program based on coding and dynamic replication strategy

This paper presents a new and improved solution based on GE code and dynamic replication strategy - Noah (Not Only A Hadoop). This solution based on Hadoop platform architecture and improvements to the underlying file system of HDFS and to achieve a low redundancy and high-performance disaster recovery for the huge amounts of data.

3.1 HDFS EXISTING MULTI-COPY STRATEGIC ANALYSIS

The operation mode of HDFS cluster is "master-slave". The cluster contains two main types of nodes: Namenode

(master) and Datanode (slave). Namenode can only have one, but Datanode can have a plurality. The mode of data disaster recovery of the existing HDFS is the multi-copy technical. The specific approach is saving the three copies of the file blocks on the different DataNode of HDFS cluster respectively. The Namenode responsible for complete copy work, it uses regular round-robin fashion to receive the heartbeat of each DataNode in the cluster. If there has the heartbeat signal, which means that the DataNode is working, conversely indicates DataNode is not working properly. If the Namenode cannot receive a DataNode heartbeat, Namenode will release the blocks of DataNode to the other nodes. Namenode use this method to keep the number of copies [5].

The main role of multiple-copy in HDFS as follow.

- 1) Fault tolerance: to ensure system reliability.
- 2) Load balancing: The size of the DataNode load is determined by how much of the data it has.

Therefore, HDFS average distributes the data to each DataNode to achieve the load balancing. During the running process, HDFS may transfer the load by the way of moving copies.

3.2 THE DESIGN IDEAS FOR NOAH

In summary, the main reason for HDFS system has higher storage cost and lower load balancing capability is multi-copy strategy with a fixed number of copies. Therefore, this paper tries to provide a more flexible load balancing solution, which not only ensure data security but also reduce the cost of storage. This program design ideas is based on (1). Use code-based disaster recovery technology to replace the multiple-copy disaster recovery (2). Abandon the old fixed number of copy strategies, using GE code-based disaster recovery mechanism to realize the strategy that change the number of copies dynamically [6].

4 Noah disaster recovery technology realizations

Noah disaster recovery program abandoned HDFS mirror copy policy and using coding fragment solutions to re-encoded the data blocks in HDFS. This way makes the whole system cluster save only the code section that corresponding with the data block. Although the mapping table and the name space is maintained in the Namenode, but the actual data storage is based on code section and the data read and disaster recovery is also in the Namenode which operation on the code section that corresponding to the data block. Noah using dynamic replication strategy instead of a fixed copy of the original HDFS strategy makes the whole system to keep the load balancing in a good state. The architecture of Noah is shown in Figure 2 [7].

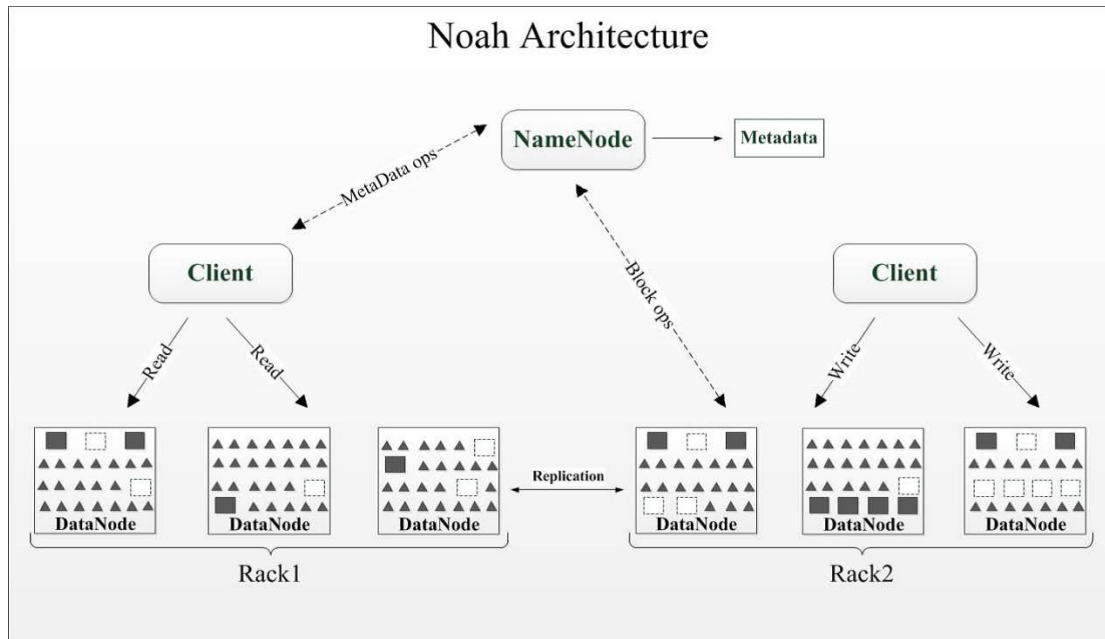


FIGURE 2 Noah architecture

4.1 DYNAMIC REPLICATION STRATEGY

4.1.2 Dynamic copies arrangement

According to the different needs of users for data and the system in support of the original the Hadoop file management mechanism, the dynamic replication strategy can make the system to maintain the load balancing in a good state by dynamic generation, delete and deploy the copies. For the hot data, dynamic replication strategy consists of the following two parts:

4.1.1 Dynamic replication generation strategy

Dynamic replica generation strategy has changed the way of saving a fixed number copies, it dynamically change the number of copies to achieve overall system load balancing. It is based on the needs of runtime system to decide the number of the copies for each data blocks. When generating copies, NameNode will be maintenance the number of the access-waiting queue for each data block that in the upper application. The maintenance principles are:

- 1) If the number of copies of a data block can be meet the application requests, the application requests do not need to wait in the queue, it can access to the copies directly.
- 2) If the number of copies of a data block can not be meet the application requests, the new application requests are entered into the data block's waiting queue, the system will generating several new copies for the application requests.
- 3) If the number of copies of a data block far beyond the number of application requests, it will produce a larger system consumes and will lead to a serious waste of storage resources. So, in order to improve utilization of storage resources, the system will delete the extra copies.

As mentioned earlier, the heartbeat is used to detect a DataNode working properly. Therefore, the DataNode notification the NameNode the system resource loads using the characteristic of the heartbeat. When the NameNode received the information from DataNode, the NameNode will judge the situation for each DataNode, and then the NameNode will move the higher load data block into the lower load data block and record the new position information of the transferred data blocks. Therefore, dynamic copies arrangement strategy is using this load transfer method to achieve a system load balance [8].

4.2 DATA RECOVERY STRATEGIES

The data recovery strategy of HDFS is replica the replication. This program presents a method that collect and decoding the section to achieve damaged data block recovering.

The section collection complete by the NameNode. NameNode first found the section that belongs damaged data blocks and then arbitrarily choose 70% of these sections put into a DataNode to decoding and data recovery operations. In the data recovery process, there will appear two situations:

- a) Code section is lost or damaged. In this case, NameNode will record the information of the coding segment until the quantity of missing or corrupted code section exceeds a threshold (generally 30%). NameNode will notify the coding segment locations to the DataNode that own these coding segments. DataNode will decoding and parsing these coding segments and then distribute these re-encoded code section to different DataNode. So as to ensure the

number of code section maintained at a safe threshold value [9].

- b) When the data of a hot spot data blocks are damaged or lost, NameNode will find the coding segments that corresponding this hot spot data blocks and then sent the coding segments to the specified DataNode. The DataNode decoding of these coding segments and thereby restoring damaged or lost data blocks, ultimately achieve the data recovery.

5 Experiments and conclusions

Experiments goal. For a distributed data storage system that consisted by N ($3 < N < 1200$) data equipment (Such as memory cells, memory or a server), we need to check that when any of r data devices have corrupted ($r < N/4$), whether the remaining $N - r$ undamaged data devices can recover all the data and if the total redundancy of system equipment is less than 1:1.

The experimental results. In this study, we used 35 sets of equipment as servers and according to GE code divide the data into 35 parts. Each part of data size ranging from 1M to 600M. Then we imitate the data corruption status, shut down seven sets arbitrary and using the remaining 28 sets of equipment to restore the corrupted data.

The experimental results showed that: when the storage system of any r ($r < N / 4$) data device was damaged, all data can be immediately and automatically recover by the remaining $N - r$ undamaged data devices and the total redundancy of the system storage equipment less than 1:1.

Figure 3 is a data disaster recovery resource consumption comparison chart for HDFS and Noah.

From figure 3, we can see that compared with HDFS 1:1 backup replication, Noah saves about 30% of the resources. And we can see the larger the backed up data the greater the Noah saved resources. Therefore, we can infer that Noah is more suitable for large-scale distributed data disaster recovery.

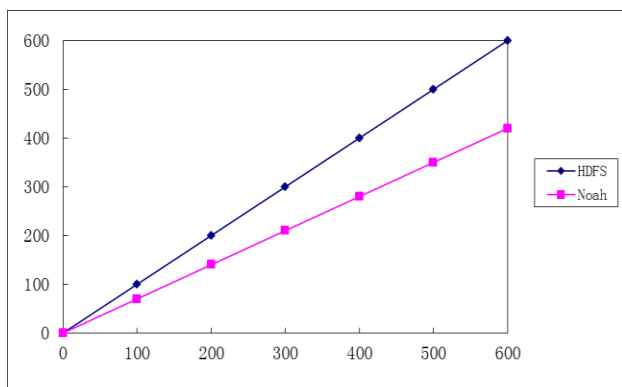


FIGURE 3 Resource occupancy comparison chart for Data disaster recover

Figure 4 and Figure 5 provide the compare for HDFS node load and Noah load node.

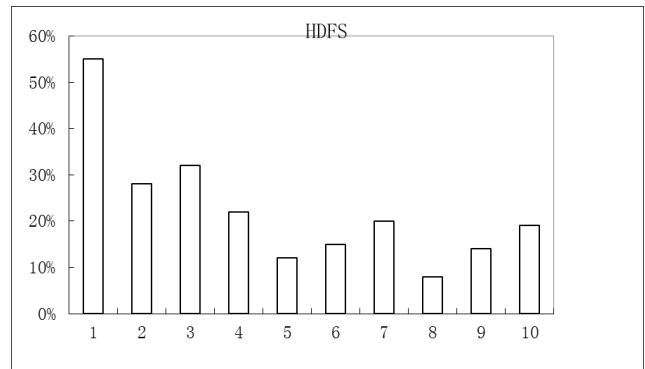


Figure 4 HDFS node load diagram

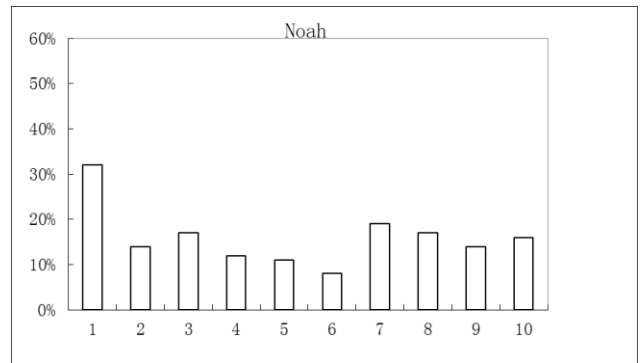


FIGURE 5 Noah node load diagram

From Figure 4 and Figure 5 we can see that the original HDFS load extremely uneven and fluctuation is larger, but Noah use dynamic replication strategy and allocates the load of a single node to each node in entire server cluster. and ultimately causing each node resource utilization more evenly and achieve load balancing.

6 Conclusions

The improvement program that proposed by this paper relative to original HDFS file storage system has following advantages.

This program lead encoding technology into HDFS file system and use an innovative way to replacing the original multi-copy disaster recovery technology. Original HDFS file system needs 1:1 redundant space, but this program just need no more than 70% redundant space.

This program adopts dynamic replication strategy and flexibility to change the number of copies replaces a fixed number of copies. It provides a more efficient system server load balancing ability and makes the distribution of resources of the entire file system more reasonable, more rapid and smooth running.

Because this program using the coding techniques, the users need not retrieve all the data when they requesting data and they can using the decoding method to restore the full data. When the parts nodes in the storage server failure or network congestion, this program has higher practicality.

References

- [1] Benzmueller Ch, Sorge V, Jamnik M, Kerber M 2008 *Journal of Applied Logic* 6(3) 318–42
- [2] Shuang K, Wang C, Su S 2010 A Novel Disaster Recovery Strategy in NGN core network based on P2P Technologies *Computational and Information Sciences* 601-4
- [3] Fei Song, Wang Xiao-Jing, Zhe Cui 2012 A trust model of P2P in cloud computing environment *Advances in Mechatronics and Control Engineering* 1962-5
- [4] Li Mingqing, Shu Jiwu, Zheng Weimin 2009 GRID Codes: Strip-Based Erasure Codes with High Fault Tolerance for Storage Systems *ACM Transactions on Storage* 4(4) 15
- [5] Dimakis A G 2011 A survey on network codes for distributed storage *Proceedings of the IEEE* 99.3 476-89
- [6] Porter G 2010 *ACM SIGOPS Operating Systems Review archive* 44(2) 41-6
- [7] Sethia P, Karlapalem Kr 2011 *Engineering Applications of Artificial Intelligence* 24(7) 1120–7
- [8] Kambatla K, Pathak A, Pucha H 2009 Towards optimizing hadoop provisioning in the cloud *Proc of the First Workshop on Hot Topics in Cloud Computing* 118
- [9] Shvachko K, Kuang H, Radia S 2010 The hadoop distributed file system *Mass Storage Systems and Technologies* 1-10

Authors



Fei Song, born in October 31, 1982, Chengdu, Sichuan, China

Current position, grades: Doctor studies

University studies: Chengdu Computer Institute of Chinese Academy of Science

Scientific interest: Data Disaster Recovery, Reliability Engineering and Network coding

Publications: 4

Experience: Song Fei is a doctoral student in Chengdu Institute of Computer Application, Chinese Academy of Sciences, China, has published more than four articles in reputed international journals and International Conferences, research interests are in the areas of Data Disaster Recovery, Reliability Engineering and Network coding.



Zhe Cui, born in September 20, 1970, Chengdu, Sichuan, China

Current position, grades: Professor

University studies: Chengdu Computer Institute of Chinese Academy of Science

Scientific interest: Trusted Computing, Embedded Systems and Reliability Engineering

Publications :15

Experience: Cui Zhe is currently a professor in Chengdu Institute of Computer Application, Chinese Academy of Sciences, China. His research interests lie in Trusted Computing, Embedded Systems and Reliability Engineering

Design a media art installation based on fuzzy controlling system

Zheng Wang^{1*}, Zhenjiang Miao²

¹ School of Architecture and Design, Beijing Jiaotong University, Shangyuancun. 3, Beijing, China

² School of Computer and Information Technology, Beijing Jiaotong University, Shangyuancun. 3, Beijing, China

Received 11 March 2014, www.tsi.lv

Abstract

As an art installation showed at Houtan station of Shanghai Metro Line7 and 2010 Shanghai World Expo Museum, "smart suspension ball" system displayed the rational sense of form and order of the controlling. The article focused on how to use appropriate fuzzy strategy to make the movements of the art installation more accurate under detailed experimental data. Another point of the article is to consider how to make this art installation to be a product with network, being combined and modular after upgrading the hardware and software of the installation. The performance of this upgraded product will bring more beautiful visual effects of controlling and technique. It will also be a successful case of integrating of science and technology into product designing for development of creative industries in China.

Keywords: Digital Media Art, Fuzzy Controlling, Installation Design, Arduino

1 Introduction

In recent years, a number of digital art works more or less related with "controlling" theory showing in many international or domestic art exhibitions attracted much attention. The visual presentation of the works in common is to control a large number of simple geometry objects moved regularly and performed in a real space. Audiences may feel a strong sense of order spatial transformation from the work. The performed form of the art works implies a particular technology beauty. This kind of visual effect gives audiences an extremely rational aesthetic enjoy.

From a technical analysis, most of these art works used stepper motors to control objects in space for precise displacement. Special computer programs were also used to control multiple real objects shaping different visual effects. After recently years developing, this kind of "controlling" art works gradually becomes a popular trend in the creating of digital art. It not only subverts the concept that only monitor, projectors could belong to the final output media, but also enriches the ways of digital art works showing. Therefore, more and more digital interactive artists prefer to use such means of expression in their art works.

2 Creative idea and system components

As one of the representatives of digital art works showing rational aesthetics and order of controlling, "smart suspension ball" was originally designed for participating the 2009 Liverpool Biennial. The art installation was

planned to place in the Mersey (a canal in Liverpool) side. When audiences went into the display area, the spherical objects in the glass tubes were triggered to move up and down by computer-controlled fans. The shapes made up by spherical objects symbolize waves of the ancient canal and cargo ships shipping in the busy waterway. The idea of the art work makes the residents left abandoned riverside gradually to replant recalls a better life when they lived in the Mersey surrounded communities.

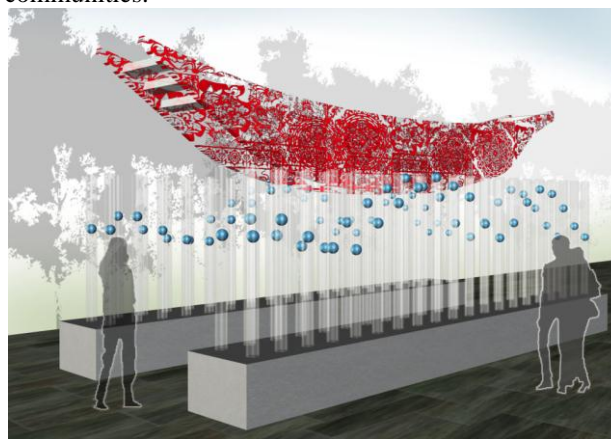


FIGURE 1 Effect drawing of "Smart Suspension Ball" placed in the Mersey side of Liverpool

For some reasons, finally this art installation didn't be showed in Liverpool. The creative idea was selected by Shanghai Shentong Metro Corporation Limited in 2010. The size of the art installation was re-adjusted to fit the metro station. It was also renamed to "Bright Wave" because the dancing spherical objects symbolize

*Corresponding author e-mail: wellington711@gmail.com

Shanghai's "urban pulse." The total length of this digital art installation was more than 15 meters and the height of each transparent tube was up to 2.5 meters. It was placed in the hall of main station of 2010 Shanghai World Expo metro Line 7 (Houtan Station of Shanghai Metro Line 7).



FIGURE 2 Effect drawing of "City Wave" at metro station

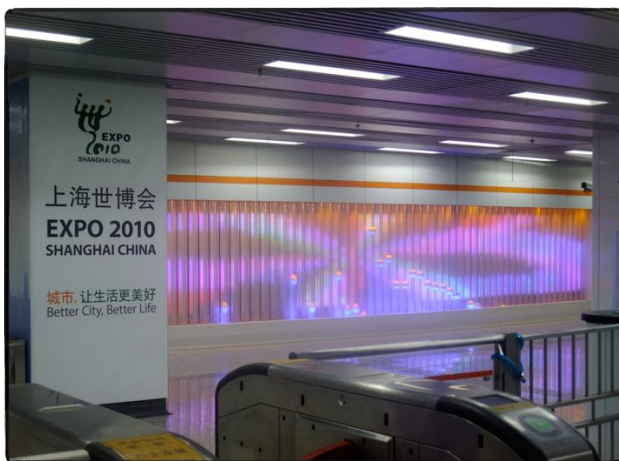


FIGURE 3 "City Wave", at Houtan station of Metro Line 7 in Shanghai

After making a number of successful shows, "Smart suspension ball" digital installation received some feedbacks from users and audiences. This year it was redesigned for meeting more and higher technical requirements:

- 1) Suspended position accuracy of the spherical objects. Accuracy means that the control hardware and software upgrades to using height sensors to get feedback to form a closed loop control.
- 2) Showing a variety of shapes. Improved the accuracy of the suspension height, so that the style is not just showing linear variation array of three-dimensional graphics rendering possible.
- 3) Miniaturization of the device. Miniaturization means that in the more space you can show this type of work, but also the recreation of industrial design.

In response to above ideas, the optimal display space for this creative product should be interior environments such as bars, cafes, clubhouse and other semi-public space. So, the product will be upgraded in several important aspects as follows:

- 1) The overall height of the glass tube in product is controlled in about 80cm. The sphere diameter will be no more than 4cm.
- 2) Unitization and modular design. Each spherical object in the tube of product is like a display pixel. User can set up a different number of units to constitute entities of varying resolution display system.
- 3) Limits the noise of industrial fans built inside the product under 35db or less.
- 4) Using an ultrasonic distance sensor (Sharp) to get the height data of each spherical object, which forms a closed loop control system. The new controlling system makes it possible to control the height of the spherical object more precisely and form them to be a shape.

3 Model of system

In a prototype of the product design, accuracy of experimental data depends on the level of craftsmanship in making. Different batches of materials such as the thickness of the wall of tube, the weight of the spherical objects or the voltage level of digital port on Arduino panel will have a direct impact on the movements of spherical objects. Therefore, it will be unavoidable to have some errors in the data of the system. Followed the laws of mechanics and based on fuzzy control theory under, we proposed a suspension control model.

First of all under physical laws, we will derive the relationship between the height of the spherical object and the wind force.

The cross-sectional area of the spherical object is

$$S_{ball} = \pi R_{ball}^2 = 3.14 \times 0.02^2 = 0.001256 \text{ m}^2, \quad (1)$$

As for the tube diameter is 0.04m, the wall is 0.0025m; as for the spherical object sphere diameter is 0.04m and a radius of 0.02m.

Thus, a formula of wind force the spherical object suffered can be derivate by Bernoulli equation:

$$F_{fan} = W_p \times S_{ball} = \frac{v_{fan}^2}{1600} \times 0.001256(\text{kN}) = \frac{v_{fan}^2}{1600} \times 1.256(\text{N}), \quad (2)$$

where v_{fan} means industrial fan wind speed, velocity near fans approximately $v_{fan} = 2.2\text{m} / \text{s}$, while another outlet of tube (about 0.8 meters away from the fan) winds $v_{fan} = 1.8\text{m} / \text{s}$, show wind speed and the distance from the outlet on.

$$v_{fan} = f(D_{ball}) = kD_{ball} + C, \quad (3)$$

Here D_{ball} represents the distance from the bottom of the tube to the bottom of the spherical object. By the

$$\begin{cases} 2.2 = k \times 0.1 + C \\ 1.8 = k \times 2.2 + C \end{cases} \Rightarrow k = -0.1905, C = 2.219,$$

$$v_{fan} = f(D_{ball}) = k - 0.1905D_{ball} + 2.219. \tag{4}$$

Therefore, by the formula (1) to (3) can be obtained

$$F_{fan} = \frac{(-0.1905D_{ball} + 2.219)^2}{1600} \times 1.256 = 0.000028D_{ball}^2 - 0.00066D_{ball} + 0.0039(N). \tag{5}$$

On this basis, the experimental observation will focus on the relationship between the height of the spherical object and the speed of the wind.

3.1 DETECTION OF SUSPENDED HEIGHT

According to the product used in the sphere diameter, tube diameter sizes, the actual height of the sphere from the target height is divided into nine levels: -0.3 m, -0.2 m, -0.1 m, -0.04 m, 0 m, 0.04 m, 0.1 m, 0.2 m, 0.3 m and the distance fuzzy language to describe size: ultra low(HN4), very low(HN3), lower(HN2), slightly lower(HN1), middle(H0), slightly higher(HP1), higher(HP2), very high(HP3), ultra high(HP4), as shown below.

TABLE 2 Fuzzy function of fans speed

	255	196	189	188	187	186	185	184	183	181	179	177
FN4	0	0	0	0	0	0	0	0	0	0.2	0.5	1
FN3	0	0	0	0	0	0	0.5	1	0.5	0	0	0
FN2	0	0	0	0	0	0.5	1	0.5	0	0	0	0
FN1	0	0	0	0	0.2	1	0.5	0.2	0	0	0	0
F0	0	0	0	0.5	1	0.5	0.2	0	0	0	0	0
FP1	0	0	0.5	1	0.2	0	0	0	0	0	0	0
FP2	0	0	1	0.2	0	0	0	0	0	0	0	0
FP3	0.5	1	0.2	0	0	0	0	0	0	0	0	0
FP4	1	0.5	0.2	0	0	0	0	0	0	0	0	0

3.3 FUZZY CONTROL STRATEGY

- 1) If the ball is close to the target position “ultra low” (HN4), the wind should be “ultra strong” (FP4).
- 2) If the ball is close to the target position is “very low” (HN3), the wind should be “very strong” (FP3).
- 3) If the ball is close to the target position is “lower” (HN2), the wind should be “strong” (FP2).
- 4) If the ball is close to the target position “slightly lower” (HN1), the wind should be “slightly stronger” (FP1).
- 5) If the ball from the middle of the target location (H0), the wind should be in middle (F0).
- 6) If the ball is close to the target position “slightly higher” (HP1), the wind should be “slightly weak” (FN1).
- 7) If the ball is close to the target position “higher” (HP2), the wind should be “weak”(FN2).
- 8) If the ball is close to the target position “very high” (HP3), the wind should be “very weak” (FN3).

TABLE 1 Fuzzy language of distance

	-0.3	-0.2	-0.1	-0.04	0	0.04	0.1	0.2	0.3
HN4	1	0.4	0	0	0	0	0	0	0
HN3	0.4	1	0.4	0	0	0	0	0	0
HN2	0	0.4	1	0.1	0	0	0	0	0
HN1	0	0	0.1	1	0.2	0	0	0	0
H0	0	0	0	0.2	1	0.2	0	0	0
HP1	0	0	0	0	0.2	1	0.2	0	0
HP2	0	0	0	0	0	0.1	1	0.4	0
HP3	0	0	0	0	0	0	0.4	1	0.4
HP4	0	0	0	0	0	0	0	0.4	1

3.2 FUNCTION OF CONTROL

This controllable value of the issue refers to the value of the speed of a industrial fan as a variable F_{fan} . The product can be transformed into the following proposition: the level of wind speed of fans is controlled by the voltage U_{fan} . The level of voltage, U_{fan} is up to the voltage from the Arduino board's analogue output ports. It can output PWM data which range is [0 255]. The fuzzy language to describe PWM data: ultra weak(FN4), very weak(FN3), weak(FN2), slightly weak(FN1), middle(F0), slightly strong(FP1), strong(FP2), very strong(FP3), ultra strong(FP4). The control function is as follows:

- 9) If the ball is close to the target position “ultra high” (HP4), the wind should be “ultra weak” (FN4).

3.4 FUZZY MATRIX

$$R = (HN4 \rightarrow FP4) \cup (HN3 \rightarrow FP3) \cup (HN2 \rightarrow FP2) \cup (HN1 \rightarrow FP1) \cup (H0 \rightarrow F0) \cup (HP1 \rightarrow FN1) \cup (HP2 \rightarrow FN2) \cup (HP3 \rightarrow FN3) \cup (HP4 \rightarrow FN4) =$$

$$\begin{aligned} & (HN_4^T \wedge FP4) \vee (HN_3^T \wedge FP3) \vee (HN_2^T \wedge FP2) \vee \\ & (HN_1^T \wedge FP1) \vee (H0^T \wedge F0) \vee (HP_1^T \wedge FN1) \vee \\ & (HP_2^T \wedge FN2) \vee (HP_3^T \wedge FN3) \vee (HP_4^T \wedge FN4). \end{aligned} \tag{6}$$

By the (6) finally get fuzzy control strategy matrix,

$$R = \begin{bmatrix} 1 & 0.5 & 0.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 1 & 0.4 & 0.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.4 & 0.4 & 1 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 1 & 0.2 & 0.2 & 0.2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0.5 & 1 & 0.5 & 0.2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.2 & 0.2 & 1 & 0.5 & 0.2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.2 & 0.5 & 1 & 0.5 & 0.4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.4 & 0.5 & 1 & 0.5 & 0.2 & 0.4 & 0.4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.4 & 0.4 & 0.4 & 0.2 & 0.5 & 1 \end{bmatrix} \quad (7)$$

3 Software Design

4

Programming language of System Control module of the software uses Java-core Processing language. The flow chart is as follows.

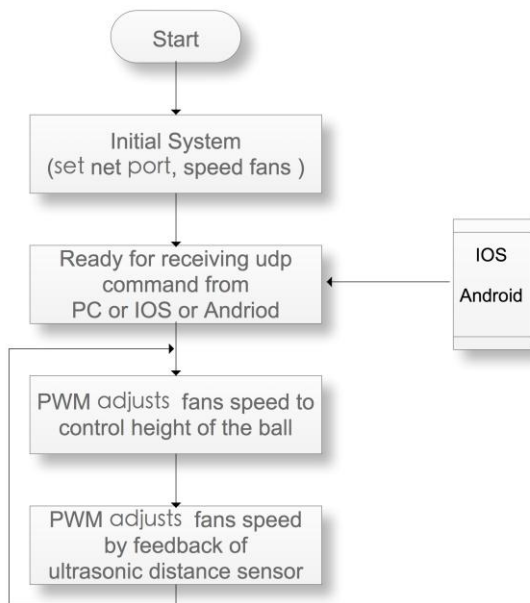


FIGURE 4 Flow chart of software design

The WiFi module is built in units' hardware to communicate with each other and the other intelligent terminals such as IOS, Android system. Transferring signals and commands in UDP protocol. The format of controlling command will be open for developers in different systems to develop different shape display mode individuation.

5 Industrial Design

"Smart suspension ball" unit consists of four transparent tubes with a hexagonal base composition. Each height of the sphere object can be controlled independently, when a user uses a plurality of unit combination, the spherical objects in the tubes formed by program control in different spatial pattern. Here is some ideas using bionic to design the figure of the base part: on the one hand the base need a large enough area to enhance the stability of the monomer product; on the other hands, analogue cellular shape achieved by the combined overall shape of a plurality of units coordinated.



FIGURE 5 Effect drawing of a unit(up); Effect drawing of a group of units in exhibition(down)

6 Conclusions

First of all, "smart suspension ball" is an innovative product with creative art and humanities works. In the second, the product combined the fuzzy control system design for the technical means. Finally, it also represents domestic design and creation of art works changing from the traditional manual way to the mode of workshop in a modern high-tech means. Meanwhile, making the physical objects as an interactive elements into interaction, means that not only tech disciplines but the visual communication, industrial modelling, sculpture, landscape and even construction and other visual arts, public art are also incorporated into the interactive art. More materials inclusiveness, diversity and technical means manifestation makes digital interactive art goes toward to a comprehensive direction.



FIGURE 6 Groups of units showed in Himalayas art gallery, Shanghai, China


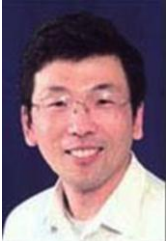
Acknowledgments

This project is supported by School of Computer and Information Technology, Beijing Jiaotong University and Fine Art College, Shanghai University. I have to

acknowledge assistance and encouragement from Virtual Lab of Fine Art College, Shanghai University, special work by Jiakang Ji, Dayi Zhu and Zhimin Zhang.

References

[1] Sun Yaojiang, Hu Zhihua, He Xiangyang 2011 *Pipe Science Technology and Engineering* 11(5) 970-4 (in Chinese)
 [2] Moore H 2012 *MATLAB for Engineers* Publishing House of Electronics Industry 44-101
 [3] Wu Shili 2008 *Basic Fuzzy Mathematics and Program Design* China's Water Conservancy and Hydropower Press 205-28 (in Chinese)
 [4] Zhang Xiaohong, Pei Daowu, Dai Jianhua 2013 *Fuzzy Mathematics and Rough Set Theory* Tsinghua University Press 161-9 (in Chinese)
 [5] Kuai Wancheng 2012 Fuzzy control algorithm for intelligent vehicle steering control system design *Friends of Science Amateurs* Oct 2012 21-2(in Chinese)
 [6] Yang Dixing 2012 *Silicon Valley* 17 163-4
 [7] Wang Rong 2009 *Computer Technology and Its Applications* No 12 128-34(in Chinese)
 [8] Wang Lu, Cui Yian, Su Hong, Cai Zixing 2005 *Computer Engineering and Applications* 15, 30-3 (in Chinese)

Authors	
	<p>Zheng Wang, born in July 11, 1978, Nanjing, China</p> <p>Current position, grades: Postdoctor in school of computer and information technology of Beijing Jiaotong University, Doctor of Art and Design, vice professor in Beijing Jiaotong University</p> <p>University studies: Art and Design(doctor degree) in Fine Art College of Shanghai University</p> <p>Scientific interest: Interactive Art, Digital Art</p> <p>Publications: 2 Patents, 3 books, 8 papers</p> <p>Experience: Wang Zheng was researching in digital interactive art. He focused on natural interaction and how to use automatic control theory in art work creation. His art work and installation had been displayed on many galleries, museums and exhibitions.</p>
	<p>Zhenjiang Miao, born on 1965, Beijing, China</p> <p>Current position, grades: professor in school of computer and information technology of Beijing Jiaotong University</p> <p>University studies: school of computer and information technology of Beijing Jiaotong University</p> <p>Scientific interest: machine visual and virtual reality</p> <p>Publications: more than 100 papers</p> <p>Experience: Miao Zhenjiang was researching in HCI and machine vision. He used to work in France and Canada. From 2004 he came back to Beijing Jiaotong University to be a professor in school of computer and information technology.</p>

Decision-making model of the urban regeneration construction project based on environment improvement

Shilong Li^{1,2*}, Hongyan Tian³

¹ Faculty of Construction Management and Real Estate Chongqing University, Chongqing City, China, 400045

² Centre for Construction Economics and Management Chongqing University, Chongqing City, China, 400045

³ Department of Building Economy Chongqing Vocational College of Architectural Engineering, Chongqing City, China, 400070

Received 6 October 2013, www.tsi.lv

Abstract

Urban regeneration is regarded as a more deliberate and harmonious progress of development, which takes a series of more sensible and multiple regeneration ways to achieve the goals, such as protection, repair, reuse and redevelopment. It is suggested that urban regeneration has positive effects on urban development and social economic system. Improvements in urban systems mainly originate from the environmental improvement, including ecological environment, social environment and the neighbourhood environment. Urban regeneration construction project and its decision-making model become the important research contents of urban regeneration. This paper discusses the urban regeneration from angle of project, and suggests a mathematical decision-making model of the urban regeneration construction project, which considers the risk, cost and environment constraints. Meanwhile, it is suggested that the joint exploitation is an important selection criterion. A case simulation is suggested in this paper in order to test the strategy model.

Keywords: Urban Regeneration, Decision-making Model, Project

1 Introduction

It is considered that urban regeneration originates from large-scale torn down and reconstructed renewal projects, which is regarded as the sustainable construction tasks in the development of urban construction. The old urban tissue may inevitably appear to physical form decline during the process of urban development and expansion. The urban pattern, public facilities and buildings are faced with the transformation of physical and economic performance or may be reconstructed considering the overall environment requirements. Due to the various constant impacts and challenges during the vicissitude of the social economy and political structure, some of the old urban areas are increasingly at the disadvantageous positions [1]. It has been pointed out that urban regeneration is the inevitable product in the process of urban development by different field researchers, which is also an intuitive performance of development and changes of the urban economy, society, environment and spatial morphology [2].

Urban regeneration is implemented by different construction projects, which could improve the urban function, ameliorate the living environment and promote the comprehensive quality of the city. The element diversifications of the space, quantity and quality are brought about by the urban regeneration, which will affect all urban systems and individuals with the implementation of new URCP (urban regeneration

construction project). Whitehead, T researched the link between urban quality improvements and economic activity, and pointed out that it was obviously important for the urban regeneration and renaissance agendas which posit attractive and well-designed environments as a way to create the right conditions for promoting economic growth [3]. Whitley, R and Prince, M recognized that urban socio-environmental conditions could affect the development and course of numerous health problems in their research, who suggested that urban regeneration programs could have an everyday functioning, coping and recovery for people with a mental illness [4]. Lawless, P examined the relationship between transport investment and urban regeneration based on an English provincial city, and pointed out that new investment had only a relatively limited impact on regeneration [5]. Biddulph, Mike examined the economic and governance context through which new forms of urban design policy and guidance have emerged, and discussed whether and how they have been applied to developments emerging across the centre [6]. Lloyd, MG researched the urban regeneration and the community development, and suggested that community planning and social inclusion partnerships were the vehicles for the urban regeneration [7]. Rogers, Chris D.F. thought that little had been done to test urban regeneration solutions in his paper and described a methodology that has developed future scenarios for the year 2050 against which to test the robustness of current engineering solutions, thereby

* Corresponding author e-mail: ls1cqu@126.com

providing unique insights into the potential impacts of present urban planning and design decisions, and thus financial investments [8].

Urban regeneration is embodied by a more deliberate and harmonious progress of development, which takes a series of more sensible and multiple regeneration ways to achieve the goals, such as protection, repair, reuse and redevelopment. Researchers suggest that urban regeneration has positive effects on urban development and social economic system. Improvements in urban systems mainly originate from the environmental improvement, including ecological environment, social environment and the neighbourhood environment. Construction project of urban regeneration and its decision making become one of the important research content of urban regeneration.

2 Problem Description

Urban regeneration is often consisted of two aspects. The first one is the regional regeneration, considering the holistic regeneration, which includes the urban environment, planning, economic growth, and transportation and so on. The other one is the construction, which mainly considers the decision of construction projects in various conditions. Urban regeneration in the aspect of construction projects has experienced several stages, including tear down and reconstruction, neighbourhood repair, economic recovery and public private partnerships and then turn to multi stakeholder partnerships. Due to the urban environmental degradation caused by urban function hysteresis, such as the traffic congestion and unmerited waste disposal, it is called to improve the urban living conditions and environment quality by urban regeneration construction project. It is suggested that urban regeneration construction project usually tends to appear by form of the modified construction in environment, economy, society or urban personal behaviour, which decision-making model is a kind of complex, hierarchical structure, multi-objective and uncertain system model. It is also considered with the random, high dimension and nonlinear characteristics [9].

Urban regeneration construction project is often chosen by the economy cost and considered as a one-time injection from the perspective of urban developers in the process of planning and development, but failed to consider long-term economic, social and cultural factors. Actually, urban regeneration often involves scientific and rational allocation of unit space increment, which is needed to grasp and update the related information of the urban unit, and then to make the choices and judgments, including the scale of the urban regeneration construction project and the project portfolio. The project decision-making model is often used to be the ANP model, system dynamics model and so on. In the research field, it is limited in the discussion about the model and method of the urban regeneration construction project selection,

which is mainly because of the complexity of the urban regeneration construction project. In this paper, it is suggested to define the decision-making model of the urban regeneration construction project as the government environmental governance project decision-making model, which is from the perspective of the government governance research.

3 Model Establishment

3.1 ANALYSIS OF MODEL ESTABLISHMENT

This paper discusses urban regeneration from the angle of project and especially the decision-making model, which often involves the conflicting economic, environmental, and socio-ecological impacts. However, too much emphasis is focused on the economic goal implementation in the former research, which may neglect the social and environmental demand. It is suggested that the DM-URCP (decision-making model of the urban regeneration construction project) is multi objective decision-making model, which is restricted by the following aspects.

3.1.1 Risk constraint

Risk is composed of risk factors, risk accidents, risk loss and so on. Generally, there are two kinds of risk definitions. The first definition emphasizes uncertainty, which illustrates the possible results of loss, profit or balance. The second definition just emphasizes uncertainty of the loss. In this paper, it is suggested that risk constraint of the urban regeneration construction project and its decision-making comes from the probability of decision-making failure and the environment safety risk in the process of the urban regeneration construction, which is described by the loss cost. Decision maker needs to undertake corresponding consequences, which cannot be transferred. The evaluation for the risk constraint of DM-URCP is the smaller the better, which usually uses the method of the fuzzy mathematics.

3.1.2 Cost constraint

In the urban regeneration construction project, finance and its source are usually to be thought and discussed in front of the decision-making. According to the classical research results, it is suggested that weighing the finance constraints and agency costs of the capital budgeting is very important, which are concern with the construction cost. Resources are limit in any projects. Supposing the limitation of the resource reflecting in capital, it is difficult to estimate the cost of URCP, which is affected by the long construction cycle and various contents. Furthermore, funds will be limited in the urban regeneration construction project when regeneration result cannot be confirmed. According to the theory of

value engineering, cost and function regeneration have a matching degree between them. It will reduce the cost where two or more projects are implemented at the same time. Therefore, it will cause the different cost by selecting the different project portfolio selection in the DM-URCP.

3.1.3 Environment constraint

Bearing capacity of resources and environment is limited. It is widely recognized that intense of the energy supply, unbalance of the energy consumption, low utilization of the energy use and environmental pollution caused by energy consumption inevitably lead to the environmental pollution. In the process of the urban regeneration construction project, carbon emissions will be produced during the demolition of buildings. Meanwhile environment will be occupied during the process of the new construction. In this paper, it is suggested that environmental constraints embodied in the water resource and land constraints. If the construction or demolition amounts of the urban regeneration construction project exceed certain amounts, more urban land will be occupied and will cause the land damage, which means that more broken lands cause environmental problems. Then, consumption of water resources will rise with the increase of the construction scales.

3.1.4 Technology constraint

Technology constraint of the urban regeneration construction project usually appears in some special projects or areas, which are caused by position of plan and construction project stationary, such as harmless treatment of the soil and super building construction etc. For most of the urban regeneration construction projects, it is not to be considered of the technology constraint. Then, in this paper, it is thought that the technology constraint is without any influence on the decision-making.

3.1.5 Culture constraint

Culture is very important towards the city. One of the culture carriers is the construction projects in the city. The culture may be destroyed with the implement of the project. It is recognized that culture constraint consists of attitudes of the urban residents to the urban regeneration construction project and the elasticity alteration of the urban veins, which is a kind of soft constraints and difficult to measure. In this paper, culture constraint is not considered in the decision-making model of the urban regeneration construction

3.2 PARAMETERS DEFINITION AND THE MODEL

In decision making of the urban regeneration construction project, it is important to reduce the risk, cost and the environment effects as well as implement the programs in

accordance with the technology constraint and the culture constraint.

Suppose i ($i=1,2,3,\dots,n$) is the urban regeneration construction project point. Firstly, analyse and establish the model for a single project. It is suggested that the decision making of the urban regeneration construction project depends on the cost and the risk, which consists of the construction, demolition and the environmental effects, if do not consider the culture and technology constraints. According to the principle of construction, urban regeneration construction project will bring out the environment pollution and the resources consumption, which can also be translate into the cost, defined as the environmental loss cost.

Demolition cost of the urban regeneration construction project cost_i^d is a function of demolition scale Q_i^d . Environment capacity in specific region is limited. Thus, if demolition scale of the urban regeneration construction project outnumber the maximum capacity, which usually estimated as a certain demolition amount Q_0^{d1} . If demolition scale of the project is less than the min capacity Q_0^{d2} , it will cause the land fragmentation which is also the environmental loss cost. Then,

$$\text{cost}_i^d \begin{cases} = \alpha Q_i^d + \lambda |Q_i^d - Q_0^{d1}| + C_0^{d1} \\ \text{s.t. } \alpha > 0; \\ \lambda > 0; \\ Q_i^d > Q_0^{d1}; \\ = \alpha Q_i^d + \lambda |Q_i^d - Q_0^{d2}| + C_0^{d2}, \\ \text{s.t. } \alpha > 0; \\ \lambda > 0; \\ Q_i^d > Q_0^{d2} \end{cases} \quad (1)$$

where cost_i^d represents the demolition cost of the urban regeneration construction project i . α is the unit of demolition cost of the urban regeneration construction project. λ is the unit of environmental loss cost. C_0^{d1} and C_0^{d2} are the constants. Meanwhile, the risk will rise corresponding with the increase of the demolition cost. Construction cost of the urban regeneration construction project cost_i^c is a function of construction scale Q_i^c . In the decision making of the urban regeneration construction project, large scale of construction may mean better quality and more comfortable environment. Certainly, the cost_i^c will rise with the ascendant of the construction scale.

Decision should balance the function of the urban regeneration construction project and the cost. Meanwhile, scale of the regeneration will bring out the risk corresponding with the cost. Risk reserve should be considered in the decision making model. Construction of

the project is different from the demolition, which is relevant with the use of funds system. If the construction scale outnumbers a certain scale Q_0^{b1} or less than a scale Q_0^{b2} , cost will rise or decrease correspondingly. Thus, it is suggested that the cost of the construction includes the direct construction cost, the exceed cost and the risk cost. Then,

$$\text{cost}_i^2 = \begin{cases} wQ_i^b + \rho \frac{|Q_i^b - Q_0^{b1}|}{Q_i^d} + C_0^{b1} \\ \text{s.t. } w > 0; \\ \rho > 0; \\ Q_i^b < Q_0^{b1} \\ wQ_i^b + \rho \frac{|Q_i^b - Q_0^{b2}|}{Q_i^d} + C_0^{b2} \\ \text{s.t. } w > 0; \\ \rho > 0; \\ Q_i^b < Q_0^{b2} \end{cases}, \quad (2)$$

where cost_i^2 represents the construction cost of the urban regeneration construction project i . w is the unit of construction cost of the urban regeneration construction project. ρ is the compound cost unit of exceed cost and risk cost, which is a function of the proportion exceed cost and risk cost in the same project. C_0^{b1} and C_0^{b2} are the constants. Meanwhile, the risk will rise corresponding with the increase of the construction cost.

$$M_i = \eta_i \text{cost}_i^1 + \delta_i \text{cost}_i^2 + \theta_i \left(\frac{1}{f(U)} \right) \\ \text{s.t. } 0 < \eta_i < 1; \\ 0 < \delta_i < 1 \quad (3)$$

where M represents the single urban regeneration construction project cost and the $\theta_i \left(\frac{1}{f(U)} \right)$ represents the risk utility level of the project implementation achievements, which is to judge the cost risk and environmental impact of the urban regeneration construction project. η_i and δ_i represent the demolition and construction coefficients, which are not constants. According to the description above, define

$$f(U) = r_i' x_i = r_i' \left(\frac{1}{\text{cost}_i^2} \right), \quad (4) \\ \text{s.t. } 0 < r_i' < 1$$

where r_i' is the coefficient of the risk utility.

In the urban regeneration, there is more than one project, which may have the compound possibility of

construction. It is suggested that joint exploitation will reduce the cost in the urban regeneration construction project. Set,

$$f(c) = \sum_{i=1}^n M_i - \sum_{i=1}^{n-1} \sum_{j=i+1}^n c_{ij} (M_i + M_j) \\ - \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n \beta_{ijk} (M_i + M_j + M_k) - \dots \\ - \sum_{i=1}^1 \sum_{j=i+1}^2 \sum_{k=j+1}^3 \dots \sum_{t=n}^n \varpi_{ijk\dots t} (M_i + M_j + M_k + \dots + M_t) \quad (5) \\ \text{s.t. } c_{ij} > 0; \\ \beta_{ijk} > 0; \\ \dots \\ \varpi_{ijk\dots t} > 0; \\ f(c) < C^0$$

where $f(c)$ represents joint exploitation cost of URCP, c_{ij} , β_{ijk} and $\varpi_{ijk\dots t}$ represent cost saving coefficient of the two projects joint exploitation of i, j , cost saving coefficient of the three projects joint exploitation of i, j, k and cost saving coefficient of n projects joint exploitation of $i, j, k \dots n$. The cost saving coefficient between each other is relevant with the attribute of the project and often determined by the ANP method. C^0 is a budget threshold.

Actually, DM-URCP is to search the combination of the projects, which joint exploitation cost, is minimal. Thus, DM-URCP is set as,

$$\min R = \min f(c) \\ \text{s.t. } \min f(c) < C^0 \quad (6)$$

4 Solving the Model

According to the model established in this paper, it is a typical combinatorial problem to search the min risk, max construction scale and the min cost. Particle swarm optimization is a swarm intelligence optimization method, which finds the overall optimum in a complex search space through the competition and collaboration among the particles and has the advantage of solving speed and stability. It has mature calculation program and acceptable for the model solving if variable value is controlled.

In a basic particle swarm optimization algorithm, it establishes an effective information sharing mechanism through the simulation of the birds flying [10]. Compared with traditional optimization method, the particle swarm optimization algorithm has characteristics of such as faster and convergence speed, smaller scale calculation etc. According to the particle swarm optimization algorithm, if there are M particles in N -dimensional space, the spatial coordinates of particle i is $x_i = \{x_{ij}\}^T$

and the displacement is $v_i = \{v_{ij}\}^T$, the optimal solution could be found as the following iterative method: $v_i = w \cdot v_i + c_1 \cdot rand_1(\cdot) \cdot (pbset_i - X_i) + c_2 \cdot rand_2(\cdot) \cdot (gbset - X_i)$, $x_i = x_i + v_i$, where $i=1,2,3,\dots,M; j=1,2,3,\dots,N$.

w is a N-diagonal matrix, which represents the inertia weight. $pbset_i$ is the optimal value of particle i in the spatial location before. $gbset$ is the optimal value of all particles in the spatial locations before. C_1 and C_2 are the constant acceleration. $rand_1(\cdot)$ and $rand_2(\cdot)$ produce the figure between 0 and 1 randomly. In addition, in the process of iteration, define the displacement threshold v_{ijmax} , and then, $\forall v_{ij} > v_{ijmax}, v_{ij} = v_{ijmax}$.

5 Case Simulation

5.1 CASE DEFINITION

Because of the urban regeneration demand, construction is associated with the demolition. The decision is that the less of the demolition the better corresponding with the more of the construction the better, which should also meet the risk and cost constraints. In this paper, Decision-making model of urban regeneration construction project does consider the culture and technology constraints temporarily and define the risk and environment constraints into measurable cost constraint. Suppose there are nine urban regeneration construction project points. The amount of demolition and construction are showed in Table 1.

TABLE 1 Demolition and Construction Amount

Points	Demolition(m ²)	Construction(m ²)
1	1200000	3000000
2	450000	750000
3	600000	1150000
4	800000	1600000
5	550000	350000
6	1250000	1000000
7	620000	1150000
8	700000	650000
9	4500000	3500000

References

[1] Han Zhang, Linfei Song 2008 British and American Urban Regeneration Research Review from the Domestic Scholars *Urban Problems* 2(2) 78-83 (In Chinese)
 [2] Zhao Yinghui 2010 Low-carbon Design Strategy in Urban Renewal Planning – The Case of Mutoulong Project in Shenzhen City Renewal Programme *Urban Planning Forum* 192(7) 44-7 (In Chinese)
 [3] Whitehead T, Simmonds D, Preston J 2006 The Effect of Urban Quality Improvements on Economic Activity *Journal of Environmental Management* 80(1) 1-12
 [4] Whitley R, Prince M 2006 *Health Promotion International* 21(1) 19-26
 [5] Lawless P 1999 Transport Investment and Urban Regeneration in a Provincial City: Sheffield, 1992-96 *Environment and Planning C-Government and Policy* 17(2) 211-26

5.2 Parameters setting and the result

According to application of particle swarm optimization research, it suggested to solve the model by particle iteration. The parameters of ρ and r_i value form the empirical data. Adopt the method of ANP, which is a mature method, c_{ij} , β_{ijk} and $\varpi_{ijk\dots}$ value [0,1] placed in the calculation process. The other parameters setting are shown as following Table 2.

TABLE 2 Parameters setting

Parameter	Value	Parameter	Value
α	0.35	w	1
λ	0.5	Q_0^{b1}	2000000
Q_0^{d1}	1500000	Q_0^{b2}	1000000
Q_0^{d2}	1000000	C^0	200000
C_0^{d1}	0	C_0^{b1}	0
C_0^{d2}	0	C_0^{b2}	0



By the model established in this paper and the solving method, which is set by controlled stages, the result, which is processed by the planning empirical data is

$$\begin{cases} i = (4, 7) \\ (\eta_4, \delta_4) = (0.75, 0.8) \\ (\eta_7, \delta_7) = (0.55, 0.9) \end{cases}$$

6 Conclusion

This paper discusses the urban regeneration from angle of project, and tries to describe a situation of the decision making model of the project. It can determine the amount and the joint exploitation strategy through the model established in the paper and solving method. However, it is only a simulation, which defines many constraints, it is suggested that urban regeneration consist of constraints and patterns research besides the project. It is only a simulation, which defines many constraints, it is suggested that urban regeneration consist of constraints and patterns research besides the project.

[6] Biddulph M 2011 *Progress in Planning* 76(2) 63-103
 [7] Lloyd M G 2002 Urban Regeneration and Community Development in Scotland: Converging Agendas for Action *Sustainable Development* 10(3) 147-54
 [8] Rogers C D F, Lombardi D R, Leach J M, Cooper R F D 2012 *Proceedings of the Institution of Civil Engineers: Engineering Sustainability* 165(1) 5-20
 [9] Hai Zheng 2007 Algorithm Study on Models of Multiple Objective Risk Decision under Principal and Subordinate Hierarch Decision-making *Operations Research and Management Science* 16(1) 1-8 (In Chinese)
 [10] Shen S Y, Liu Y K 2010 A New Class of Fuzzy Location-Allocation Problems and Its Approximation Method *Information-an International Interdisciplinary Journal* 13(3) 577-91 (In Chinese)

Authors	
	<p>Shilong Li, born in September, 1981, Shapingba District, Chongqing City, P.R. China</p> <p>Current position, grades: the Assistant Professor of Faculty of Construction Management and Real Estate, Chongqing University, China. University studies: received his B.Sc. from Chongqing University in China. He received his M.Sc. from Chongqing University in China. Scientific interest: His research interest fields include Urban Planning, Environmental Engineering and Management. Publications: more than 10 papers published in various journals. Experience: He has teaching experience of 7 years, has completed five scientific research projects.</p>
	<p>Hongyan Tian, born in August, 1972, Nanan District, Chongqing City, P.R. China</p> <p>Current position, grades: the Associate Professor of Chongqing Vocational College of Architectural Engineering, China. University studies: received her B.Sc. from Chongqing University in China. She received her M.Sc. from Chongqing University in China. Scientific interest: Her research interest fields include Project Management, Environmental Engineering and Management. Publications: more than 15 papers published in various journals. Experience: She has teaching experience of 12 years, and has completed six scientific research projects.</p>

Displacement prediction of Liangshuijing landslide based on time series additive model

Q X Zhang¹, Y P Wu^{1, 2*}, G Zh Ou³, X G Fan¹, J H Zhou⁴

¹ China University of Geosciences, Faculty of Engineering, Wuhan, Hubei, China, 430074

² Three Gorges Research Center for geo-hazard, Ministry of Education, Wuhan, Hubei, China, 430074

³ Zunyi Normal College, Faculty of Engineering, Zunyi, Guizhou, China, 563002

⁴ College of Civil Engineering and Architecture, Guilin University of Technology, Guilin, Guangxi, China, 541004

Received 18 March 2014, www.tsi.lv

Abstract

The evolution of landslide displacement is affected by many factors. This paper studied the displacement monitoring data of Liangshuijing Landslide with Factor Analysis Method and found that the dominant factors influencing landslide displacement were in decreasing sequence: cumulative rainfall of anterior two months> rainfall of current month> the average reservoir level of current month>reservoir level fluctuation of current month. The paper selected three typical GPS monitoring points (ZJC09, ZJC11, ZJC13) of Liangshuijing Landslide to forecast their displacements by adopting the time series additive model on basis of the conclusion of previous factor analysis. The accumulative displacement of Liangshuijing Landslide can be divided into trend term and random term. The polynomial fitting was used for trend term displacement prediction. BP neural network model was used for the random displacement prediction. The final calculation results indicated that combination of factor analysis method and time series additive model could generate a reasonable and accurate prediction of landslide.

Keywords: Displacement prediction, Time series, Liangshuijing Landslide, Factor analysis, BP neural network

1 Introduction

Monitoring data of landslide displacement is a macroscopic dynamic response to external triggering factors (reservoir water level fluctuation and rainfall, etc.), reflecting the whole evolution process of landslide. It also supports important information for landslide inducing mechanism, formation process, prediction and stability evaluation. However, the characteristic of landslide displacement is complex due to the complexity and diversity of landslide formation mechanism and evolution process. Besides, there are many factors affecting the stability of landslide. Usually, it is difficult to find out the dominant factors because the selection of these variables is subjective. Therefore, it is necessary to figure out the dominant influence factors on landslide inducing mechanism and evolution process from these variables [1].

R W Jibson and D K Keefer (1989) [2] took advantage of discriminant analysis and multiple linear regression to study 220 large landslides influenced by 1811-12 New Madrid earthquake, and found that slope steepness had a stronger effect than the other three factors: slope height, slope aspect and stratigraphic variation. S Lee and J A Talib (2005) [3] used GIS and image processing to interpret satellite spatial data of landslide area, then selected topographic slope, topographic aspect, topographic curvature, lithology, land

use, vegetation index to analyse factors inducing landslide with probability-frequency ratio method. Deping Guo, Masanori Hamada (2013) [4] analysed landslide influential factors during 2008 Wenchuan Earthquake by using qualitative and quantitative method to find out that slope height, horizontal peak ground acceleration and geological structure had a stronger effect on the sliding area and volume than slope angle and rock type. He Keqiang (2005) [5] found rainfall was the most important dynamic factor among the six factors controlling displacements, using the monitoring data of displacements of Xintan Landslide, China. The above researches mostly focused on landslide stability influence factors including internal factors and external factors. Moreover, more and more attentions have been paid to external factors, (such as rainfall and earthquake). While there is less research focusing on reservoir water level, nevertheless it is an indispensable factor for reservoir bank landslides.

And once the dominant factors are determined, displacement prediction of landslide can be more appropriate and rational as considering multiple exterior influence factors. Since the 1980s, many intelligence theory and methods have been applied in the prediction of landslide displacement, such as the regression analysis, time series, grey system, artificial neural network, wavelet analysis and Kalman filter, obtaining great results [6]-[16]. When bearing a certain load, landslide

* Corresponding author e-mail addresses: ypwu@cug.edu.cn

slope will adjust itself into another stable state with small changes in its shape, size, location and time dimension. In addition, it is feasible to use the periodic monitoring data processing and analysis results to reflect physical properties of the landslide mathematically, so as to find out its laws of deformation and failure trend.

In this paper, three typical monitoring points (ZJC09 ZJC11 ZJC13) of Liangshuijing Landslide are selected. On the basis of these monitoring data and geological environment, formation mechanism, deformation and failure characteristics, dominant impact factors of the landslide are obtained from factor analysis method, an improved time series additive model is proposed to forecast its displacement.

2 Methodology

2.1 FACTOR ANALYSIS

Factor Analysis Method was first put forward in 1904 by Charles Spearman when using statistics method to study the scores of intelligence test. In the 1950s, this method began to prevail in Psychology, Medicine, Geology and Economics [17]. Factor analysis which derived from the ideas of dimension reduction can study internal structure of original variable matrix. It is also a multivariate statistical analysis method to transfer some variables with complicated relationship to a few integrated factors. When using factor analysis method, variables are divided into several groups based on correlation value to make high correlation between variables within the same group and low relevance between variables in different groups. Each group of variable represents a basic structure of unobservable variables, and the basic structure is called common factor [18]. Using the sum of linear function of reduced number of unobservable variables (common factor) and specific factor to describe the original observation, factor analysis can deduce the inner link between observed variables. Zhang T T and Yan E C (2012) [19] adopted the factor analysis method to analyse landslide deformation mechanism using monitoring data displacement of one landslide in Three Gorges Reservoir area, and the main factors influencing the landslide deformation is reservoir water level with rainfall followed. Sun X R, Ma F H, etc., (2012) [20] proposed the main influence factors on dam deformation based on Factor Analysis Method and monitoring data. Miao H B, Yin K L, etc., (2010) [21] established the comprehensive evaluation model of landslide displacement to improve the landslide prediction effect by using factor analysis. By using Factor Analysis Method to analyse monitoring data, Bi J L, Liu X, etc., (2010) [22] identified the main factors affecting seepage around the dam. Ma S S, Wang Z W (2002) [1] studied the sliding model and evolution law of one landslide by using factor analysis method which was used to analysis the landslide during different development period from the perspective of quantitative. There is few application of factor analysis method in

engineering. Therefore, it is significant and it has bright prospects to study inducing mechanism and evolution process of landslide by using factor analysis.

Suppose we have a set of m samples, each of which has p observable random variables. For the convenience of contrast, the samples were standardized by eliminating the effects of dimension and magnitude. Suppose the new observable random variables are X_1, X_2, \dots, X_p , and after standardization, the common factors are $F_1, F_2, \dots, F_m(m < p)$. The model can be described as follows:

$$\begin{cases} X_1 = a_{11}F_1 + a_{12}F_2 + \dots + a_{1m}F_m \\ X_2 = a_{21}F_1 + a_{22}F_2 + \dots + a_{2m}F_m \\ \dots\dots\dots \\ X_p = a_{p1}F_1 + a_{p2}F_2 + \dots + a_{pm}F_m \end{cases} \quad (1)$$

Equation (1) can be called factor model. The matrix term is:

$$X = AF \quad (2)$$

Here, common factor (also principal factor), F_1, F_2, \dots, F_m refer to the factors commonly possessed in the expression of observable random variables, and they are also independent and unobservable.

In the equation (2), the a_{ij} are defined as factor loadings with an absolute value smaller than equation (1). The more its absolute value is, the better X_i and F_j correlate, or the greater loading common factor F_j works on X_i . Therefore, a_{ij} can also be called common factor loadings, and matrix A refers to the factor loadings matrix.

2.2 TIME SERIES ADDITIVE MODEL

As discussed in the previous sections of landslide deformation characteristics, the occurrence of landslide displacement is related to the coupling effect of inherent geological condition and exterior inducing factors (such as rainfall, fluctuation of reservoir water level and so on.). The total displacement can be divided into different parts according to their influences. In a time series model, the total displacement, $y(t)$, contains three response components as expressed in equation (3) [23]:

$$y(t) = f(t) + p(t) + \varepsilon(t), \quad (3)$$

where, $f(t)$ refers to the trend term described as an increasing function reflecting the tendency of the series; $p(t)$ is the periodic term illustrating the fluctuation of displacement under the influence of periodic exterior factors, such as rainfall, reservoir water level and so on; $\varepsilon(t)$, known as the random term, denotes the random variable caused by the uncertainty. Because of small magnitude of $p(t)$ and unobvious periodic fluctuation of landslide displacement, the trend term in this series model

is ignored. As for $\varepsilon(t)$, back propagation(BP) neural network can be used to reflect its complex nonlinear series[24].

Each term in the time series additive model can correspond to each component of landslide displacement, making displacement prediction a significant mathematical and physical meaning. Therefore, based on

the factor analysis of displacement characteristics for Liangshuijing Landslide, multivariable neural network was established for extracting the random term of the landslide displacement by using time series additive model. Fig. 1 shows the structure chart of the landslide prediction model.

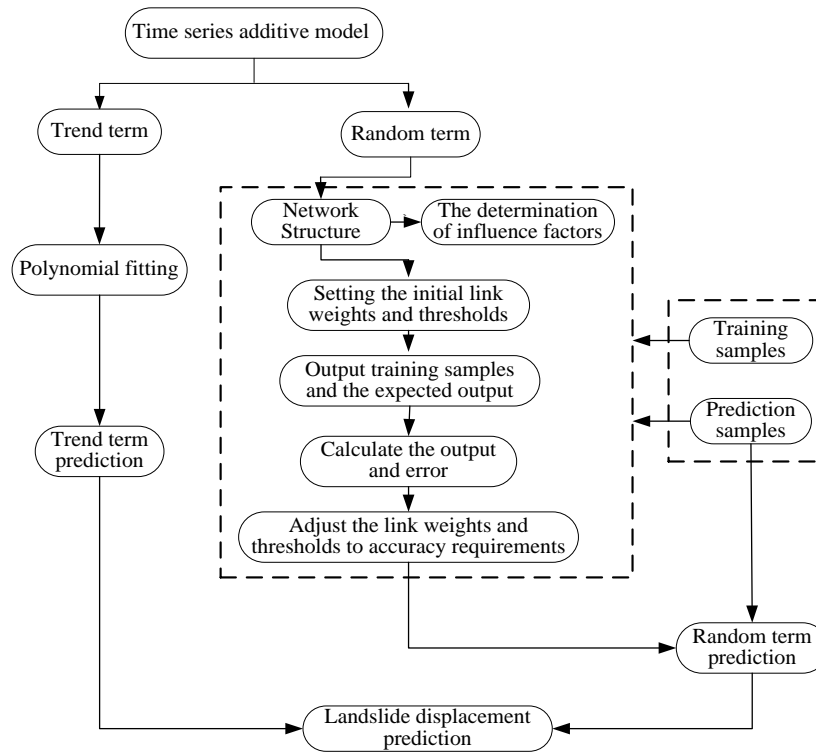


FIGURE 1 Structure chart of landslide prediction model

2.2.1 Polynomial fitting

Trend item of accumulative displacement reflects the long-term development model of landslide deformation and the overall trend. Taking the monitoring information of surface displacement and geological features of Liangshuijing Landslide into consideration, the change cycle of displacement curve can be regarded as one year. Then, Moving Average Method is adopted to smooth the landslide displacement curve, aiming at weakening the mutation effects on the displacement curve caused by influence factors. Moreover, the displacement curve after smoothing can be seen as the trend term of landslide displacement.

As one dimensional time series, landslide displacement can be described by $X_i=\{X_1, X_2, X_3, \dots, X_t\}$, and X_t denotes the displacement trend term, as in equation (4):

$$X_t = \frac{x_t + x_{t-1} + \dots + x_{t-n+1}}{n}, (t-n, n+1, n+2, \dots, t), \quad (4)$$

where, n refers to the periodic value, and in this paper, n=12, accounting for the characteristics of landslide

accumulative displacement. Depending on the shapes of displacement curve, different trend term model can be determined [25]. In general, if the trend term displacement has an ideally linear trend or exponential trend, then the least squares method and polynomial fitting method can be used, which have a wide application in landslide displacement forecast. On the contrary, GM (1, 1) and GM (2, 1) can be utilized to describe the case of exponential growth, exponential decay and periodic oscillation, respectively.

Based on the characteristics of trend term displacement of Liangshuijing Landslide, polynomial fitting method with least squares is adopted. Its mathematical model is shown in equation (5):

$$X_t = at^3 + bt^2 + ct + d. \quad (5)$$

2.2.2 BP neural network model

With the development of artificial intelligence, neural network has gradually been applied to landslide analysis and forecasting, which has obtained good results. Neural network is composed of large amounts of information processing units (neurons) connected into a wide range of

artificial networks to simulate the brain's structure and function of the system, and automatically summarizes laws from the known data to obtain the intrinsic connection between these data. Therefore, it has a strong nonlinear mapping ability. Currently there are many neural network architectures and algorithms. While BP neural networks, one of the widely used neural network, is an error back propagation network and composed by a large number of neurons self-organizing, adaptive dynamic systems, prevailing in a strong learning ability of revealing nonlinear relationship between inherent sample data.

A typical BP neural network has three layers: input layer, hidden layer and output layer. The one layer is fully connected to another; output value of several neurons forming a layer is determined by the input node values, the role of the function and the threshold. Learning process of BP neural network includes information forward propagation and error back propagation. In the process of forward propagation, the training samples spread from the intermediate layer of the input layer to the output layer, and the output values between the function and threshold can be calculated. After comparing with the expected output value, if there is an error, the connection weights should be revised in the direction of reducing errors. And so forth, until the output of the results meets the accuracy requirements. The procedure of the training BP neural network can be described as follows:

The input samples of network are:

$$A_i = (x_{i1}, x_{i2}, \dots, x_{in}), \quad (i = 1, 2, \dots, m), \quad (6)$$

where m is the learning model logarithmic; n denotes the number of input layers.

Corresponding output vector is:

$$B_i = (y_{i1}, y_{i2}, \dots, y_{ik}), \quad (i = 1, 2, \dots, m), \quad (7)$$

where the m is the number of output model, in relation to input model; the k refers to the number of input layers.

Calculation of the input of hidden layers:

$$S_j = \sum w_{ij} x_n - \theta_{aj}, \quad (8)$$

where, the W_{ij} are the collection weighs between input layer and intermediate layer; the θ_{ij} reflect the thresholds of hidden layers; j is the number of neurons of hidden layer.

Calculation of the output of hidden layers:

$$b_j = f(S_j) = \frac{1}{1 + e^{-S_j}} \quad (\text{S type function}), \quad (9)$$

$$\text{Or } b_j = f(S_j) = \frac{1}{1 + e^{-2S_j}} \quad (\text{tangsigs function}). \quad (10)$$

Calculation of the input, output of output layers:

$$L_i = \sum_{j=1}^n v_{ij} b_j - \gamma_i, \quad (11)$$

$$Y_i = f(L_i), \quad (12)$$

where Y_i are outputs; V_{ij} reflects the collection weighs between intermediate layer and output layer; γ_i are the thresholds of output layers; $f(L_i)$ is one kind of S type function.

Repeated training and adjustment of the error:

$$E = \sum_{i=1}^k (Y_i - B_i)^2 / 2. \quad (13)$$

If the value of E is less than a certain prediction accuracy request, it can be proved that the network has been trained well enough to predict the displacement of landslide using the new value.

3 Case study

Fig. 2 shows that Liangshuijing Landslide is located on the right bank of Yangtze River, Three Gorge Reservoir, China. The plane form of this landslide shows U-shape, with the terrain of approximate round-backed armchair at the rear part. The front and trailing edge elevation are nearly 100m and 319.5 m respectively, with a total volume of about $407.79 \times 104 \text{ m}^3$. After trail storage of Three Gorges Reservoir to 172 m, the landslide began to deform in November 2008, and then obvious deformation appeared in April 2009. With a rate (1cm/day) of landslide fissure development, maximum tension cracks has amounted to 60 cm. Due to the narrow area of the Yangtze River waterway and large volume of sliding materials, the instable landslide will cause a large swell after falling into the river. It is a great threat to the safety of passing ships and passengers, and the economic losses and social influences are incalculable. As a result, the local government paid high attention to the landslide, and began monitoring all-round in May 2009. Fig. 3 shows the monitoring data studied in this paper until April 2012.

4 Displacement prediction of landslide

4.1 DETERMINATION OF DOMINANT INFLUENCE FACTORS

Rainfall and reservoir water level fluctuation are important factors of inducing reservoir bank landslides [26]. Influence of reservoir level fluctuation on the landslide displacement has been introduced in the

previous study, and landslide displacement is closely related to submerged body under the reservoir water level. According to the result of Factor Analysis Method, the largest change of current monthly reservoir water level can be regarded as influence factor of landslide deformation. There are a large amount of studies on the

relationship between landslide and rainfall, which show that 1-2 month rainfall before the occurrence of landslide has promoting effect on landslide deformation [27, 28].

As a result, rainfall in the current month and rainfall two months before are selected as the influence factors of landslide deformation.

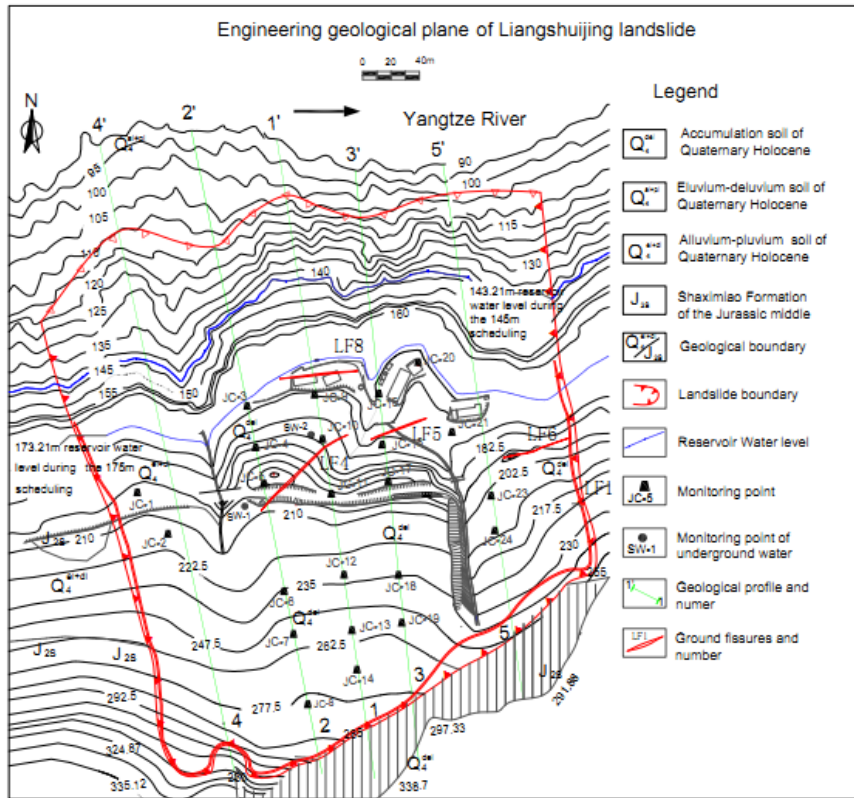


FIGURE 2 Engineering geological plane of Liangshuijing Landslide

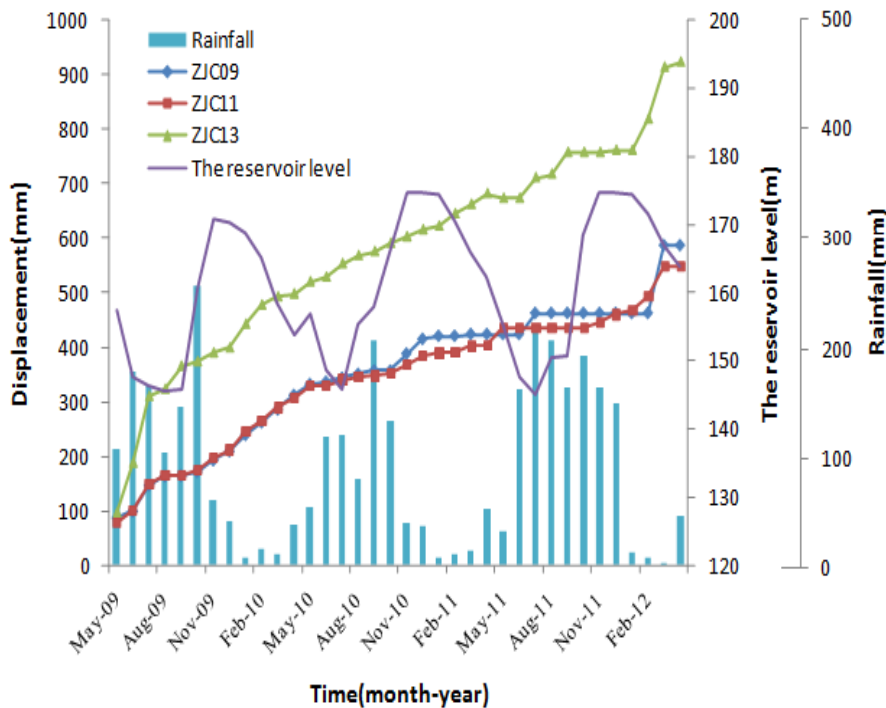


FIGURE 3 Monitoring data of rainfall reservoir water level and displacement

4.2 COMPREHENSIVE ANALYSIS OF PREDITION RESULTS

4.2.1 Displacement prediction of trend term using polynomial fitting

Taking advantages of extraction of trend term and polynomial fitting theory, we can get the parameters of trend term displacement of the monitoring points (ZJC09, ZJC11, ZJC13) after polynomial fitting, which can be seen in Table 1.

The parameters in Table 1 are adopted for landslide displacement prediction. The results are shown in Fig.4. It shows that all these three monitoring points can predict

the displacement well, mainly because of large number of monitoring data, relatively gentle development of displacement and great effect of eliminating the influence of the periodic item by moving average method. Increasing the moving cycle n appropriately can eliminate the influence of periodic term more obviously, but at the same time, more early monitoring data are needed.

TABLE 1 Parameters of trend term displacement after polynomial fitting

GPS No	Parameters				Accuracy /R ²
	a	b	c	d	
ZJC09	-0.0126	0.6128	4.5612	87.852	0.9969
ZJC11	-0.0079	0.3438	8.3983	73.754	0.9965
ZJC13	0.0054	-0.5087	29.718	98.359	0.9963

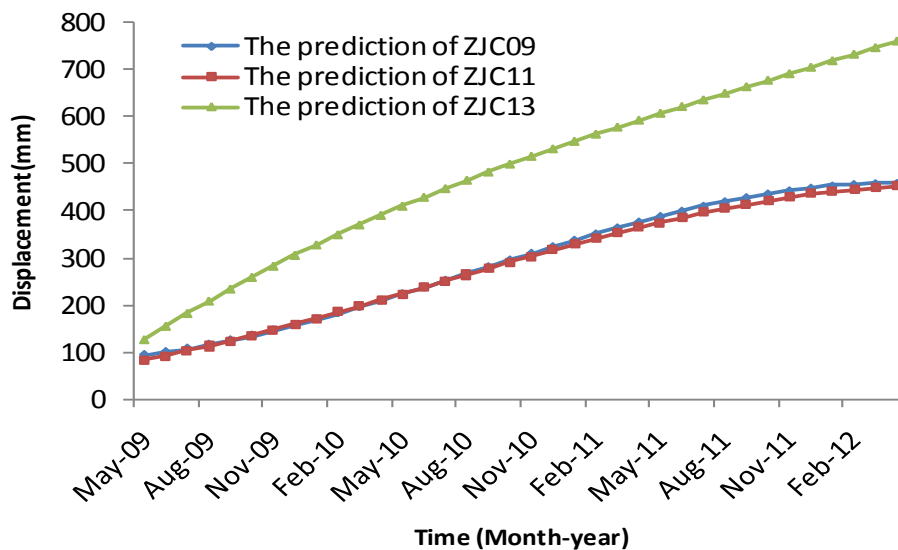


FIGURE 4 The curve of trend term displacement prediction

4.2.2 Displacement prediction of random term

Random term of displacement can be gained after eliminating the trend term from the accumulative displacement, when time series additive model is adopted. However, it is a complex nonlinear time series influenced by multiple factors. Hence, BP neural network mentioned above is used to forecast the displacement of random term by selecting rainfall in the current month, rainfall before two months, monthly average reservoir water level and monthly average reservoir water level variation as the input layer. Prediction results are shown in Table 2.

4.2.3 Accumulative displacement prediction

As mentioned above, trend term and random term of landslide displacement can be obtained by least squares curve fitting and BP neural network respectively. Comprehensive displacement prediction can be obtained after adding these two terms. Finally, the effect of

prediction can be tested by comparison with the test sample in advance, which is shown in Fig. 5. Fig.5 shows that the changes of the comprehensive displacement prediction and accumulative monitoring displacement of landslide are consistent, reaching the expected requirements of prediction accuracy. Here, the smoothing in the moving average method is used to extract trend term and the curve fitting method is used for polynomial fitting. The development trend of landslide displacement mainly depends on trend term displacement of landslide. Therefore, the prediction of trend term displacement by using curve fitting method and smoothing in the moving average method can reduce mutation of displacement curve caused by external factors. Predicting displacement random item is a process to revise accumulative displacement trend of landslide according to a variety of stochastic factors in nature. The BP neural network enables the random term displacement to response to the change of external influence factors by sample training and learning, and thus for reaching the self-organizing, adaptive dynamic prediction.

TABLE 2 Calculation of random term displacement of Liangshuijing Landslide

Time	ZJC09			ZJC11			ZJC13		
	Random term	Prediction	relative error	Random term	Prediction	relative error	Random term	Prediction	relative error
May-09	-4.8	-4.5	0.3	-4.5	-4.9	-0.4	-29.0	-28.5	0.5
Jun-09	0.9	0.3	-0.6	7.5	6.3	-1.2	33.6	27.3	-6.3
Jul-09	40.8	36.8	-4.0	46.5	42.6	-3.9	128.1	118.7	-9.4
Aug-09	49.3	48.4	-0.9	52.9	52.2	-0.7	114.8	116.1	1.3
Sep-09	41.0	41.8	0.8	42.0	43.1	1.1	131.6	129.9	-1.7
Oct-09	37.8	38.2	0.4	39.5	39.8	0.3	115.0	116.7	1.7
Nov-09	48.1	47.0	-1.1	50.6	49.4	-1.2	107.3	108.1	0.8
Dec-09	53.0	52.5	-0.5	52.7	52.4	-0.3	93.3	94.7	1.4
Jan-10	70.8	69.0	-1.8	74.4	72.2	-2.2	114.6	112.5	-2.1
Feb-10	81.0	79.9	-1.1	80.0	79.4	-0.6	128.4	127.0	-1.4
Mar-10	90.9	89.9	-1.0	92.4	91.2	-1.2	122.3	123.0	0.7
Apr-10	103.9	102.6	-1.3	97.2	96.7	-0.5	107.5	109.0	1.5
May-10	109.9	109.2	-0.7	106.4	105.4	-1.0	110.4	110.1	-0.3
Jun-10	99.9	100.9	1.0	93.3	95.1	1.8	99.8	100.9	1.1
Jul-10	92.7	93.5	0.8	90.1	90.5	0.4	106.1	105.4	-0.7
Aug-10	85.3	86.1	0.8	81.4	82.3	0.9	103.0	103.4	0.4
Sep-10	77.0	77.8	0.8	70.3	71.4	1.1	93.1	94.1	1.0
Oct-10	62.5	64.0	1.5	62.9	63.7	0.8	91.6	91.8	0.2
Nov-10	79.4	77.7	-1.7	64.3	64.1	-0.2	86.8	87.3	0.5
Dec-10	92.3	91.0	-1.3	68.7	68.3	-0.4	83.5	83.9	0.4
Jan-11	84.0	84.9	0.9	60.8	61.6	0.8	74.6	75.5	0.9
Feb-11	70.6	72.0	1.4	50.5	51.6	1.1	81.8	81.1	-0.7
Mar-11	61.0	62.0	1.0	50.0	50.1	0.1	84.2	83.9	-0.3
Apr-11	48.5	49.8	1.3	39.0	40.2	1.2	89.1	88.5	-0.6
May-11	36.6	37.9	1.3	62.0	59.7	-2.3	65.9	68.3	2.4
Jun-11	25.4	26.6	1.2	51.5	52.5	1.0	51.6	53.1	1.5
Jul-11	53.4	50.6	-2.8	41.5	42.5	1.0	75.4	73.0	-2.4
Aug-11	43.8	44.8	1.0	32.1	33.1	1.0	67.8	68.5	0.7
Sep-11	35.0	35.9	0.9	23.4	24.3	0.9	93.6	91.0	-2.6
Oct-11	27.2	28.0	0.8	15.3	16.2	0.9	79.8	81.2	1.4
Nov-11	20.4	21.1	0.7	16.2	16.1	-0.1	66.0	67.4	1.4
Dec-11	14.7	15.3	0.6	25.3	24.3	-1.0	56.3	57.3	1.0
Jan-12	10.3	10.8	0.5	28.1	27.8	-0.3	42.6	44.0	1.4
Feb-12	7.1	7.4	0.3	48.4	46.3	-2.1	85.5	81.2	-4.3
Mar-12	129.3	117.1	-12.2	98.8	93.7	-5.1	166.7	158.6	-8.1
Apr-12	128.9	128.9	0.0	95.8	96.1	0.3	162.0	162.5	0.5

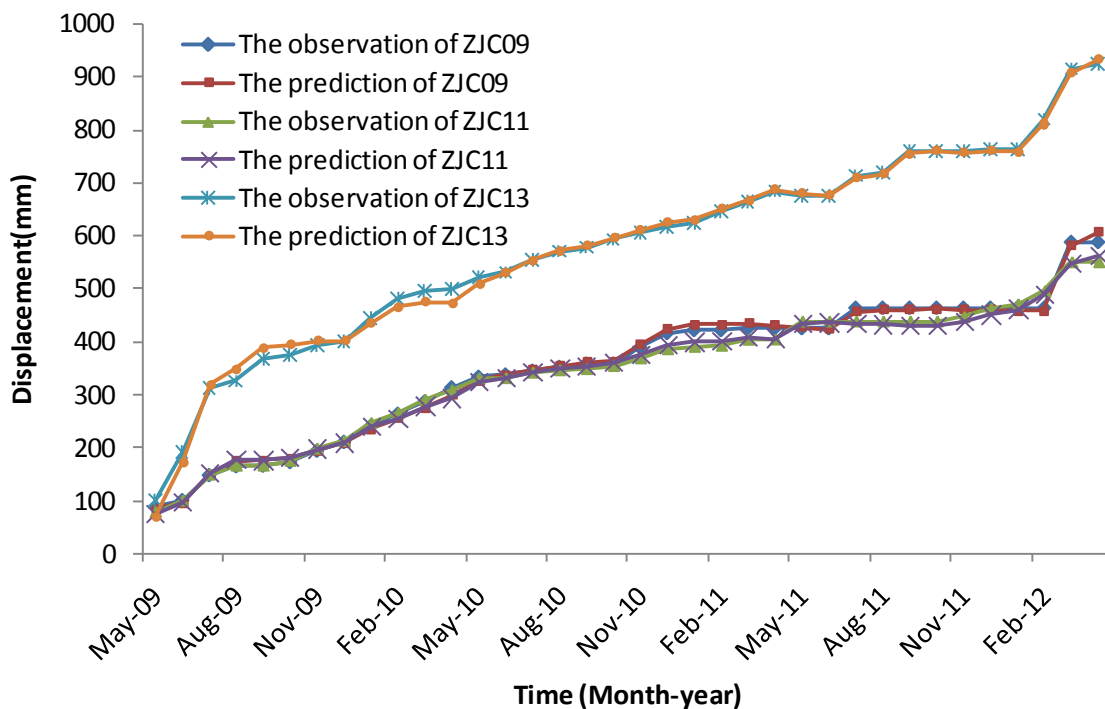


FIGURE 5 The curve of the observation and prediction accumulative displacement

5 Conclusions

This paper takes Liangshuijing landslide as a study case that relates physical movement process of landslide to external influence factors. Accumulative displacement of Liangshuijing Landslide can be divided into trend term and random term. The changes of rainfall and reservoir water level, which both have great effects on landslide, are selected for random term of displacement prediction when using time series additive model. The combination of mathematical function and neural network organically is applied to establish corresponding prediction model, which obtains good prediction effect. The conclusions can be drawn as follows:

(1) The dominant influence factors of Liangshuijing landslide displacement are in the decreasing sequence: cumulative rainfall of anterior two months> rainfall of current month> the average reservoir level of current month>reservoir level fluctuation of current month.

References

- [1] Ma S S, Wang Z Z, Zhang M 2004 Study on stability of landside of reservoir area *Rock and Soil Mechanics* **25**(11) 1837-40 (*In Chinese*)
- [2] Jibson R W, Keefer D K 1989 Statistical analysis of factors affecting landslide, distribution in the New Madrid seismic zone, Tennessee and Kentucky *Engineering Geology* **27**(1) 509-42
- [3] Lee S, Jasmi Abdul Talib 2005 Probabilistic landslide susceptibility and factor effect analysis *Engineering Geology* **47**(7) 982-90
- [4] Guo D, Hamada M 2013 Qualitative and quantitative analysis on landslide influential factors during Wenchuan earthquake: A case study in Wenchuan County *Engineering Geology* **152**(1) 202-9
- [5] He K Q, Yang J B, Wang S J 2005 Analysis of dynamic factors of debris landslide by means of the model of quantitative theory - using the Xintan landslide, China, as an example *Environmental geology* **48**(6) 676-81
- [6] Keefer D K, Wilson R C, Mark R K 1987 Real-time landslide warning during heavy rainfall *Science* **238** 921-5
- [7] Hayashi S, Komamura F, Park B 1988 On the forecast of time to failure of slope (II)-approximate forecast in the early period of the tertiary creep *Jpn Landslide Soc* **23** 1-16
- [8] Fukuzono T 1990 Recent studies on time prediction of slope failure *Landslide News* **4**(9) 9-12
- [9] Mayoraz F, Vulliet L 2002 Neural networks for slope movement prediction *Int J Geomech* **2** 153-73
- [10] Coe J A, Ellis W L, Godt J W 2003 Seasonal movement of the Slumgullion landslide determined from Global Positioning System surveys and field instrumentation, July 1998-March 2002 *Eng Geol* **68** 67-101
- [11] Wang F W, Wang G, Sassa K 2005 Displacement monitoring and physical exploration on the Shuping Landslide reactivated by impoundment of the Three Gorges Reservoir, China *Risk Analysis and Sustainable Disaster Management: proc. 1st General Assembly of the International-Consortium-on-Landslides (Washington DC, USA, 12-14 October 2005)* ed Sassa, K; Fukuoka, H; Wang, F: Berlin pp 313-9
- [12] Wu Y P, Teng W F, Li Y W. 2007 Application of grey-neural network model to landslide deformation prediction *Chin J Rock Mech Eng* **26**(3) 632-6 (*In Chinese*)
- [13] Ferentinou M D, Sakellariou M G 2007 Computational intelligence tools for the prediction of slope performance *Computers and Geotechnics* **34** 362-84
- [14] Biswajeet P, Saro L 2010 Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models *Environ Earth Sci* **60** 1037-54
- [15] Du Juan, Yin Kunlong, Suzanne Lacasse 2013 Displacement prediction in colluvial landslides, Three Gorges Reservoir, China *Landslides* **10** 203-18
- [16] Avinash K G, Ashamanjari K G 2011 Landslide susceptibility modeling of Aghanashini River catchment in Western Ghats of Uttara Kannad district, Karnataka, India *Nature Environment and Pollution Technology* **10**(2) 251-4
- [17] Zhao P D, Hu W L, Li Z J 1983 *Statistical Prediction of Mineral Deposit* Geological Publishing House: Beijing (*In Chinese*)
- [18] Xiang D J, Li H W 2005 *Applied Multivariate Statistical Analysis* China University of geosciences. Press:China-Wuhan (*In Chinese*)
- [19] Zhang T T, Yan E C 2012 Landslide deformation analysis based on factor analysis *Journal of Yangtze River Scientific Research Institute* **29**(4) 21-5 (*In Chinese*)
- [20] Sun X R, Ma F H, Zhang S 2012 Model of Dam Deformation Safety Monitoring Based on Factor Analysis *Water Resources and Power* **30**(4) 34-7 (*In Chinese*)
- [21] Miao H B, Yin K L, Xu F 2010 Comprehensive evaluation on multiple predictions of the landslide displacements based on component analysis *Journal of Wuhan University of Technology* **32**(19) 65-70 (*In Chinese*)
- [22] Bi Q L, Liu X, Xu B 2010 Application of factor analysis to analysis of seepage bypass dam *Water Resources and Power* **28**(5) 62-5 (*In Chinese*)
- [23] Popescu F A, M Elia G 2011 Prediction of time to slope failure: a general framework *Environmental Earth Sciences* **66**(1) 245-56 (*In Chinese*)
- [24] Lin L S, Feng X T, Bai S W 2002 Application of artificial neural network to prediction of sliding slope *Rock and Soil Mechanic* **23**(4) 508-10 (*In Chinese*)
- [25] Miu H B, Yin K L, Chai B 2009 Landslide deformation prediction Based on non-stationary time series analysis *Geological Science and Technology Information* **28**(4) 107-11
- [26] Cheng X J 2005 *Neural Network and Its Application* National Defense Industry: Beijing (*In Chinese*)
- [27] Zhang G R 2005 Spatial prediction and real-time warning of landslides and it's risk management based on WebGIS *Doctoral Dissertation China University of Geosciences, Wuhan,China (In Chinese)*
- [28] Guidicini G, Iwasa Y 1977 Tentative correlation between rainfall and landslides in a humid tropical environment *Bulletin of IAEG, Prague* **16** 13-20

Authors	
	<p>Y P Wu, born in October, 1971, Haining City, Zhejiang Province, P.R. China</p> <p>Current position, grades: the Professor of School of Engineering, China University of Geosciences, China. University studies: She received her M.S degree and D.S degree in Geological engineering from the faculty of Engineering in China University of Geosciences. Scientific interest: Her research interest fields include the mechanism of geology disaster, geology disaster prevention and its forecasting and prediction technology. Publications: more than 50 papers published in various journals. Experience: She has teaching experience of 20 years, has completed more than 40 scientific research projects.</p>
	<p>Q X Zhang, born in September, 1989, Chengdu City, Sichuan Province, P.R. China</p> <p>Current position, grades: a Master student of School of Engineering, China University of Geosciences, China University studies: She received her B.S degree in Geological engineering from the faculty of Engineering in Lanzhou University, China. Scientific interest: Her research interest fields include the mechanism of geology disaster, geology disaster forecasting and prediction technology. Publications: 4 papers published in various journals. Experience: She has completed 3 scientific research projects.</p>
	<p>G Zh Ou, born in May, 1989, Jinzhou City, Hubei Province, P.R. China</p> <p>Current position, grades: a Master degree candidate of School of Engineering, China University of Geosciences, China. University studies: He received his B.S degree in Geological engineering from the faculty of Engineering in China University of Geosciences. Scientific interest: His research interest fields are the mechanism of geology disaster, stability of geotechnical engineering . Publications: 4 papers published in various journals. Experience: he has completed more than 3 scientific research projects.</p>
	<p>X G Fan, born in October, 1989, Anyang City, Henan Province, P.R. China</p> <p>Current position, grades: a Master student of School of Engineering, China University of Geosciences, China University studies: He received his B.S degree in Geological engineering from the faculty of Engineering in China University of Geosciences. Scientific interest: His research interest fields are the mechanism of geology disaster, stability of geotechnical engineering. Publications: No paper published in various journals. Experience: No completed scientific research projects.</p>
	<p>J H Zhou, born in August, 1971, Zhouping County, Guangxi Province, P.R. China</p> <p>Current position, grades: the lecturer of College of Civil Engineering and Architecture, Guilin University of Technology, China. University studies: She received her B'S degree in Geological Engineering from China University of Geosciences and M'S degree in Disaster Prevention and Mitigation Engineering from Guilin University of Technology. Scientific interest: Her research interest fields include mechanism of unsaturated soil and special soil such as expansive soil. Publications: more than 10 papers published in various journals. Experience: She has teaching experience of 20 years, has completed more than 5 scientific research projects.</p>

Study on the Decision Value of Analysts' Recommendations

Liu Wanli*

Business Administration School of Henan University, Kaifeng City, Henan Province, China, 475004

Received 8 January 2014, www.tsi.lv

Abstract

This paper documents a relationship between analysts' recommendations and the stock price reaction in China. Using a new methodology that combines the event of stock dividends and transfer of reserves to common shares, the author provides evidence of the decision value of analysts' recommendations that is different from the mature market. The results show that analysts' cumulative rating values positively relate to the cumulative abnormal returns. Favourable ratings result in the lower cumulative abnormal returns. The cumulative number of analyst rating agencies negatively relates to the cumulative abnormal returns. In general, analysts' information does not bring abnormal returns for investors.

Keywords: Analysts' Recommendations, Rating, Stock Dividends and Transfer of Reserves to Common Shares

1 Introduction

The security analyst contributes to the improvement of the response speed of stock price to the information [1-2]. Ref [3] found that most investors lack the time, skills, sources of information and the ability to account for the financial statements. Therefore, analysts' professional ability becomes one of the main ways to make the accounting information more effectively reflect the stock price, and analysts process information efficiently, which is helpful to enhance the efficiency of the stock price reaction. China holds the world's largest number of stock investors, but many of them do not have investment knowledge and they are desirous to find out "news" in order to obtain abnormal returns in the stock. Common recommendation columns of financial media become their convenient way to get the message. Their investment decisions are heavily influenced by all kinds of recommendations. If a large number of investors believe these recommendations and follow analysts' recommendations to trade, these investors' transactions are bound to influence the returns of recommended stocks. The purpose of this article is to investigate whether analysts' recommendations have the decision value to investors in the Chinese securities market.

Dividend information is the principal financial information that the company disclosed, and also is one of the sources of information that the security analyst can rely on. Before or after the annual report is released, stock dividends and transfer of reserves to common shares are the subjects of speculation and analysts also depend on that information to evaluate, recommend stocks timing. Financial information disclosed by companies is interpreted by security analysts and then released to the community in the name of the experts' recommendations, which is the most intuitive information

to the stock investors. Analysts' recommendations could affect investors' investment decisions, and investors' stock exchange is bound to affect the market returns of recommended stocks.

Using a new methodology that combines the event of stock dividends and transfer of reserves to common shares, the paper examines the decision value of analysts' recommendations to provide a reference for investment decision-making of investors. The remainder of the article is organized as follows. Section II lays out the hypotheses; Section III presents the data and methodology; Section IV reports and discusses the empirical results; Section V concludes.

2 The literature review and hypotheses

2.1 LITERATURE REVIEW

Scholars' conclusions on the decision value of analysts' recommendations are not consistent in term of the abnormal returns, which investors could obtain according to the analysts' recommendations. By studying the decision value of recommendations through the market reaction of the recommendation, the literature mainly concentrates on two aspects: (1) studies on the recommendations do not distinguish the recommended strength and grade. In 1933, Cowles, who study the stock recommendations firstly, found that analysts' recommendations cannot get more returns than that of the market benchmark index [4-6]. Through studying on the "Value Line" the stock market reaction, ref [7] found that according to the "value line" stocks recommended to operate, you can get abnormal returns; Ref [8-9] also found that without considering transaction costs and information costs, the operation according to the value Line can obtain abnormal returns; (2) the studies about

* *Corresponding author* e-mail: Liuwanli2007@sina.com

recommendations distinguish the different intensity and the number of analysts on the stocks recommended. Ref [10] found that "buy" portfolio has significantly positive abnormal returns. Other scholars have also presented evidence that the analysts' recommendations in the short-term can obtain abnormal returns [11]. Ref [12] reported that the recommendation reaction of a single agency was significantly greater than that of its multi-agency, Beneish whose research is about "the Heard on the Street" column of the Wall Street Journal got a similar conclusion [13].

China's analyst industry was established later, and the work of decision value of analysts' recommendations is late. Nonetheless, they still made a useful research results. Restricted by the availability of data, the attention focuses on two aspects: (1) studies mainly based on financial newspaper media's stocks recommended, and there is no distinction between the recommended strength and grade. Ref [14] found that before stocks recommendations issued the abnormal returns appeared positive. After publication, the abnormal returns were negative; Ref [15-16] obtained the similar conclusions. Ref [17-18] reported the evidence that people following analysts' recommendations in the short term can obtain abnormal returns; (2) few studies have distinguished the strength or the number of analysts. Ref [19] noted that the returns of the "strong buy" rating on the stock in the two intervals are lower than that of a "strong sell" rating. Ref [20] showed that, along with the increase in the number of security analysts, the company-level information is more likely to be included in the stock price, making the price drop synchronization. Ref [21] found that the higher the concentration of stock analysts after the base date of the announcement is, the higher the negative abnormal returns are, and the greater breadth of information dissemination is. Ref [22] observed the returns of the companies that have analysts following are lower than that of companies without analysts.

However, few studies have been done on the decision value of analysts' recommendations in China. Chinese existing literature in the following areas is to be discussed and deepened: (1) no systematic research has been dedicated to five ratings. The references on stocks recommended discussed above are based on the stocks recommended data of print media as samples that are all strong buying or buy ratings; (2) there are few studies on the number of analysts that have the impact on the stock, but different ratings would produce obviously different market reaction in different directions. So, it is necessary to use the comprehensive value of different analysts; (3) the work based on specific events related research on methodology has not been taken into account in the mentioned above studies, which use the data in a certain period of analysts stocks recommended in events selection. Giving full consideration to the existing research results, based on stock dividends and transfer of reserves to common shares as the event, this paper tests the impact of analysts' recommendations on the

cumulative abnormal returns to study the decision value of recommendations using the quantified ratings and the number of rating agencies during the event.

2.2 HYPOTHESES

The paper uses quantified ratings from the database. Low rating means high returns expected. As the analyst's rating based on the performance of stocks in the following period of time compared to the performance of the market to divide, if the analyst's rating is accurate, the lower the rating value is, the higher the abnormal return should be. Chinese existing researches use basically data recommended in the financial newspaper as samples, which can be regarded as strong buy rating studies. We can find that after the stocks recommended information issued, and it had significant negative abnormal returns; "Strong buy" rating in two intervals is lower than "strong sell" rating on the stock return [19]; no evidence can prove analyst's recommendations on long-term profitability. In our sample, analysts' rating values mainly concentrate within the range of 1 to 3, and the strong sell or sell rating values were very little. Within the 41 days of the plan event, strong buy ratings account for 11.08% of the sample, buy ratings represent 69.65% of the sample, neutral ratings account for 19.12%, sell ratings account for 0.15%. There is no strong sell rating in the sample. The rating is mainly to buy. Sell and strong sell rating just remind investors to avoid risk and reduce losses, and would produce the pressure of market selling price, which might damage the holding investors' interest. Sell and strong sell rating show high costs and risks, and cannot conducive to the maintenance of relations between analyst and management, so that the analyst might lose first-hand information channels from the listed company to obtain, and also affect the investment banking division to undertake the brokerage business. Pressure from the company's management and investment banking division would make the analysts tend to rate the "strong buy" and "buy", so that "strong buy" and "buy" rating by analysts might be affected by stakeholders, and the accuracy of their rating is low. According to scholars and our findings in the rating distribution of the sample, the following hypothesis is proposed:

Hypothesis 1: There is a positive correlation between the rating values and cumulative abnormal returns during the plan event.

The researchers concluded that if the concentration of the stocks is higher, the negative abnormal return is higher after the date of the publication. The more the number of analysts is, the higher the degree of information of the stock is, and the abnormal return of the stock split announcement and the degree of stock information changed inversely. Thereafter, the following hypothesis is proposed:

Hypothesis 2: There is a negative correlation between the number of rating agencies and the cumulative abnormal returns during the plan event.

3 Study design

3.1 METHODOLOGY

Considerable prior literature on the stock price reaction of the analysts' recommendations usually uses the event-study methodology. Considering prior studies and the needs, our core methodology is the event-study methodology. In this paper, using stock dividends and transfer of reserves to common shares plan announced as the event date to discuss the effect of analysts' recommendations on stocks abnormal returns during the event. Depending on the research on dividend income before and after the announcement [23], the event

window starts from 30 days before the plan announcement to 10 days after the plan announcement.

The econometric model used in empirical research is the merger data model, the general form as follows:

$$y_{it} = \beta_{it}^T x_{it} + \varepsilon_{it}, i=1, \dots, N, \tag{1}$$

wherein, y_{it} is the dependent variable, x_{it} , β_{it} are $K \times 1$ column vector, respectively as the vector of variables and coefficient vector, ε_{it} is random disturbance, T represents the matrix transpose, N is the number of sectional units (individual), t for different year.

TABLE 1 Definitions of variables

Variable nature	Variable Name	Variable Sign	Variable Description
Explained variables	The CAR during the 41 days of the plan event	CMCAR (-30,10) TMCAR (-30,10)	The current market value weighted The total market value weighted
	The CAR during the first 31 days of the plan event	CMCAR (-30,0) TMCAR (-30,0)	The current market value weighted The total market value weighted
	The cumulative rating value	CRATING1	The cumulative rating value during the 41 days
		CRATING2	The cumulative rating value during the 31 days
Explanatory variables	The cumulative number of rating agencies	CRIN1 CRIN2	The cumulative number of rating agencies during the 41 days The cumulative number of rating agencies during the 31 days
	Scale	SIZE	The natural logarithm of the total share capital
The control variables	Send transfer	SZ	Bonus per share and the total number conversed
	Cash dividend	PAI	Cash dividends amounts per share
	Profitability	EPS	Basic earnings per share
	Growth	SG	Revenue growth rate
	Industry	IND	SEC industry dummy variables

3.2 VARIABLE SELECTIONS

3.2.1 Explained variables.

The explained variables are the cumulative abnormal returns(CAR) during the 41 days of the plan event and the cumulative abnormal returns(CAR) during the first 31 days of the plan event, that is CMCAR (-30, 10), TMCAR (-30, 10), CMCAR (-30, 0) and TMCAR (-30, 0). Analysing the cumulative abnormal returns during the event, we find that the lowest cumulative abnormal returns(CAR) (-30, 0) also accounts for 83% of event returns or more during the three years, so these two cumulative abnormal returns are tested simultaneously. Each cumulative abnormal return has two types of circulation, the current market value weighted and total market value weighted.

3.2.2 Explanatory variables

(1) The cumulative rating value and cumulative number of rating agencies during 41 days of the event, denoted as CRATING1 and CRIN1.

$$CRATING1 = \sum_{t=-30}^{10} RATING, (t=-30, \dots, 0, \dots, 10), \tag{2}$$

$$CRIN1 = \sum_{t=-30}^{10} RIN, (t=-30, \dots, 0, \dots, 10), \tag{3}$$

wherein, RATING represents rating value, RIN represents the number of rating agencies.

(2) The cumulative rating value and cumulative number of rating agencies during the first 31 days of the event, denoted as CRATING2 and CRIN2.

$$CRATING2 = \sum_{t=-30}^0 RATING, (t=-30, \dots, 0), \tag{4}$$

$$CRIN2 = \sum_{t=-30}^0 RIN, (t=-30, \dots, 0), \tag{5}$$

3.2.3 The control variables

According to the literature, controlling the other factors may affect the cumulative abnormal return during the plan event. The control variables are size, send transfer, cash dividends, profitability, growth, industry. Table 1 describes the definition of specific variables.

3.3 SAMPLE SELECTION

According to the practice of scholars [24], excluding financial companies and missing data companies, there

have 691 observations in the sample from 2008 to 2010; the market experienced a fall, rise and relatively stable fluctuation in the period 2008-2010, in which interval that avoid the impact of market trends on abnormal returns and the measure of stock price reaction of stocks recommended to a certain extent. Data were available from CSMAR and WIND.

4 Empirical results

4.1 EMPIRICAL RESULTS DURING THE 41 DAYS OF THE EVENT

Firstly, the paper tests the impact of analysts' cumulative rating values on the cumulative abnormal returns of stocks during the 41 days of the event. Secondly, the paper tests the impact of the cumulative number of analyst rating agencies on the cumulative abnormal returns of stocks during the 41 days of the event. Finally, the paper incorporates the analysts' cumulative rating values and the number of rating agencies in the model to test the impact on the cumulative abnormal returns of stocks during the 41 days of the event. According to Table 2, the probability accompanied by an F-test corresponding statistic is less than 5%, whose results show that the coefficients in all the models are statistically highly significant.

Table 2 shows the estimated coefficients of explanatory variable CRATING1 are 0.0007 and 0.0008 in the regression model 1 and 2, and both of them are significantly positive at the level of 5%. The different rating values exist systematic difference, compared with the low value of the cumulative rating, the high cumulative rating value can produce the higher cumulative abnormal returns during the 41 days of the event, and the low cumulative rating value corresponds to

the lower cumulative abnormal returns, which support our hypothesis 1.

Consistent with negative abnormal returns on analysts' recommendations after the announcement, it can be considered relative to the mean of abnormal returns of the sample combination that announce stock dividends and transfer of reserves to common shares, analysts' low rating values make the recommended stocks produce the negative abnormal returns. Positive correlation between the rating value and cumulative abnormal returns, perhaps the reason is the lower rating value corresponding to lower cost and risk. We also can explain from the analysts' independence. We can consider that the cost and the risk of the negative rating are higher than the positive rating; therefore, it requires higher abnormal returns to compensate.

Table 2 presents the estimated coefficients of explanatory variable CRIN1 that in the regression model 3 and 4 are -0.0001, and are significantly negative at the level of 5% and 10% respectively, which support our hypothesis 2. The larger cumulative number of rating agencies is, the lower the cumulative abnormal returns of the plan are during the event. The results show that the larger number of rating agencies is, the greater the breadth of information released by the company is, the more concern that the company can obtain. Investors are not susceptible to a single analyst opinion. It will reduce the wrong investment risk that misled by the individual analyst's opinion, thus contributing to investors' understanding of publicly listed companies in the information transmitted and other operating conditions and other aspects of information, and reduce the degree of asymmetry of information between listed companies and investors, and then reducing the cumulative abnormal returns.

TABLE 2 Empirical results for 41 days

	CMCAR (-0,10)	TMCAR (-0,10)	CMCAR (-0,10)	TMCAR (-0,10)	CMCAR (-0,10)	TMCAR (-,10)
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
CRATING1	0.0007** (2.2019)	0.0008** (2.3615)			0.0006* (1.7887)	0.0007** (1.9899)
CRIN1			-0.0001** (-2.0589)	-0.0001* (-1.8903)	-0.0001 (-1.6102)	-0.0001 (-1.4017)
SIZE	-0.0238*** (-3.0169)	-0.0254*** (-3.1077)	-0.0175* (-1.9574)	-0.0199** (-2.1394)	-0.0170* (-1.9030)	-0.0193** (-2.0801)
SZ	0.1671*** (7.8135)	0.1719*** (7.7594)	0.1627*** (7.6555)	0.1668*** (7.5686)	0.1671*** (7.8253)	0.1719*** (7.7669)
PAI	0.0737 (1.5099)	0.0497 (0.9838)	0.0803 (1.6404)	0.0559 (1.1017)	0.0794 (1.6246)	0.0549 (1.0835)
EPS	-0.0493*** (-2.8855)	-0.0401** (-2.2675)	-0.0474*** (-2.7155)	-0.0399** (-2.2018)	-0.0417** (-2.3510)	-0.0332* (-1.8098)
SG		0.0008 (0.1761)	-0.0001 (-0.0293)	0.0008 (0.1712)	-0.0004 (-0.0823)	0.0005 (0.1127)
IND	Controlled	Controlled	Controlled	Controlled	Controlled	Controlled
Constant	0.4116** (2.4618)	0.4427** (2.5564)	0.3753** (2.1346)	0.4281** (2.3482)	0.3029* (1.6816)	0.3446* (1.8463)
F Value	5.6645	5.7095	5.6232	5.5741	5.5079	5.5101
P Value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: * significant at the 10% level (two-tailed).

** significant at the 5% level (two-tailed).

*** significant at the 1% level (two-tailed).

Table 2 presents the estimated coefficients of explanatory variable CRIN1 that in the regression model 3 and 4 are -0.0001, and are significantly negative at the level of 5% and 10% respectively, which support our hypothesis 2. The larger cumulative number of rating agencies is, the lower the cumulative abnormal returns of the plan are during the event. The results show that the larger number of rating agencies is, the greater the breadth of information released by the company is, the more concern that the company can obtain. Investors are not susceptible to a single analyst opinion. It will reduce the wrong investment risk that misled by the individual analyst's opinion, thus contributing to investors' understanding of publicly listed companies in the information transmitted and other operating conditions and other aspects of information, and reduce the degree of asymmetry of information between listed companies and investors, and then reducing the cumulative abnormal returns.

In the model 5 and 6, incorporating the two explanatory variables CRATING1, CRIN1 and the control variables, the results show that the cumulative rating value is statistically significantly positive, but the cumulative number of rating agencies is negative, statistically insignificant.

Table 2 reports that the estimated coefficients of the SIZE in all models are significantly negative symbol in line with expectations, indicating that the SIZE is smaller, the cumulative abnormal returns during the 41 days of the event is higher. The estimated coefficients of SZ in all models are positive at the level of 1% statistical significance, the symbol in line with expectations, indicating that the proportion SZ on the plan has a significant positive effect on cumulative abnormal returns during the 41 days of the event. The estimated coefficients of profitability in all models are significantly negative. The estimated coefficients of cash dividends and the growth are not significant in all models.

TABLE 3 Empirical results for 31 days

	CMCAR (-30,0)	TMCAR (-30,0)	CMCAR (-30,0)	TMCAR (-30,0)	CMCAR (-30,0)	TMCAR (-30,0)
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
CRATING2	0.0012*** (3.5206)	0.0013*** (3.3923)			0.0012*** (3.1942)	0.0012*** (3.1087)
CRIN2			-0.0001* (-1.8773)	-0.0001* (-1.6619)	-0.0001 (-1.1737)	-0.0001 (-0.9790)
SIZE	-0.0212*** (-3.2430)	-0.0236*** (-3.3714)	-0.0180** (-2.4232)	-0.0209*** (-2.6233)	-0.0171** (-2.3151)	-0.0199** (-2.5187)
SZ	0.1331*** (7.5242)	0.1369*** (7.2302)	0.1270*** (7.1697)	0.1305*** (6.8852)	0.1332*** (7.5312)	0.1369*** (7.2339)
PAI	0.0375 (0.9281)	0.0225 (0.5206)	0.0423 (1.0374)	0.0270 (0.6196)	0.0407 (1.0060)	0.0254 (0.5859)
EPS	-0.0144 (-1.0156)	-0.0052 (-0.3445)	-0.0184 (-1.2697)	-0.0101 (-0.6529)	-0.0098 (-0.6660)	-0.0011 (-0.0708)
SG	0.0025 (0.6775)	0.0032 (0.8219)	0.0026 (0.7118)	0.0034 (0.8616)	0.0023 (0.6232)	0.0030 (0.7761)
IND	Controlled	Controlled	Controlled	Controlled	Controlled	Controlled
Constant	0.3433** (2.4858)	0.3915*** (2.6488)	0.3847*** (2.6292)	0.4444*** (2.8386)	0.2775* (1.8621)	0.3327** (2.0856)
F Value	6.3892	6.3063	5.7862	5.7121	6.1145	6.0088
P Value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: * significant at the 10% level (two-tailed).

** significant at the 5% level (two-tailed).

*** significant at the 1% level (two-tailed).

4.2 EMPIRICAL RESULTS DURING THE FIRST 31 DAYS OF THE EVENT

Same with the above tests on the cumulative abnormal returns during the 41 days of plan event, first, the paper tests the effect of analysts' cumulative rating values and the cumulative number of rating agencies on the cumulative abnormal returns of the 31 days respectively. Then, the paper incorporates the analysts' cumulative rating values and the cumulative number of rating agencies in the model to test their effect on the cumulative abnormal returns during the 41 days of the event. According to Table 3, the probability accompanied by an F-test is less than 1%, and the results show that the

coefficients of all the models are statistically highly significant in the whole.

Table 3 reports that the estimated coefficients of explanatory variable CRATING2 are 0.0012 and 0.0013 in the regression model 1 and 2, and both of them are significantly positive at the 1% level. The estimated coefficients are larger than the coefficients of CRATING1 in Table 2. Cumulative rating values have more effect on the cumulative abnormal returns during the first 31 days. Different rating values have systematic differences. The low values correspond to the lower cumulative abnormal returns, which support our hypothesis 1.

Table 3 presents that the estimated coefficients of explanatory variable CRIN2 are -0.0001 in the model 3 and 4, and are significantly negative at the level of 10%. The results support our hypothesis 2. The larger number of rating agencies is, the lower cumulative abnormal returns are during the plan event. Combining statistics on the number of rating agencies during the plan event, a relatively small number of rating agencies before the date of the event on reducing the degree of information asymmetry is weaker than a big number of rating agencies. The more rating agencies represent the higher degree of stock information, more help to reduce information asymmetry, decreasing the corresponding cumulative abnormal returns.

In the model 5 and 6, incorporating the two explanatory variables CRATING2, CRIN2 and the control variables, the results show that the cumulative rating values are significantly positive, but the cumulative number of rating agencies is negative, statistically insignificant.

Table 3 indicates the SIZE is still a factor influencing the cumulative abnormal returns during the first 31 days, the symbol of which is negative. The estimated coefficients of SZ are negative, and the SZ has a significant impact on the cumulative abnormal returns during the plan event. Nevertheless, the influence of profitability on the cumulative abnormal returns is no longer significant. The estimated coefficients of cash dividends and growth are not yet statistically significant.

5 Conclusions

The paper uses quantified rating values and the number of rating agencies to empirically study the decision value of analysts' recommendations. Security analysts are intermediaries of information gathering and dissemination. Information is the critical factor affecting the decision-making of investor, and analysts' recommendations affect investors' stock trading, and then affecting the market returns of recommended stocks. Considering the market speculation to stock dividends and transfer of reserves to common shares which are also the themes of analysts' recommendations in China's security market, constructing a recommendations combination based on a specific event, the paper tests the

References

- [1] Brennan M J, Jegadeesh N, Swaminathan B 1993 Investment analysis and the adjustment of stock prices to common information *Review of Financial Studies* 6(4) 799-824
- [2] Hong H, Lim T, Stein J 2000 Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies *Journal of Finance* 55(1) 265-96
- [3] Beaver W H 2002 Perspectives on recent capital market research *The Accounting Review* 77(2) 453-74
- [4] Cowles A 1933 Can stock market forecasters forecast? *Econometrica* 1(3) 309-24
- [5] Diefenbach R E 1972 How good is the institutional brokerage research? *Financial Analyst Journal* 28(1) 54-60

effect of analysts' recommendations on stock market returns and their direction. Specifically, analysts' cumulative rating values have significantly positive impact on the cumulative abnormal returns during the 41 days and the first 31 days of the event, and a lower rating released corresponds to lower cumulative abnormal returns. The cumulative number of rating agencies has a significant negative impact on the cumulative abnormal returns during the 41 days and the first 31 days of the event, indicating the larger number of rating agencies the higher degree of stock information is, with the wider information dissemination and the higher degree of concern to the company. Different analysts give different views of recommended stocks based on their information and expertise, which reduce the risk of analysts' misleading ratings by minorities, and then significantly reducing the degree of information asymmetry and the cumulative abnormal returns during the event.

Analysts' recommendations seem no help to investors. Perhaps the main reasons are that (1) the cost and the risk of the negative rating are higher than the positive rating. Pressure from the company's management and investment banking division would make the analysts tend to rate the "strong buy" and "buy"; (2) the overall quality of the listed companies is not high, and investors prefer to short-term operation. The analyst may also meet the investor to give a rating based on their own interests in this market environment.

The contribution of this paper is to quantify the rating and construct the sample to test the decision value of the recommendations by using the specific event instead of the specific time period, and it will enrich the analyst's literature research and provide references for investment decision-making of investors. However, the research limitation is that the stock returns might be affected by various factors, analysts' recommendations may be one aspect of affecting the stock market reaction. Despite the measure of market reaction used market-adjusted return to minimize the different market conditions' impact on stock returns, maybe there still exists deviations.

Acknowledgments

This project is funded by Humanities and Social Science Fund of Education Ministry (2009JJD790041).

- [6] Logue E, Tuttle D 1973 Brokerage house investment advice *Financial Review* 8(1) 38-54
- [7] Black F 1973 Yes, Virginia, there is hope: test of the value line ranking system *Financial Analyst Journal* 29(5) 10-4
- [8] Holloway C 1981 A note on testing an aggressive investment strategy using value line ranks *Journal of Finance* 36(3) 711-9
- [9] Copeland T, Mayers D 1982 The value line enigma (1965-1978): a case study of performance evaluation issues *Journal of Financial Economics* 10(3) 289-321
- [10] Groh J C, Lewellen W G, Schlarbaum G G, Lease R C 1979 An analysis of brokerage house securities recommendations *Financial Analyst Journal* 35(1) 32-40
- [11] Bjerring J H, Lakonishok J, Vermaelen T 1983 Stock prices and financial analysts' recommendations *Journal of Finance* 38(1) 187-204

- [12] Liu P, Smith S D, Syed A A 1990 Stock price reactions to the Wall Street journal's securities recommendations *Journal of Financial and Quantitative Analysis* 25(3) 399-410
- [13] Beneish M D 1991 Stock prices and the dissemination of analysts' recommendations *Journal of Business* 64(3) 393-416
- [14] Lin X 2000 Analysis on the Chinese securities advisory body to forecast *Economic Research Journal* 2 56-65 (in Chinese)
- [15] Tang J, Song F M 2002 An analysis of selecting stocks prediction of the recommendation *Collected Essays on Finance and Economics* 1 44-9 (in Chinese)
- [16] Xu Y X, Chen C 2009 Motivation of market reaction on media recommended stocks *Management World* 11 65-73 (in Chinese)
- [17] Zhu B X and Wang Y K 2001 An empirical analysis of the effect of the securities media recommendations *Economic Research Journal* 4 51-7 (in Chinese)
- [18] Ding L, Sun H 2001 Effect of Chinese stock recommendations *Management World* 5 111-6 (in Chinese)
- [19] Huang J, Dong X L 2006 The empirical study on the security analysts ability of stock recommendation and earnings forecasts *Application of Statistics and Management* 6 742-9 (in Chinese)
- [20] Zhu H J, He X J, Tao L 2007 Chinese securities analysts can improve the efficiency of the capital markets *Financial Research* 2 110-21 (in Chinese)
- [21] Xu L P, Liu J H 2008 Study of the recommending effect of the securities analyst on market *Collected Essays on Finance and Economics* 4 70-6 (in Chinese)
- [22] Liu W L 2013 Security analysts' recommendations and the returns of stock dividends *Search* 6 31-3 (in Chinese)
- [23] Liu W L, Xue Z Y 2010 The effect of stock dividends on shareholders' value: evidences from Chinese listed companies *The Theory and Practice of Finance and Economics* 31(3) 53-57 (in Chinese)
- [24] Yang R X 2013 Bibliometrical analysis on the big data research in China *Journal of Digital Information Management* 11(6) 383-90

Authors



Wanli Liu, born in March, 1981, Kaifeng City, Henan Province, P. R. China

Current position, grades: the lecturer of Business Administration School, Henan University, China.

University studies: She graduated from Xiamen University in 2011, received a doctor's degree in management.

Scientific interest: Financial Management, data mining.

Publications: more than 10 papers published in various journals.

Experience: She has teaching experience of 3 years, has presided over 1 scientific research projects.

A dual capacity sourcing model of disruption management for an injured power system

Bao Xing^{1, 2*}

¹ Zhejiang Provincial Committee Party School of CPC, West Wenyi Str. 1000, 311121 Hangzhou, China

² Zhejiang Scientific Development Research Centre, West Wenyi Str. 1000, 311121 Hangzhou, China

Received 1 March 2014, www.tsi.lv

Abstract

Great loss would be caused when power system lost its critical capacity by the impact of extreme events. Disruption management of State Grid Zhejiang Electric Power Company of China (shorted for SGZEPC) suffered in 2008 was firstly investigated, and then a dual sourcing model of regular and expedite capacity during recovery periods is correspondingly presented in this paper. A mathematical model of capacity procurement in a multi recover periods is constructed at the aim of minimizing the disruption cost of injured power system. Three meaningful managerial insights are obtained through sensitivity analysis on key parameters, which is helpful for manager to make decision during the disruption period.

Keywords: disruption management, injured power system, dual capacity sourcing, multi-period

1 Introduction

Power system is one of most important public facilities which highly relies on the safety and steady operation of its critical capacities, such as power transmission network, distribution facilities and power plants. Despite a lot of protection on its safety operation have been done, power system still cannot eliminate every potential threats arose by unexpected events, such as natural disasters, man-made operational defaults and even intentional attacks.

In recent years, serious blackouts were frequently witnessed that power systems lost their critical capacities. For example, State Grid Zhejiang Electric Power Company of China (shorted for SGZEPC) suffered great loss on its operation history, because of a hundred-year frozen rain attacked its transmission network in Jan 2008, and about 75 percent of its power transmission lines were crippled by heavy ice. This unexpected frozen rain caused a serious blackout among Zhejiang Province of China, which caused the consequence of over 20 million people suffered power shortage for over two weeks. Disruption management was started at the first time, Manager of SGZEPC on one hand started recovery process on the injured capacity, on the other hand acquired as much as possible the temporary electricity from its neighbouring provincial partners and mobile generators.

In this paper, we would like to present a multi-period model under the scenario of SGZEPC's disruption management in 2008. The paper is organized as follows: In section 2, the most related literatures to our research are reviewed. In section 3, the basic description and notations of our model are presented according to our

background case. In section 4, the mathematical model and proofs of optimal decision sequences are presented at the aim of minimizing the disruption cost. In section 5, sensitivity analyses on key parameters of our model are given to illustrate the impact on disruption cost. Finally, managerial insights and conclusions are presented in section 6.

2 Literature review

In the past decade, many researchers in the fields of operation management and risk management have done voluminous studies in order to enhance the system's operational robustness in facing fluctuations from both inside and outside [1]. However, some researchers argued that some traditional risk mitigation methods should be reinvestigated under the scenario of operation system (or supply chain) might be attacked by those unexpected events, because of "Snow Ball" effects could cause risks spreading quickly along the whole supply chain from the disrupted node [2]. Losses caused by abnormal disruptions are much more serious than those caused by normal risks (operational risk, seasonal fluctuation risk et al). Different from those traditional risk management researches, disruption management puts more emphasis on risks arose by those unexpected events with great negative impacts and extremely low probabilities. Researches exist in disruption management nowadays can be categorized into two streams. The first stream adopts the methodologies of empirical and framework studies, for one purpose is to verify whether disruption really has negative impact both on company's long-term performance through statistical models or cases [3-4], and

* *Corresponding author* e-mail: goldbxing@gmail.com

for another purpose is to introduce disruption mitigation methods, guidelines and strategies through framework models [5-6]. The second stream attracts the most of researchers in giving mathematic models and simulations in order to give some effective mitigation methods before\during\after the break out of disruption, and much of the works follow the framework of [1].

Once the disruption happens on the critical capacity of power system, there three kinds of mitigation methods in operational level that manager could take [5]. They are lagging demand along the time scale by price coordination [7], increasing temporary capacity by sourcing from suppliers or partners [8-9], and mitigation the cost by production rescheduling [10-11]. Although disruption cost could be greatly cut down, the mitigation methods/policies are always confined by the operation system's public importance. For example, shifting the demand to the future time span may alleviate the current shortage of capacity by increasing the price temporarily [12]. However, this might not be a good idea, because increasing the electricity spending might easily incur the protest or pressure from the public or government [13]. Procuring or sourcing inventory from backup suppliers or capacity-sharing partners are extensively studied in supply disruption management, and how to make the best order-split decisions among different suppliers are mainly focused [14-18]. We are glad to find that dual capacity sourcing described in [14] was still adopted by SGZEPC in 2008 when acquiring temporary electricity during the disruption management. For example, temporary electricity was provided from its neighbouring provincial power system and mobile generators. However, capacity damage are much more serious than capacity shortage, which could causes a sharp fall in the service level and contributes great loss towards the power system. Following this point of view, some researchers paid their attention to generate sub-optimal operational plans/routings by dispatch the residual capacity in order to obtain a satisfied cost as well as save the calculation time [19-20]. However, the assumption of recovered capacity can be only reused at the last period might not be proper according to the practice of SGZEPC, because recovered capacity was gradually put into reoperation during the disruption periods [21]. Thus, we are inspired to make a further extension of [21] under the scenario of power system's injured capacity is gradually recovered.

3 Problem description and notation

According to the practice of SGZEPC in 2008, electricity provision in Zhejiang province is seriously deteriorated due to the injury of critical capacity. SGZEPC started the procedure of disruption management at first time. On one hand, the injured capacity should be recovered to normal state; on the other hand, the service level of electricity provision should be tried best to keep in certain satisfied level by acquiring more temporary electricity from both inside and outside of SGZEPC's operation system. There

were two kinds of capacity that could SGZEPC provide extra electricity. For example, regular capacity was procured form its neighbouring provincial partners through SGZEPC's residual capacity, and the expedite capacity was procured by starting the backup power systems, such as mobile generators. Therefore, the disruption management model could be illustrated in figure 1.

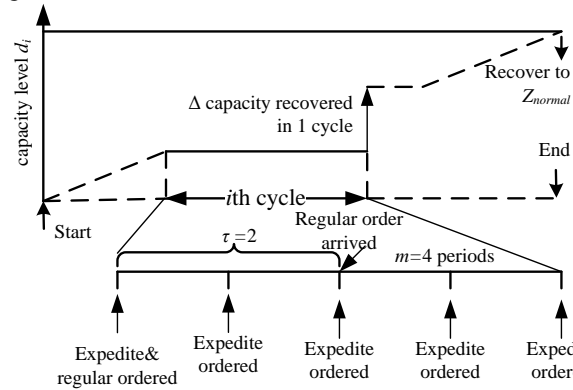


FIGURE 1 Disruption management model when operation capacity is injured

For convenience of creating mathematical model, notations are described as bellow:

Z_{normal} is the power system's planned capacity in normal state.

d_0 is the residual capacity that is still in operation when disruption take place, and $d_0 < Z_{normal}$.

N is the recovery cycles that is needed to recover all the injured capacity to normal state Z_{normal} , and there are m periods in 1 cycle.

r is the recovered capacity in 1 recovery cycle which is bounded by recovery technique, and $\Delta = (Z_{normal} - d_0) / N$ is the recovery up-limit. Further, we assume $\Delta = r$. Recovered capacity need m periods that can be put into re-operation. In other words, recovered capacity can only be reused at every end of recovery cycle.

$P(\Delta)$ is recovery cost function.

U is regular capacity that capacity-sharing partner could provide, and is ordered only once at the beginning of each cycle with a fixed value. U is received after τ periods, therefore, τ is the lead time of regular capacity, where $2 \leq \tau \leq m$.

y is electricity demand in each period, and y is a non-negative random variable with the probability density and distribution function are $f(y)$ and $F(y)$ respectively. Electricity demand is satisfied by the output of system's capacity at the end of each period and unit capacity satisfied unit demand. Unsatisfied demand is backlogged to next period.

The subscript (i, j) is a time indicator which means i^{th} cycle and j^{th} period, and $i = 1, \dots, N$, $j = 1, \dots, m$. Some additional notations are defined as follows:

$z_{i,j}$ is the planned capacity in (i, j) period which is a decision variable, where $0 \leq z_{i,j} \leq d_i$ when $j=1, \dots, m-1$, and $0 \leq z_{i,j} \leq d_i + \Delta$ when $j=m$.

$v_{i,j}$ is the quantity of expedite capacity procurement, which is ordered at the beginning of each period and received at the end of each period, therefore the lead time of $v_{i,j}$ is set to 1 period. Additionally, $0 \leq v_{i,j} \leq V$, where V is the maximum that expedite capacity could be acquired.

$x_{i,j}$ is the net inventory at the beginning of (i, j) period, demand that was not satisfied in last period is backlogged. $x_{i,j}$ is also the state transition variable, and $x_{i,j} = \min\{0, x_{i,j-1}\}$ which means excessive electricity cannot be backlogged into next period.

c_0 is the unit cost of using system's own capacity. c_1 and c_2 are the unit cost of using regular and expedite capacity respectively, the expedite capacity is much more expensive, thus $c_1 < c_0 < c_2$.

k is the unit shortage cost when electricity cannot meet the demand.

α is the manager's risk attitude factor during the disruption periods, which take the form of equation (1). Manager will be more aggressive in using expedite capacity $v_{i,j}$ when α increases. Otherwise, decision maker will be more conservative when α decreases.

$$\alpha = \frac{v_{i,j}}{Z_{normal} - z_{i,j}} \tag{1}$$

4 Mathematic model

Based on the description of section 2, we could construct the cost function $\phi_{i,j}$ in normal state as equation (2), where the first item is operation cost of planned capacity, the second item is electricity shortage cost.

$$\phi_{i,j} = c_0 z_{i,j} + k(z_{i,j} - y)^- \tag{2}$$

It is rather interesting that $\phi_{i,j}$ is obviously a newsvendor model. Thus, we could easily obtained the optimal planned capacity $z_{i,j}^*$ in normal state by minimizing the expectation cost of equation (2), and $z_{i,j}^*$ is regarded as Z_{normal} as equation (3), where $F^{-1}(\cdot)$ is the inverse function of $F(\cdot)$.

$$Z_{normal} = F^{-1}\left(1 - \frac{c_0}{k}\right) \tag{3}$$

Further, we could construct the cost function during the disruption period (i, j) as equation (4), where the second item is the procurement cost of regular capacity, the third is procurement cost of expedite capacity, and the last item is injured capacity recovery cost.

$$\varphi_{i,j} = c_0 z_{i,j} + c_1 U \delta(j - \tau) + c_2 v_{i,j} + k[x_{i,j-1} + z_{i,j} + v_{i,j} + U \delta(j - \tau) - y]^+ + P(\Delta) \tag{4}$$

$\delta(\cdot)$ is dirichlet function, where $\delta(0) = 1$ when $j = \tau$, otherwise $\delta(\cdot) = 0$ when $j \neq \tau$. $x_{i,j}$ in equation (4) is the state transition function, which takes the form as equation (5).

$$x_{i,j} = \min\{x_{i,j-1} + z_{i,j} + v_{i,j} + U \delta(j - \tau) - y, 0\} \tag{5}$$

The expectation disruption cost of equation (5) takes the form as equation (6), where $\theta = x_{i,j-1} + z_{i,j} + v_{i,j} + U \delta(j - \tau)$.

$$\Phi_{i,j} = c_0 z_{i,j} + c_1 U \delta(j - \tau) + c_2 v_{i,j} + k \int_{y=\theta}^{\infty} (y - \theta) f(y) dy + P(\Delta) \tag{6}$$

It can be easily proved that $\Phi_{i,j}$ is convex in $z_{i,j}$, so we omit the proof. And we could get the optimal planned capacity $z_{i,j}^*$ in period (i, j) by solving $\frac{\partial \Phi_{i,j}}{\partial z_{i,j}} = 0$, and the expedite capacity can be obtained through equation (1), that is $v_{i,j}^* = \alpha(Z_{normal} - z_{i,j}^*)$, which is obvious an "order up to" decision.

However, due to the net inventory variable $x_{i,j}^*$, we could not expect the optimal decision $[z_{i,j}^*, v_{i,j}^*]$ in single period to be also optimal in a multi-period decision. Here, we define a cost function $J_{i,j}(x, z, v)$, which is a total expected cost by aggregating expected discounted from $(i, j)^{th}$ period to the (N, m) period. And $J_{i,j}(x, z, v)$ takes the form of equation (7).

$$J_{i,j}(x, z, v) = \Phi_{i,j} + \rho \min\{J_{i,j+1}(x, z, v)\} \tag{7}$$

ρ is discounted cost which is set to 1 in the following equations. Further, let $T_{i,j+1}(x, z, v) = \min\{J_{i,j+1}(x, z, v)\}$, then we could obtain $J_{i,j}(x, z, v)$ as equation (8):

$$J_{i,j}(x, z, v) = c_0 + c_1 U \delta(j - \tau) + c_2 v_{i,j} + \sum_1^i P(\Delta) + k \int_{y=\theta}^{\infty} (y - \theta) f(y) dy + \int_{y=0}^{\infty} T_{i,j}(\theta - y) f(y) dy \tag{8}$$

Proposition 1: $T_{i,j}(x, z, v)$ is convex in $x_{i,j}$, $z_{i,j}$ and $v_{i,j}$.

Proof: Backward induction is adopted in proving proposition 1. Firstly, expectation cost of last period $J_{N,m} = \Phi_{N,m}$, and it can be easily proved

that $\frac{\partial^2 J_{N,m}}{\partial x_{N,m}^2} \geq 0$. Due to $T_{N,m} = \min\{J_{N,m}\}$, it can be

easily proved that $\frac{\partial^2 T_{N,m}}{\partial x_{N,m}^2} \geq 0$. Secondly, we assume

$\frac{\partial^2 T_{i,j+1}}{\partial x_{i,j+1}^2} \geq 0$, as long as $\frac{\partial^2 T_{i,j}}{\partial x_{i,j}^2}$ can be proved to be non-

negative, then we could obtain the proposition 1. According to equation (8), we could get

$\frac{\partial^2 J_{i,j}}{\partial x_{i,j}^2} = kf(\theta) + \int_{y=0}^{\infty} \frac{\partial^2 T_{i,j+1}(\theta - y)}{\partial x_{i,j+1}^2} f(y) dy \geq 0$, and then

$\frac{\partial^2 T_{i,j}}{\partial x_{i,j}^2} \geq 0$. According equation (5) and duality principle,

$T_{i,j}(x, z, v)$ is also convex in z and v .

Proposition 2: Discounted cost $J_{i,j}(x, z, v)$ is convex in $z_{i,j}$ and $v_{i,j}$. There is a unique optimal decision sequence $\{z_{i,j}^*, v_{i,j}^*\}$ that minimizes $J_{i,j}(x, z, v)$, and takes the form as:

$$z_{i,j}^* = \max\{0, \min(z_{i,j}^*, d_i + \Delta\delta(j - m))\} \tag{9}$$

$$v_{i,j}^* = \max\{0, \min(V, Z_{normal} - z_{i,j}^* - x_{i,j} - U\delta(j - \tau))\} \tag{10}$$

Proof: The second order derivative of $J_{i,j}(x, z, v)$ with x has been verified non-negative, that means

$\frac{\partial^2 J_{i,j}}{\partial x_{i,j}^2} \geq 0$, and according to the duality principle,

$\frac{\partial^2 J_{i,j}}{\partial z_{i,j}^2} \geq 0$ is also non-negative. That means the optimal

$z_{i,j}^*$ is the solution of

$$\frac{\partial J_{i,j}}{\partial z_{i,j}} = c_0 - k + \int_{y=0}^{\infty} T'(\theta - y) f(y) dy = 0$$

Considering that $z_{i,j}$ is non-negative, the optimal

$z_{i,j}^* = \max\{0, \min(z_{i,j}^*, d_i + \Delta\delta(j - m))\}$. And the optimal

$v_{i,j}^*$ could be obtained by "order up to" policy by

equation (1), so that

$$v_{i,j}^* = \max\{0, \min(V, Z_{normal} - z_{i,j}^* - x_{i,j} - U\delta(j - \tau))\}$$

5 Numerical simulation

Despite the optimal decision sequence $\{z_{i,j}^*, v_{i,j}^*\}$ being proved, no analytical solution could be foreseen in section 4. So in this section, numerical simulation is presented by genetic algorithm, in order to give much more managerial insights by sensitivity analysis to some key parameters. Basic parameters are set as below: electricity demand y in each period during disruption is normal distributed, of which the mean and variance are $\mu = 1$ and $\sigma^2 = 0.1$ respectively. $d_0 = 0.8$, and $d_0 < \mu$. $P(r) = 10^4 r^2$ which means recovery cost is rather expensive and presents characteristic of "diseconomy of scale". $k = 100$, $c_0 = 2$, $c_1 = 3$, $c_2 = 4$, $Z_{normal} = 1.65$. Sensitivity analyses on parameters of α , τ , N , m , U and V are assumed to be carried.

5.1 SENSITIVITY ANALYSES ON α

Z_{normal} is the optimal planned capacity of each period, the decision variable $z_{i,j}^*$ should be as close as possible to Z_{normal} , so that the optimal $z_{i,j}^* = d_i + \Delta\delta(j - m)$. Then $\{v_{i,j}^*\}$ can be calculated by equation (10), and $\{v_{i,j}^*\}$ is given in FIGURE 2, where $N = 5$, $m = 5$, $U = 0.2$ and $V = 0.1$. It is quite interesting that the optimal expedite capacity ordering sequence $\{v_{i,j}^*\}$ converges into a rectangular wave with its amplitude equals to V before the 12th period when $\alpha \geq 1.0$, while $\{v_{i,j}^*\}$ presents to be a bell-shape curve with its peak decreases when α varies from 0.9 to 0.1. It means that a much more stable and easier decision sequence of $\{v_{i,j}^*\}$ could be obtained when manager takes much more aggressive risk attitude, while the lowest disruption cost is promised when he maximizes the procurement of expedite capacity at the first several disruption periods.

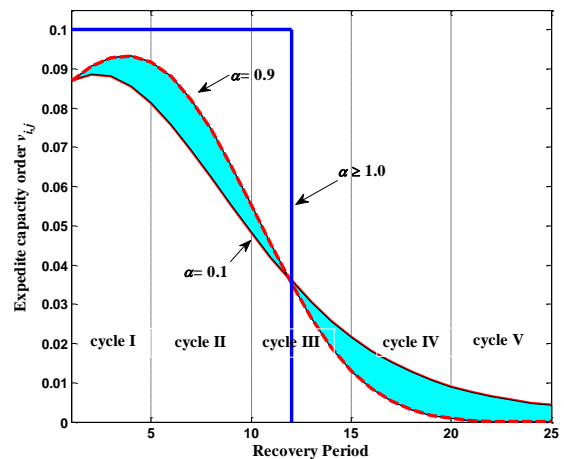


FIGURE 2 Expedite capacity order sequence $\{v_{i,j}^*\}$ when α varies

5.2 SENSITIVITY ANALYSES ON τ , U AND V

Following the parameters setting of section 5.1, the total costs are given by TABLE 1 to show the trends of cost when the lead time of regular capacity τ varies. In generally, total cost would be cut down when τ decreases. However, costs presented in TABLE 1 announce that only 0.13% of cost reduction is contributed when τ decrease from 5 period to 1period. It means that lead time of regular capacity has merely no impact on total cost.

TABLE 1 Total cost when $[\alpha, \tau]$ varies

	$\tau=2$	$\tau=3$	$\tau=4$	$\tau=5$
$\alpha=0.1$	34829	34841	34852	34874
$\alpha=0.5$	34813	34832	34836	34848
$\alpha=1.0$	34808	34823	34839	34855
$\alpha=1.5$	34808	34823	34839	34855
$\alpha=2.0$	34808	34823	34839	34855

Furthermore, we would like to investigate whether total costs will be changed when the upper bound of U and V vary. TABLE 2 gives the numerical results when U varies from 0.0 to 0.4 and V varies from 0.0 to 0.3. It is interesting that cost decreases when either U or V increases. However, the increment of U or V has a marginal decreasing effect in cutting down the cost. And even more, costs presented in TABLE 2 announce that only 2.07% of cost reduction in average is contributed when U or V increases. Hereunto, we could draw the conclusion that manager could not have to get more regular and expedite capacity to the best of his ability during the capacity recovery process, because it contributes very small in cutting down the cost. This conclusion also gives the possible implication that manager might not procure any regular or expedite capacity due to the small reduction in cost.

TABLE 2 Total cost when $[U, V]$ vary when $\tau = 2$

	$V=0.0$	$V=0.1$	$V=0.2$	$V=0.3$
$U=0.0$	35396	34837	34806	34800
$U=0.1$	35249	34835	34773	34763
$U=0.2$	35135	34818	34742	34728
$U=0.3$	35053	34813	34714	34694
$U=0.4$	35001	34809	34691	34661

5.3 SENSITIVITY ANALYSES ON N AND m

Further, we would like to investigate the influence of $[N, m]$ on the total cost with $U = 0.4$ and $V = 0.3$. TABLE 3 gives the simulation results. It is interesting that shortening the recovery cycles (N) has nearly no contribution in cutting down the total cost, while cost is greatly cut down by shortening the lead time (m) that recovered capacity is putting into reuse again. Some ways could help to shorten m , such as improving the labour skill of maintenance department; requiring suppliers

provide spare parts as soon as possible, seeking the maintenance support from partners, and so on.

TABLE 3 Total cost when $[N, m]$ vary when $\tau = 2$

	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$
$N=3$	20813	27758	34610	41671	48641
$N=4$	20838	27793	34656	41729	48712
$N=5$	20867	27833	34661	41795	48795
$N=6$	20899	27878	34769	41873	48892
$N=7$	20935	27928	34834	41958	49001

6 Conclusions and future researches

In this paper, a dual capacity sourcing model for an injured power system is presented by investigating the disruption management of SGZEPC in 2008, and correspondingly mathematical proofs are given to verify the existence of optimal decision, and numerical analysis is presented towards several key parameters by genetic algorithm. Three managerial insights are obtained through our research. Firstly, we find that an stable optimal decision sequence in expedite capacity ordering sequence could be obtained when manager takes aggressive risk attitude in recovery process, and the more aggressive the manager is, the more easier for him to make decisions. Secondly, disruption cost could be cut down by shortening the lead time of regular capacity and maximizing the procurement of regular and expedite capacity. However, the reduction in disruption is rather small, which implies that manager could procure none capacity when the procurement coordination cost is relatively high. Thirdly, disruption cost could be greatly cut down when shortening the lead time of recovered capacity being put into reoperation, which implies skilled maintenance labour, fast provision of spare parts and technique support in capacity recovery from partners could do great contribution in cutting down the disruption cost.

Future works will be carried on two respects. First, since the possibility of manager in procuring non capacity, which implies lower service level of power system is during disruption periods. Supervision penalty as well as its optimal boundary should be further verified to prompt manager providing more electricity. Secondly, whether the conclusions mentioned above are still proper when electricity demand is correlated between neighbouring periods should be further researched.

Acknowledgments

Author wishes to thank the support of National Natural Science Foundation of China under grant No.71302033, Specialized Research Fund for the Doctoral Program of Higher Education under grant No.2011332612002, Zhejiang Provincial Natural Science Foundation of China under grant No.LQ12G01004, and Zhejiang Provincial Social Science Foundation of China under grant No.13ZJQN048YB.

References

- [1] Tang C S 2006 Review: Perspectives in Supply Chain Risk Management *International Journal of Production Economics* **103** 451-488
- [2] Swierczek A 2013 The impact of supply chain integration on the "snowball effect" in the transmission of disruptions: an empirical evaluation of the model *International Journal of Production Economics* <http://dx.doi.org/10.1016/j.ijpe.2013.08.010> 20 August 2013
- [3] Hendricks K B, Singhal V R 2005 An empirical analysis of the effect of supply chain disruptions on long-run stock price performance and equity risk of the firm *Production and Operations Management* **14**(1) 35-52
- [4] Oke A, Gopalakrishnan M 2009 Managing disruptions in supply chains: a case study of a retail supply chain *International Journal of Production Economics* **118** 168-174
- [5] Chopra S, Sohi M S 2004 Managing risk to avoid supply chain breakdown *MIT Sloan Management Review* **46**(1) 53-62
- [6] Kleindorfer P R, Saad G H 2005 Managing disruption risks in supply chains *Production and Operation Management* **14**(1) 53-68
- [7] Serel D A 2008 Inventory and pricing decisions in a single-period problem involving risky supply *International Journal of Production Economic* **116**(1) 115-128
- [8] Sarkar A, Mohapatra P K J 2009 Determining the optimal size of supply base with the consideration of risks of supply disruptions *International Journal of Production Economics* **119**(1) 122-135
- [9] Meena P L, Sarmah S P, Sarkar A 2011 Sourcing decisions under risk of catastrophic event disruptions *Transportation Research Part E* **47** 1058-1074
- [10] Yu G, Qi X T 2004 *Disruption Management: Framework, Models and Applications* World Scientific Publishing Co., River Edge, NJ. Chapter 3
- [11] Poullikkas A 2013 Optimization analysis for pumped energy storage systems in small isolated power systems *Journal of Power Technologies* **93**(2) 78-89
- [12] Huang C, Yu G, Wang S, et al 2006 Disruption management for supply chain coordination with exponential demand function *Acta Mathematica Scientia* **26**(4) 655-669
- [13] Aoki M, Rothwell G 2013 A comparative institutional analysis of Fukushima nuclear disaster: Lessons and policy implications *Energy Policy* **53** 240-247
- [14] Berge P D, Gerstenfeld A, Zeng A Z 2004 How many suppliers are best? A Decision-analysis approach *Omega* **32** 9-15
- [15] Ruiz-Torres A J, Mahmoodi F 2007 The optimal number of suppliers considering the costs of individual supplier failures *International Journal of Production Economics* **35**(1) 104-115
- [16] Kopytov E, Greenglaz L, Muravjov A, et al 2007 *Computer Modelling & New Technologies* **11**(1) 21-30
- [17] Meena P L, Sarmah S P, Sarkar A 2011 Sourcing decisions under risk of catastrophic event disruptions *Transportation Research Part E* **47** 1058-1074
- [18] Tomlin B T 2006 On the value of mitigation and contingency strategies for managing supply chain disruption risks *Management Science* **52**(5) 639-657
- [19] Yang J, Qi X, Yu G 2005 Disruption management in production planning *Naval Research Logistics* **52**(5) 420-442
- [20] Qi X, Bard J F, Yu G 2006 Disruption management for machine scheduling: the case of SPT schedules *International Journal of Production Economics* **103**(1) 166-184
- [21] Bao X 2010 Capacity management model of production operation system after the disruption *Journal of Industrial Engineering/Engineering Management* **24**(1) 45-50 (In Chinese)

Authors



Bao Xing, born in October, 1981, Hangzhou, Zhejiang, P.R.China

Current position, grades: Associate Professor of Business administration Department, Zhejiang Provincial Committee Party School of CPC, Hangzhou, Zhejiang, China.

University studies: received his doctoral degree in Business administration from Shanghai Jiaotong university in China.

Scientific interest: His research interest fields include disruption management, risk management, supply chain management, financial risk management.

Publications: more than 25 papers published in various journals.

Experience: He has teaching experience of 6 years, has completed 5 scientific research projects which are supported by MOE (Ministry of Education in China) Project of Humanities and Social Sciences, National Natural Science Foundation of China, Specialized Research Fund for the Doctoral Program of Higher Education under grant No.2011332612002, Zhejiang Provincial Natural Science Foundation of China, and Zhejiang Provincial Social Science Foundation of China.

Analysis on the dynamic effects of the aggregate supply, aggregate demand and macroeconomic policies of china based on SVAR model

Jianxin Bi^{1,2*}, Lianghai Lei²

¹ Faculty of Management, University of Shanghai for Science and Technology, 516 Jun Gong Road, Shanghai, China

² Faculty of Computer and Information, Zhejiang Wanli University, No.8, South Qian Hu Road Ningbo, Zhejiang, China

Received 1 March 2014, www.tsi.lv

Abstract

After studying the financial crisis using the AS-AD model and the SVAR model, the paper analyses the dynamic effects of the aggregate supply, aggregate demand and macroeconomic policies of China. Then, combining the real macroeconomic environment of China at present, the paper discusses the Keynesian AD-AS model, gives the constraint conditions of SVAR model according to the economic meaning of China, makes an empirical study based on five selected variables including supply, demand, fiscal expenditure, monetary expenditure and interest rate and their monthly data correspondingly, and lastly analyses the results of empirical study to make recommendations on current macroeconomic policy adjustment of China.

Keywords: AD-AS model, SVAR model, impulse response, variance analysis

1 Introduction

The financial crisis caused by American sub-prime mortgage crisis in 2007 had different shocks to economies of different countries in different degrees from different channels. During the financial crisis, many scholars, at home or abroad, studied the effects on Chinese economic fluctuations focusing on the effect issue of fiscal policy & monetary policy and exogenous shocks. Main document researches are as follows: Wenfu W [1] studied the dynamic shock effects of government expenditure on total output, social investment and resident consumption using a structural vector auto regression (SVAR) model and came to a conclusion that government expenditure had a positive effect on total output and a complementary effect on social investment and residents' consumption. Qiang D [2] built a dynamic stochastic general equilibrium model based on financial accelerator theory to study external shock effects during economic fluctuations with the quarterly data of China. Numerical simulation results consisted with real economic operation and thus had a strong explanatory power to the fluctuation of output and investment in real economy. He draw such a conclusion that shocks of exchange rate and foreign demand had an obvious dual function: the appreciation of RMB had a greater adverse impact, while the recession in foreign demand was not as terrible as imagined. Chinese economy would rise from the bottom by expanding domestic demand. Within a M-F model framework, Rong L [3] introduced the conclusion of the classic paper of Bernanke & Blinder

[20] to build an economic structure model of China and established a SVAR model under the model to estimate the dynamic adjustment characteristics of economic system. He calculated the overall pulling effect of current round of expansionary monetary policy on economy by impulse response function and then analysed and compared the effects of different channels. In conclusion, most document researches only explored economic fluctuation by focusing on one factor, such as monetary policy, fiscal policy or external shock. Research literatures of domestic and foreign scholars on the dynamic effects of aggregate supply, aggregate demand and macroeconomic policy in Chinese economic fluctuations are few.

American subprime mortgage crisis has inflicted a severe impact on Chinese economy. The fluctuations and instability of Chinese economy require the government to take active measures to adjust economy correspondingly. Under current market economy conditions, the government makes macroeconomic regulations with *The General Theory of Employment, Interest, and Money* of Keynes as the source of theory. Keynes proposed macro-control policies of fiscal policy and monetary policy for market imbalance and economic fluctuation. For either policy, the goal is demand management, which is fixing market imbalance by regulating aggregate supply – aggregate demand. Any regulation & control theory is based on corresponding theory and model describing macroeconomic fluctuations. As to Keynes' theory, its main idea describing macroeconomic fluctuations can be reflected in the famous AS-AD model.

* Corresponding author e-mail: greygirl0511@126.com

After studying the financial crisis using an AS-AD model and a SVAR model, the paper analyses the dynamic effects of the aggregate supply, aggregate demand and macroeconomic policy in Chinese economy. AS-AD-model-based SVAR theoretical research literatures are as follows: Min G [5] estimated a SVAR model containing output and price level based on the AS-AD model to reveal the changes of AS-AD forces driving Chinese economic fluctuations during 1996-2005; Jian G [6] built an AS-AD model with Chinese characteristics and estimated & tested the model using quarterly data. Researches show that the macroeconomic model with Keynes characteristics can explain the economic fluctuations in China to some degree. The paper first uses Jian G's [6] AS-AD model with Chinese characteristics and Wenfu W's theoretical analytical framework for reference and then improves the model on this basis. Jian G estimated the equation using Cochrane-Orcutt program and LSM; while the paper makes estimations using a SVAR model and the SVAR model is a structured VAR model. The VAR model is modelling based on the statistical property of data and building the model using each endogenous variable as the function of lagged value of all the endogenous variables in the system, thus extending the single-variable autoregressive model to a vector autoregressive model composed by multivariate time series variables. VAR model does not give the exact form of current correlation among variables, that is, there is no current value of endogenous variables in the right hand of model. Current correlations hide in the related structure of error term and cannot be explained, so the structured VAR model placing the current value of endogenous variable into the right had of model, i.e. the SVAR model, can reduce estimated parameters by imposing constraint conditions to the parameter space. Next, the empirical study of Min G [5] only analysed the shock effects of aggregate supply and aggregate demand in Chinese economic fluctuations. When Wenfu W tested the matching or applicability of Keynesian AD-AS model in China, he only selected four variables of supply, demand, fiscal expenditure and monetary expenditure and used quarterly data as the endogenous variables of SVAR model without regarding the variable of interest rate, in which case the persuasiveness of conclusion was weakened to some degree in the empirical analysis. Facts proved that since 2007 the People's Bank of China (PBOC) has exercised low control and regulation to interest rate. Currently, PBOC controls loan interest rate ceiling & floor, and other market interest rates are almost open, such as the monetary market, constant return interest rate and inter-bank rate, which are all open. In 2007, Shanghai Interbank Offered Rate (SHIBOR) was established as a reference rate. On this basis, when selecting variables, besides four variables of supply, demand, fiscal expenditure and monetary expenditure, the paper also adds interest rate into the endogenous variables of SVAR model; in addition, the literature empirical researches mentioned above all used quarterly

data as the object of study, but the SVAR model in the paper only considers short-term restraints and the characteristics of macroeconomic policy and selected monthly data for empirical analysis. At last, basing on a SVAR theoretical research, the paper discusses Keynesian AD-AS model combining the real macroeconomic environment in China at present, proposes the constraint conditions of SVAR model according to the economic meaning of China, recognizes SVAR structural equation form from VAR simple equation form, and makes an empirical analysis using the monthly data from September of 2008 to February of 2012, thus is of innovation.

The structure of the paper is as follows: the second section discusses Keynesian AD-AS model; the third section analyses the shock effects of supply, demand and macroeconomic policy on Chinese economic fluctuations using the data from September of 2008 to February of 2012 and a modified AS-AD model through the impulse response function and variance decomposition means; the fourth section gives a conclusion and recommendations.

2 Discussions on Keynesian AD-AS Model

Within the basic framework of real business cycle theory and using AS-AD model as a major analysis tool, the section discusses traditional AS-AD model combining the real economic background in China currently to make the modified AD-AS model match the real economic background of China currently.

2.1 CURVE FUNCTION OF AGGREGATE SUPPLY

Generally, the aggregate supply curve is expressed as:

$$p_t = AS(p_{t-1}^e, y_{t-1}), \quad (1)$$

where y_t represents GDP, p_t represents the general price level in period t and subscript t is time, p_{t-1}^e represents the expected value of price level of period t during period $t-1$, y_{t-1} represents the total output in period $t-1$ and first derivative $AS'_p > 0$, $AS'_y > 0$ means price level is decided by output and expected price level and aggregate supply is their increasing function respectively. For the expectation p_{t-1}^e , many literature researches try to make an explanation using a rational expectation hypothesis, but a lot of empirical researches don't support the hypothesis [15-17]. According to Keynesian adaptive expectation principles, the expectation on price can be identified as $p_{t-1}^e = f(p_{t-1})$, and then Chinese aggregate supply function is:

$$p_t = AS(p_{t-1}, y_{t-1}). \quad (2)$$

Make a Taylor series expansion to equation (2), and after a linear approximation it can be expressed as:

$$p_t = \beta_0 + \beta_p p_{t-1} + \beta_y y_{t-1} + \theta_t, \tag{3}$$

where β_0 is a constant, β_p and β_y are coefficients and θ_t represents the changes of aggregate supply or price. $\{\theta_t\}_{t=0}^\infty$ is a random walk process, i.e., $\theta_t = \theta_{t-1} + u_t^s$, where u_t^s is supply shock or price shock, the cost factor shock affecting price, such as the price rise of raw materials of bulk commodity. $\{u_t^s\}_{t=1}^\infty$ is an independent identically distributed random process where u_t^s follows normal distribution $N(0, \sigma_s^2)$. σ_s^2 represents the variance of supply shock.

2.2 CURVE FUNCTION OF AGGREGATE DEMAND

Aggregate demand curve comes from the famous IS-LM model. Traditionally, the model is expressed as:

$$y_t = IS(r_t, d_t), \tag{4}$$

$$\frac{M_t}{p_t} = LM(y_t, r_t). \tag{5}$$

In the equation, r_t is the interest rate of period t , d_t is fiscal policy variable, M_t is nominal money supply (in which case $m_t = M_t / p_t$ is real money supply), and $IS_r' < 0$, $IS_d' > 0$, $LM_y' > 0$, $LM_r' < 0$. IS equation decides the equation of output, which is the aggregate social demand decided by consumption demand, investment demand and government expenditure demand, may be adding net export demand. LM equation reflects market equilibrium which means real money supply equals to real currency demand. The aggregate demand curve got by solving Equation (3) and (4) and is as follows:

$$y_t = AD(m_t, d_t), \tag{6}$$

where $m_t = M_t / p_t$ and $AD_m' > 0$, $AD_d' > 0$. In the equation, assume money supply is exogenous, while interest rate is endogenously determined by the balance of currency market. However, in western countries, as the evolution and development of modern financial and banking system, the increase in money supply is more and more decided by the money demand generated from the demand for production and circulation within economic society. That is, PBOC can affect interest rate by monetary policy tool and then affect the cost of money supply, but it cannot control money supply at will.

Therefore, whether from the perspective of the basic model of modern money supply or from the perspective of the development of money supply theory, money supply is significantly endogenous. In the empirical analysis of Jing W [8], the data from quarter 2 of 2001 to quarter 2 of 2011 indicated that Chinese money supply showed more and more endogenous property and the existing compulsory settlement & sales system in China must be reformed to promote interest rate liberalization reform actively.

In addition to the endogeneity of money supply and interest rate, the lagging effect of monetary policy and fiscal policy on economy is also taken into consideration, and together with prior revenue's effect on current demand, so the aggregate demand curve equation of China can be expressed as:

$$y_t = AD(y_{t-1}, m_{t-1}, d_{t-1}), \tag{7}$$

where $AD_m' > 0$, $AD_d' > 0$. Make a Taylor series expansion to equation (6) and the expression after linear approximation is:

$$y_t = \lambda_0 + \lambda_m m_{t-1} + \lambda_d d_{t-1} + \lambda_y y_{t-1} + u_t^y. \tag{8}$$

In the expression, λ_0 is a constant and constant coefficient $\lambda_m > 0$, $\lambda_d > 0$, $\lambda_y < 0$. u_t^y is the aggregate demand shock in period t . $\{u_t^y\}_{t=1}^\infty$ is an independent identically distributed random process. u_t^y follows normal distribution $N(0, \sigma_d^2)$. σ_d^2 represents the variance of demand shock.

2.3 POLICY RESPONSE FUNCTION

We first discuss the response function of fiscal policy. The change rule of d_t can be expressed as:

$$d_t = \delta_0 + \delta_y y_t + \delta_p p_t + \delta_d d_{t-1} + u_t^d. \tag{9}$$

In the equation, δ_0 is a constant and constant coefficient $\delta_y < 0$, $\delta_p < 0$, $\delta_d < 0$. u_t^d is the government expenditure shock in period t following normal distribution $N(0, \sigma_d^2)$. $\{u_t^d\}_{t=1}^\infty$ is an independent identically distributed random process. σ_d^2 represents the variance of government expenditure shock. The rule indicates that when output or price rises, fiscal policy will be restrained, and vice versa.

The next is the response function of monetary policy. When money supply is endogenous, money aggregate is endogenous in economic operation and decided jointly by the subjects of social economic activities, rather than an

exogenous variable decided and controlled by governmental monetary authority (PBOC). Similar to equation (8), money supply rule has the following form:

$$m_t = \theta_0 + \theta_y y_t + \theta_p p_t + \theta_d d_t + \theta_m m_{t-1} + u_t^m \tag{10}$$

In the equation, θ_0 is a constant and constant coefficients $\theta_y < 0$, $\theta_p < 0$, $\theta_d > 0$, $\theta_m < 0$. u_t^m is the money supply shock of PBOC in period t . $\{u_t^m\}_{t=1}^\infty$ is an independent identically distributed random process. u_t^m follows normal distribution $N(0, \sigma_m^2)$. σ_m^2 represents the variance of money supply shock.

Last, we discuss the rule of interest rate. Interest rate is generally endogenous, but when money supply is endogenous, it is difficult for PBOC to realize money supply rule using monetary policy tools (such as open-market operation, discount rate, reserve ratio, etc.). In fact, the central banks of many developed countries currently have given up money supply rule and use the indicator of interest rate as the intermediate target more frequently, and the so called Taylor rule (interest rate rule) is generated there from.

$$r_t - r_{t-1} = \phi_0 + \phi_1 (r_{t-1} - r^*) + \phi_y (y_t - y^*) + \phi_p (p_t - p^*) \tag{11}$$

In the equation, $\phi_1 < 0$, $\phi_y < 0$, $\phi_p < 0$. r^* , y^* and p^* can be interpreted as the targets of interest rate, output and price. The rule indicates that when a target deviates, interest rate will be adjusted [10, 18].

According to equation (11), it is simplified into the following form:

$$r_t = \omega_0 + \omega_r r_{t-1} + \omega_y y_t + \omega_p p_t + u_t^r \tag{12}$$

where the interest rate impact is u_t^r in period t , $\{u_t^r\}_{t=1}^\infty$ is an independent identically distributed random process, u_t^r follows normal distribution $N(0, \sigma_r^2)$ and σ_r^2 represents the variance of interest rate shock.

Because China is in a critical period of interest rate rule transition currently, according to the characteristics of transitional economy, either money supply or interest rate can be considered as a policy variable. Therefore, Chinese money policy can be presented as the implementation of both rules that is using money supply rule (10) and interest rate rule (12) simultaneously [11].

To sum up, for Chinese economic fluctuation currently, formula (3) represents aggregate supply curve; formula (8) represents aggregate demand curve; formulas (9), (10) and (12) represents the policy response rules of fiscal policy, monetary policy and interest rate rule.

According to the methods of Blanchard [19], Blanchard & Quah [21] and Gali [14], the following difference equation can be got from Equation (3), (8), (9), (10) and (11):

$$\Delta p_t = \beta_0 + \beta_p p_{t-1} + \beta_y y_{t-1} + \theta_t^p \tag{13}$$

$$\Delta y_t = \lambda_m \Delta m_{t-1} + \lambda_d \Delta d_{t-1} + \lambda_y \Delta y_{t-1} + u_t^y - u_{t-1}^y \tag{14}$$

$$\Delta d_t = \delta_y \Delta y_{t-1} + \delta_p \Delta p_{t-1} + \delta_d \Delta d_{t-1} + u_t^d - u_{t-1}^d \tag{15}$$

$$\Delta m_t = \theta_y \Delta y_{t-1} + \theta_p \Delta p_{t-1} + \theta_d \Delta d_{t-1} + \theta_m \Delta m_{t-1} + u_t^d - u_{t-1}^d \tag{16}$$

$$\Delta r_t = w_r \Delta r_{t-1} + w_y \Delta y_{t-1} + w_p \Delta p_{t-1} + u_t^r - u_{t-1}^r \tag{17}$$

The following equation form can be obtained from formulas (13) ~ (17).

$$\begin{bmatrix} \Delta p_t \\ \Delta y_t \\ \Delta d_t \\ \Delta m_t \\ \Delta r_t \end{bmatrix} = \begin{bmatrix} d_{11}(L) & d_{12}(L) & d_{13}(L) & d_{14}(L) & d_{15}(L) \\ d_{21}(L) & d_{22}(L) & d_{23}(L) & d_{24}(L) & d_{25}(L) \\ d_{31}(L) & d_{32}(L) & d_{33}(L) & d_{34}(L) & d_{35}(L) \\ d_{41}(L) & d_{42}(L) & d_{43}(L) & d_{44}(L) & d_{45}(L) \\ d_{51}(L) & d_{52}(L) & d_{53}(L) & d_{54}(L) & d_{55}(L) \end{bmatrix} \begin{bmatrix} u_t^s \\ u_t^y \\ u_t^d \\ u_t^m \\ u_t^r \end{bmatrix} \tag{18}$$

where d_{ij} ($i, j=1, \dots, 5$) represents a lagging polynomial. The economic meaning of equation (18) is the theoretical basis of the empirical analysis in the paper and the empirical analysis is made by considering the five variables of Δp_t , Δy_t , Δd_t , Δm_t , Δr_t as the endogenous variables of SVAR model.

3 AD-AS Model Empirical Analysis in Chinese Macroeconomic Fluctuations

3.1 SAMPLE SELECTION AND VARIABLE DEFINITION

The financial crisis can be divided into three stages of development: from sub-prime mortgage crisis to financial crisis, from financial crisis to the spread of substantial economy, and the financial market volatility again since 2010. In the second stage when the financial crisis happened, Chinese government adjusted macroeconomic policy immediately from the tight monetary policy and prudent fiscal policy of 2007 to the macroeconomic economy of keeping stable economic development & controlling price hikes in August of 2008, and to the proactive fiscal policy & moderately loose monetary policy in November of 2008. Because the object of study in the paper is analysing the dynamic effects of aggregate supply, aggregate demand and economic policy when China introduced economic policies successively after the crisis, the second stage's starting point, i.e. September of

2008, is the starting point of the data selected in the paper.

The data selected are monthly data of the following variables from September of 2008 to February of 2012 from the source of RESSET/DB. In the empirical part, five endogenous variables of GDP growth rate, inflation rate, monetary growth rate, fiscal expenditure and interest rate are selected. Because there is no monthly data on GDP growth rate, the paper uses monthly industrial growth rate as the substitution for monthly GDP growth rate y_t . Inflation rate p_t is calculated using consumer price indexes (CPIs) with the formula Inflation Rate= (Current CPI-Base CPI)/Base CPI. Monetary growth rate m_t is a narrow monetary growth rate. Fiscal expenditure variable d_t can be calculated with the formula (Fiscal Expenditure-Fiscal Revenue)/Aggregate Fiscal Expenditure. Interest rate r_t is Shanghai interbank offered monthly rate.

3.2 QUANTITATIVE TEST OF AS-AD AND MACROECONOMIC POLICY OF CHINESE ECONOMY AFTER FINANCIAL CRISIS

3.2.1 Test of VAR Model Lag Structure

We first inspect the chart of AR root. As shown in Figure 1, the all the inverse roots of estimated VAR model is in the unit circle, i.e. <1, indicating a stable state and meeting the conditions of pulse response analysis and variance analysis.

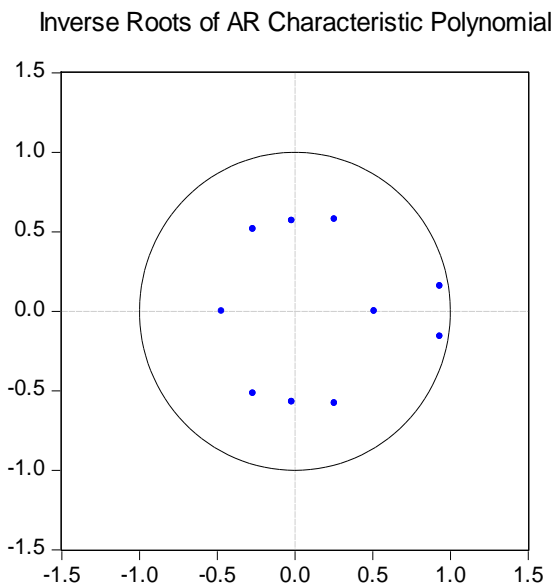


FIGURE 1 VAR Model Stationary Inspection

Next, the paper uses ADF (Augmented Dickey-Fuller) method to make a stationary test to each sequence and corresponding first difference sequence. Test results are shown in table 1. In the ADF test of original sequence, d_t 's simple-root statistic -6.982528 is less than -2.935001, ADF critical value with a significance level of 5%, so variable d_t passes the test. Similarly, m_t also passes the

test. Other variables fail the ADF test. Then we make the first-difference test. As in table 2 all the variables pass the ADF test, indicating the first difference is stationary.

TABLE 1 ADF Test of Sequence and Its First Difference Sequence

Variable	ADF Test				Result
	1% level	5% level	t-Statistic	Prob*	
d_t	-3.600987	-2.935001	-6.982528	0.0000	Stationary
m_t	-3.600987	-2.935001	-6.849287	0.0000	Stationary
p_t	-3.600987	-2.935001	-1.831882	0.3603	Non-stationary
r_t	-3.600987	-2.935001	-0.452598	0.8901	Non-stationary
y_t	-3.600987	-2.935001	-3.513982	0.0125	Non-stationary
Dd_t	-3.632900	-2.948404	-5.653387	0.0000	Stationary
Dm_t	-3.610453	-2.938987	-10.84507	0.0000	Stationary
Dp_t	-3.605593	-2.936942	-5.556035	0.0000	Stationary
Dr_t	-3.610453	-2.938987	-4.001337	0.0000	Stationary
Dy_t	-3.615588	-2.941145	-5.763493	0.0000	Stationary

Table 1 shows that each sequence is first-difference stationary, which means each sequence is a first-difference single integration sequence. Therefore, there is a need to make a co-integration test to the variables contained in the model. We use Johansen co-integration test to check if the model has a co-integration relationship. The result shows that 5 co-integration equations exist under the significance levels of 1%~5% and the endogenous variables in the model have a co-integration relationship.

At last, we determine lag order. VAR equation is tested and the result shows that the optimal lag order k is selected according to AIC criteria, SC criteria and LR criteria. Results show that the optimal orders from two information criteria are different. AIC rules suggest selecting lag period 4; SC information criteria suggest selecting period 1; LR test suggests period 4. Because AIC and LR rules both select lag period 4, we finally selected the form of lag period 4, i.e. p=4.

3.2.2 SVAR Model Recognition Constraint Condition

On the theoretical basis of AD-AS model analysis, we select five endogenous variables of the measurement models in the paper. The five variables are industrial growth rate y_t , which is a substitution for monthly GDP growth rate, inflation rate p_t , monetary growth rate m_t , fiscal policy variable d_t and interest rate r_t .

The short-term constraints of structured VAR are generally based on AB-type SVAR model. It means A & B are 5x5 invertible matrices and satisfy:

$$A\varepsilon_t = B\mu_t, t=1,2,\dots,T. \tag{19}$$

The recognition conditions are parameters that can estimate the structural formula model only when imposing $k(k-1)/2=10$ constraint conditions. The constraint conditions are either in the same period (short-term) or long-term [12].

About the constraint conditions of SVAR model in the paper, we give the following five assumptions according to Equation (13) - (17) and their economic meaning. 1) Current GDP growth rate, current fiscal expenditure, current monetary growth rate and current interest rate have no effect on current inflation rate. 2) The changes of current inflation rate affect the changes of current GDP growth rate. Fiscal policy and monetary policy have a time lag, so the changes of government expenditure and monetary growth rate do not affect GDP growth rate in current period. In the short run, interest rate does not affect GDP growth rate. 3) Current fiscal expenditure is only affected by current inflation rate, current GDP growth rate and self shock but not monetary policy and interest rate. 4) Monetary policy is affected by current inflation rate, current GDP growth rate, current government expenditure changes and the self shock of current monetary growth rate but not current interest rate. 5) Current interest rate is affected by the shocks of current inflation rate, current GDP growth rate and interest rate changes.

3.2.3 Estimate Structure Factors

Basing on the recognition model of equation (19), the paper selects SVAR model's endogenous vector $\mu_t = [\Delta p_t, \Delta y_t, \Delta d_t, \Delta m_t, \Delta r_t]$, where the first difference sequence of monthly p_t, y_t, d_t, m_t and r_t represents the time series corresponding to $\Delta p_t, \Delta y_t, \Delta d_t, \Delta m_t$ and Δr_t . The disturbing term ε_t of simple equation is a five-dimensional vector $\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}, \varepsilon_{4t}, \varepsilon_{5t}]$, of which each component is the linear combination of $u_t^s, u_t^y, u_t^d, u_t^m, u_t^r$. Therefore, ε_t represents a composite shock, and the matrices A & B got are as follows:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.13 & 1 & 0 & 0 & 0 \\ 0.41 & -1.56 & 1 & 0 & 0 \\ -0.05 & 0.23 & -0.07 & 1 & 0 \\ 0.07 & 0.20 & -0.03 & 0.17 & 1 \end{bmatrix}, \tag{20}$$

$$B = \begin{bmatrix} 1.45 & 0 & 0 & 0 & 0 \\ 0 & 0.26 & 0 & 0 & 0 \\ 0 & 0 & 2.56 & 0 & 0 \\ 0 & 0 & 0 & 0.48 & 0 \\ 0 & 0 & 0 & 0 & 0.27 \end{bmatrix}. \tag{21}$$

Then, according to the recognition principles of SVAR model, we get a structural equation and the results of impulse response function (IRF) and variance decomposition necessary for analysis.

3.2.4 Response Analysis of Impulse Function

IFR describes the effects on endogenous variables' current & future values when adding a one-off shock to a disturbing term.

1) Dynamic Effects of Supply Shock

Figure 2 shows the dynamic response process of output, price level, money supply, interest rate and government expenditure when the aggregate supply has a 1% positive shock. First, we explore price level's response to supply shock. In current period of shock, price level rose immediately by a big margin and then declined continuously and reached the lowest in 5 months. After that, it rose slowly and kept on a stable level in the 26th month or so. The final accumulative shock effect is increased. It means the positive shock of supply may cause the rise in price level [26]. Next, we explore output's response to supply shock. Output rose continuously in current period and then declined slowly after the 3rd month. It finally reached the lowest in the 5th month or so. After that, it rose slightly and reached a stable state in the 20th month. The final accumulative shock effect is increased. Then, we explore the dynamic response process of fiscal policy and monetary policy to supply shock. After the supply shock, the accumulative shock effect on government expenditure approximated to zero, an invariant state. Money supply rose slowly in early period and began to decline in the 5th month. Then it rose and declined again and reached a stable value after the 12th month or so. The final accumulative shock effect is increased. Finally, we explore supply shock effect on interest rate. Interest rate was unchanged in current period and began to decline slowly in the 6th month and reached the lowest in the 15th month. Then, it reached a stable state. The final accumulative shock effect of supply on interest rate is decreased.

The following conclusion can be drawn from the analysis above: 1) after the financial crisis, aggregate supply has a positive accumulative shock effect on price level, output and money supply; 2) after the financial crisis, the accumulative shock effect of supply on government expenditure approximates to 0, i.e. nearly no supply shock; supply has a negative accumulative shock effect on interest rate. After the financial crisis, America took a quantitative easing policy causing the rise of bulk commodity in international market. China has a high dependency on bulk commodity, such as oil, in world market. The rise in foreign commodity price may cause the increase in Chinese commodity export, thus increasing the export demand of foreign trade in China. On the other hand, the rise in foreign commodity price may reduce Chinese residents' consumption of import commodity and increase the consumption of domestic commodity. In this case, it may result in an increase in the aggregate demand of the whole society. Because of the rise in foreign commodity price, Chinese foreign trade may face large surplus, which may increase the

foreign exchange reserves of China greatly. In a pegged exchange rate system, large foreign exchange reserves may cause a great increase in domestic money supply and then cause the decline of interest rate and the increase in investment and finally result in a demand-pull inflation.

2) *Dynamic Effects of Demand Shock*

When aggregate demand has a positive shock of 1%, output, price level, money supply, interest rate and government expenditure have a dynamic response process to it. First, we explore the demand shock effect on price level. Current price dropped to a value immediately. Then the accumulative response effect of price level rose slowly and continuously. It declined slowly in the 5th month or so and began to rise in the 8th

month with small fluctuations. It fluctuated during the 8th -24th month and stabilized after the 24th month. It indicates the positive demand shock promotes the rise in price level. Next, we explore output's dynamic response to demand shock. Output rose to a level in current period suffering the demand shock, and then declined continuously. It dropped to the lowest accumulative decline value in the 1st month and then rose slowly with fluctuations. It reached the highest in the 5th month, dropped slightly and then rose again. It kept stable and constant in the 27th month. It means demand shock may help output increase. Then, we explore the responses of government expenditure and money supply to demand shock. Government expenditure rose continuously in current period and reached the top in the 1st month or so.

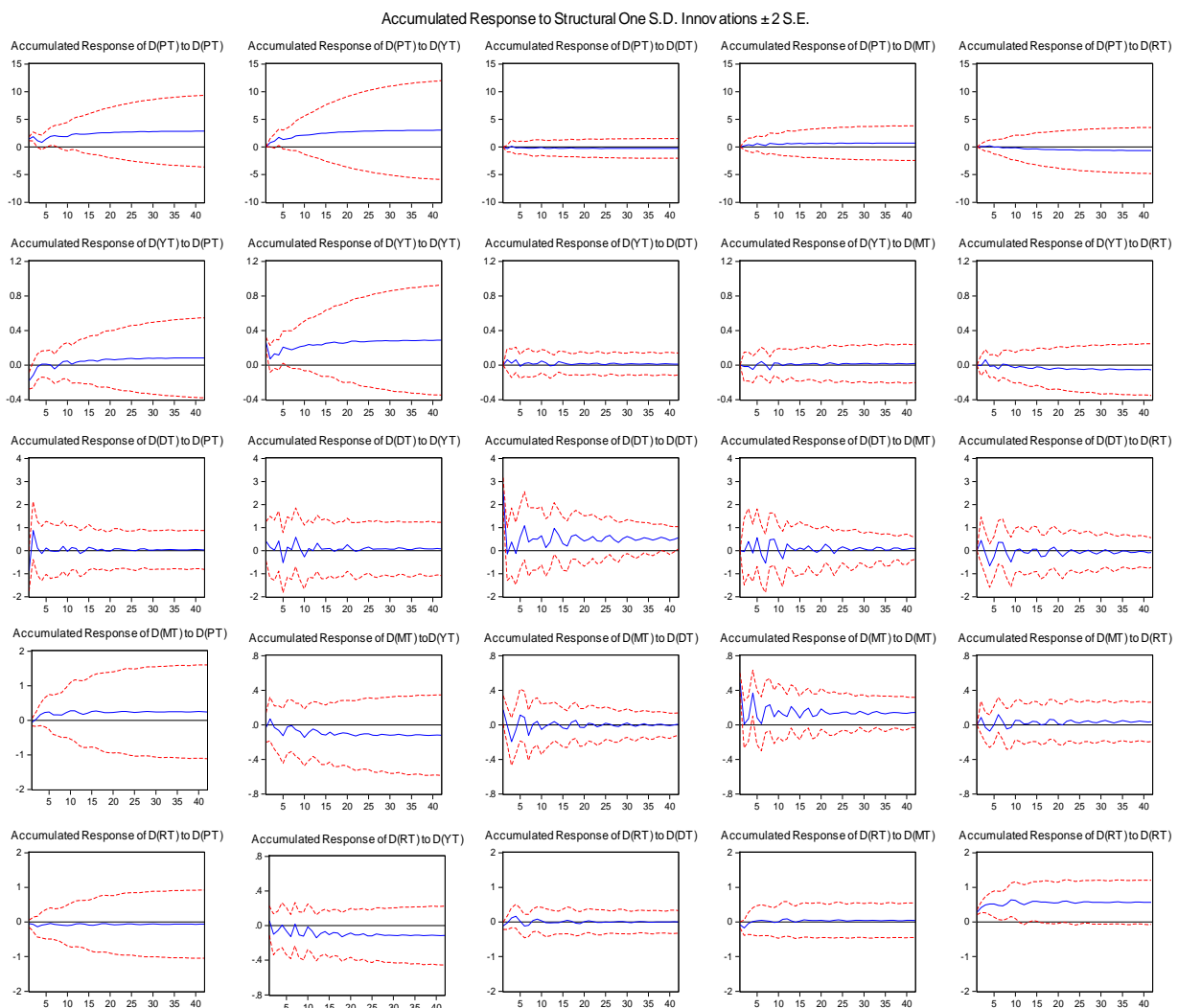


FIGURE 2 Dynamic Effects of Accumulative Shocks among Variables
(Full line represents impulse response function; dotted line represents ± 2 -times standard deviation)

Then, it declined continuously and rose slightly. It fluctuated during the 5th -27th months and reached a stable state in the 27th month. Its final accumulative shock effect approximates to 0. Money supply did not change at first in current period. In the 3rd month, it declined slightly and

then rose. It reached a stable state approximating to 0 in the 23rd month. After the demand shock, interest rate rose slightly in early days, reached the maximum in the 3rd month, and then fell. It suffered small fluctuations during the 3rd-23rd months and reached a stable state after the

23rd month. The accumulative shock effect was negative, indicating demand shock may cause interest rate decline. The following conclusion can be drawn from the analysis above: 1) after the financial crisis, demand has positive accumulative shock effects on price level and output; 2) after financial crisis, demand has a negative accumulative shock effect on interest rate and nearly no shock on government expenditure and money supply. The first conclusion consists with AD-AS model's economic meaning that demand increase results in the rise in output and price level, and it consists with the conclusion of Blanchard [19]. When adopting deflation policy after demand shock, demand has slight negative shocks to government expenditure and money supply in a short term but nearly no shock in the long run. The conclusion consists with the economic connotation of Keynesian macro demand management policy. Meanwhile, from table 2 we can see that the contribution of demand shock to output fluctuations is about 50.83%, significantly bigger than the contributions of other four shocks. It consists with Keynesian view that output fluctuations mainly come from the demand shock. Aggregate demand contains investment demand and consumption demand. The fundamental approach to increase effective demand is increasing residents' income steadily, so residents' income can be increased by approaches such as reducing tax or cutting interest rate. Therefore, positive demand shock may cause interest rate decline.

3) *Dynamic Effects of Government Expenditure Shock*

This section discusses the dynamic response process of output, price level, money supply, interest rate and government expenditure when government expenditure has a positive shock of 1%. First, we explore level price's response to government expenditure shock. When price level suffered a government expenditure shock of 1%, it rose sharply in current period and reached the highest in the 1.5th month or so. Then it dropped continuously and reached the lowest in the 4th month. After that, it rose again and suffered fluctuations during the 5th-27th months. In the 27th month, it reached a stable state. The final accumulative shock effect is positive. It indicates that government expenditure may cause the rise in price level. Next, we explore output's dynamic response to government expenditure. After a government expenditure shock of 1%, output rose slightly in current period. Then it dropped slowly continuously, rose slowly and dropped slowly again. It reached the lowest in the 5th month and then rose with a small range continuously. It suffered small fluctuations during the 6th-38th months and reached a stable value in the 38th month. The final accumulative shock effect is positive. Then, we explore the dynamic responses of government expenditure and money supply under government expenditure shock. Government expenditure is government's expenditure for random events, such as the external financial crisis of economy and natural hazards like flood or earthquake. In the case of 1% government expenditure shock, government

expenditure rose greatly to a level in current period and then the accumulative effect dropped sharply to the lowest in the 1st month. After that, it rose continuously to the highest in the 3rd month and then dropped. It suffered small fluctuations during the 4th - 42nd months and gradually stabilized in the 42nd month. As to the dynamic response of money supply, after the government expenditure shock of 1%, money supply rose slightly, dropped and rose again. It suffered small fluctuations during the 2nd - 42nd months and reaches a stable state after 42 months. It indicates that when suffering a government expenditure shock, government expenditure and money supply will both increase to eliminate the adverse effects in economy. Last, we explore the dynamic response of interest rate to government expenditure shock. When suffering the government shock of 1%, interest rate rose slightly at the beginning. It began to drop slowly in the 2nd month and reached the lowest in the 4th month. After that, it rose again. It had small fluctuations during the 7th-25th months and reached a stable state after the 25th month. The final accumulative shock effect approximates to 0.

The following conclusion can be drawn from the analysis above: 1) after the financial crisis, government expenditure has positive accumulative shock effects on price level, output, government expenditure and money supply; 2) after the financial crisis, government expenditure has a negative accumulative shock effect on interest rate. To cope with Chinese economic downturn under the impact of international financial crisis, Chinese government began to implement positive fiscal policy in November of 2008, mainly appearing as Chinese government planned to use ¥ 4 trillion government expenditure in advance for railway construction, rural infrastructure construction, social security expenditure, increasing export rebate rate, promotion of fuel oil tax, etc. Government expenditure increased in a short time, i.e. in the first 25 months (09/2008-09/2010), resulting in the increase in aggregate output. Output increase caused the increase in money demand and money market rate (i.e. the rise in capital cost). In this case, marginal cost of production rose and then caused rising prices. As the rise in output and price level, benchmark interest rate began to rise, and money supply began to rise. After 25 months, the effects on aggregate supply and interest rate disappeared and the effects on other economic variables tended to be stable. Keynesianism believes that government expenditure may increase aggregate demand and promote aggregate demand curve to move, thus causing the rise in price level; government expenditure also promotes the increase in output in a short time, which consisting with AD-AS model's economic meaning that government expenditure helps output increase.

4) *Dynamic Effects of Money Supply Shock*

The section focuses on the dynamic response process of output, price level, money supply, interest rate and

government expenditure after a 1% positive shock of money supply. First, we discuss price level's dynamic response to money supply shock. After the 1% money supply shock, the accumulative dynamic effect of price level rose slowly at first and then dropped slowly and rose slowly again. It reached a stable state in the 17th month or so. The final accumulative shock effect is positive. Then we explore the dynamic response of output to money supply shock. Output dropped with a small range immediately in current period and then rose slightly. It rose to the highest in the 2nd month and then dropped slowly. In the 42nd month, it reached a stable value. The final accumulative shock effect is negative. The next is the dynamic responses of government expenditure and money supply to money supply shock. When suffering a money supply shock, current money supply rose greatly immediately and then its accumulative effect dropped sharply and reached the minimum in the 1.5th month or so. After that, it rose continuously to the maximum in the 5th month and then dropped again. It suffered small fluctuations in the 7th - 42nd months. The final accumulative shock effect is positive. As to the dynamic effects of money supply on government expenditure, government expenditure rose in current period when suffering the shock, and then it dropped continuously and reached the lowest point in the 3rd month or so. After that, it rose continuously and began to drop continuously in the 5th month. It had small fluctuations during the 7th - 24th months and finally stabilized in the 24th month. The final accumulative shock effect is negative but small. Last, we explore the dynamic response of interest rate to money supply shock. When suffering the money supply shock, interest rate rose slightly and then began to drop in the 2nd month and rose again with small fluctuations. It reached a stable state in the 40th month. The final accumulative shock effect is positive but small.

The following conclusion can be drawn from the analysis: 1) after the financial crisis, money supply has positive accumulative shock effects on price level, money supply and interest rate; 2) after the financial crisis, money supply has negative accumulative shock effects on output and government expenditure, but the effects are so small that can be neglected. After the financial crisis, to curb Chinese economic decline, PBOC lowered deposit reserve ratio four times from the 17.5% in June of 2008 to the 14.5% in the end of the year, adjusted deposit interest rate in August of 2008 three times from 4.14% falling to 2.25%, and regulated loan interest rate four times from 7.47% falling to 5.31%. According to model estimation, we can see monetary policy has more limited effects on interest rate and government expenditure. That is, in the first eight months, money supply can affect output, but later the effect is very small, and positive monetary policy has a negative shock to output. It indicates monetary policy has limited effects on economy, consisting with AD-AS model's economic meaning. The reason is unobvious effect of monetary policy, and some scholars share the same view [3]. The

effects of money supply on price level are much bigger than on output. The effects of money supply on various variables can be analysed through variance decomposition. As shown in the table, after the 20th month, money supply is driving contribution to price level accounts for 8.79% but to output is only 4.46%. The former nearly doubles the latter. It indicates the changes of price almost completely depend on money in the long term. It accords with the principles of economics and can be confirmed by practice. After China adopted easy monetary policies in 2008, economic decline was curbed in a short period, but in the next year of 2011, a serious inflation occurred and money supply shock seemed more act on price level rather than output, causing more difficulties for realizing the multiple objects of monetary policy simultaneously.

5) *Dynamic Effects of Interest Rate Shock*

This section discusses the dynamic response process of output, price level, money supply, interest rate and government expenditure when interest rate has a 1% positive shock. First, let us see the dynamic response of price level to interest rate shock. After the money supply shock of 1%, price level dropped slightly in current period and had small fluctuations during the 0-18th months. It reached a stable state in the 18th month or so. The final accumulative shock effect is negative. It indicates the positive shock of interest rate may cause the small decline of price level. Second, the dynamic response of output to interest rate shock. Output rose slightly and immediately in current period and then dropped slowly. It had small fluctuations during the 2nd - 26th months or so and reached a stable value after the 26th month. The final accumulative shock effect is negative. It indicates the positive shock of interest rate may cause output decline, consisting with the views of some scholars [3]. Then, the dynamic responses of government expenditure and money supply to interest rate shock. When suffering the interest rate shock, government expenditure declined slightly in current period and then rose slightly. It had small fluctuations during the 3rd - 8th months. In the 8th month, its accumulative effect stayed in a stable value. Its final accumulative shock effect approximates to zero, indicating government expenditure has nearly no change under the positive shock of interest rate. As to the dynamic effects of interest rate on money supply, money supply reduced in current period when suffering the shock and then rose. It reached a stable state in the 3rd month. The final accumulative shock effect approximates to zero, indicating that money supply has nearly no change under the positive shock of interest rate. Last, the dynamic responses of market interest rate to interest rate shock. When suffering the interest rate shock, market interest rate rose greatly in current period and then rose slowly. It reached a stable state in the 25th month or so, indicating the interest rate shock may cause a sharp increase in the actual interest rate in market.

The following conclusion can be drawn from the analysis above: 1) after the financial crisis, interest rate has negative accumulative shock effects on price level and output; 3) after the financial crisis, the positive shock of interest rate only affects government expenditure for eight months and the effect approximates to 0 later; it only affects money supply for three months and the effect approximates to 0 later; it has a positive accumulative shock effect on interest rate. The reason is the positive interest rate shock (i.e., deflation monetary policy) has negative effects on price level and economy. In case of rate hike, financial institutions may increase the excess reserves depositing in PBOC, and banks' funds for lending decrease, which equalling to reducing money supply. Therefore, there are influences in a short time (i.e., in three months), but no influence in a long time. Interest rate generally does not affect government expenditure, consisting with relevant economic theory.

3.2.5 Analysis on the Relative Importance of Shock

Variance decomposition further evaluates the importance of different structural shocks by analysing the contribution of each structural shock to endogenous variable changes (generally measured by variance). Therefore, variance decomposition gives the information on relative importance of each random disturbance affecting the variables in VAR model. Table 2 shows the contribution rate of shocks among variables.

TABLE 2 Variance Decomposition of the Prediction Errors among Variables

Period	Price Level	Output Shock	Government Expenditure	Money Supply	Interest Rate
Aggregate Supply					
Month 1	100.00	0.00	0.00	0.00	0.00
Month 10	58.75	25.11	6.83	7.56	1.75
Month 20	58.38	24.38	7.15	7.89	2.20
Month 30	58.12	24.35	7.22	8.00	2.26
Month 42	58.15	24.34	7.23	8.01	2.27
Aggregate Demand					
Month 1	33.30	66.70	0.00	0.00	0.00
Month 10	25.02	53.36	6.64	7.92	7.07
Month 20	24.92	51.27	8.13	8.33	7.35
Month 30	24.76	50.90	8.31	8.60	7.42
Month 42	24.72	50.83	8.35	8.65	7.46
Fiscal Expenditure					
Month 1	10.39	2.17	87.44	0.00	0.00
Month 10	16.78	8.82	56.88	10.38	7.15
Month 20	15.78	9.29	55.64	11.75	7.54
Month 30	15.54	9.32	55.13	12.22	7.79
Month 42	15.47	9.32	55.03	12.29	7.89
Money Supply					
Month 1	1.18	0.41	11.28	87.15	0.00
Month 10	8.08	3.94	19.42	62.98	5.58
Month 20	8.78	4.40	19.35	61.61	5.85
Month 30	8.81	4.45	19.58	61.08	6.08
Month 42	8.79	4.46	19.64	60.96	6.16
Interest Rate					
Month 1	2.345	3.60	12.82	7.00	74.22
Month 10	3.64	15.10	11.20	11.08	58.99
Month 20	4.12	15.72	9.82	13.16	57.18
Month 30	4.14	15.64	9.64	13.36	57.22
Month 42	4.13	15.59	9.72	13.38	57.18

1) *After the financial crisis, expanding domestic demand became an important means to keep the economic growth of China.*

Table 2 shows, that aggregate supply has an accumulative contribution rate of 58.14% to the deep shock of price level and an accumulative contribution rate of 24.34% to output shock; aggregate demand has an accumulative contribution rate of 24.71% to price level shock and an accumulative contribution level of 50.82 to output shock. Overall, the contribution rate of aggregate demand shock to output is more than twice of aggregate supply's contribution to output. Therefore, after the financial crisis, to improve economic growth, the government must consider expanding domestic demand as a main instrument.

2) *After the financial crisis, proactive fiscal policy curbed the economic decline of China effectively, while easy money policy backfired as to its adjustment to macro economy.*

In table 2, fiscal expenditure has an accumulative contribution rate of 15.47% to price level shock and an accumulative contribution rate of 9.32% to output shock. It indicates that proactive fiscal policy curbed the economic decline of China effectively, but the effect was limited. Positive money supply has an accumulative rate of 8.79% to price level shock and an accumulative rate of 4.45% to output shock. It means in the long term expansionary money policy's shock to price level is twice of that to output (see Figure 2), and positive money supply has a negative shock to output. Therefore, taking expansionary money policy to curb economic decline is just the opposite of the wish and may bring a huge negative effect, i.e., the inflation. Therefore, it is more difficult to achieve economic goals through monetary policy.

3) *During the financial crisis, the reduction of interest rate had limited regulation effects on economy.*

As shown in table 2, interest rate has an accumulative contribution rate of 4.13% to price level shock, an accumulative contribution rate of 15.59% to output shock, an accumulative contribution rate of 9.72% to fiscal expenditure shock, an accumulative contribution rate of 13.38% to money supply and an accumulative contribution rate of 57.18% to interest rate itself. It indicates that interest rate has limited effects on price level and certain effects on output. After the financial crisis, China took policies of interest & interest rate reduction, which curbed economic decline to some extent with limited effects. The conclusion consists with Marxist interest rate's macro policy effect, and PBOC's interest rate policy has a certain asymmetry in its effect. In a period of economic prosperity, because there is a great investment demand and a relative shortage of money, PBOC's interest rate policy has a relative effective regulation result; in a period of economic recession, because there is an insufficient effective demand on

consumption, desired investment declines naturally and enterprise' demand on funds falls with it. Because of investment risk, banks' loan to enterprises also declines and money in the market is sufficient relatively. In this case, it is difficult for PBOC's interest rate policy to work [13]. Therefore, PBOC's interest rate policy should be used in asymmetric operations corresponding to different stages of economic cycle. In an inflation period with overheated economy, the regulation & control should be dominated by interest rate policy instrument; while in a deflation period with economic recession, interest rate policy should be used in conjunction with many policy instruments, such as money supply, credit scale and fiscal policy.

4 Conclusion and recommendations

On the theoretical analysis basis of AS-AD model and using the SVAR model, the paper reveals the dynamic effects of aggregate supply, aggregate demand and macroeconomic policies of China from September of 2008 to February of 2012 under the background of financial crisis.

The paper first analyses the shocks among the variables of aggregate supply, aggregate demand, government expenditure, money supply and interest rate after the financial crisis according to impulse function response, and then draws the following conclusions for analysis: 1) the aggregate supply has positive accumulative shock effects on price level and output; 2) the demand has positive accumulative shock effects on price level and output; 3) the government expenditure has positive accumulative shock effects on output, government expenditure and money supply, and it has negative accumulative shock effects on interest rate, but every shock effect is limited; 4) the money supply has positive accumulative shock effects on price level and negative accumulative shock effects on output; 5) the interest rate has negative accumulative shock effects on output.

References

- [1] Wenfu W 2010 Empirical analysis on the dynamic effects of chinese Government expenditure *Soft Science* **123**(3) 28-31 (in Chinese)
- [2] Qiang G, Li Z 2010 Research on external shocks in Chinese macroeconomic fluctuation – general equilibrium numerical simulation analysis based on financial accelerator theory *Economic Review* **110**(5) 112-38 (in Chinese)
- [3] Rong L, Xi W 2010 Effect of monetary policy addressing international financial crisis study of international finance **123**(7) 19-29 (in Chinese)
- [4] Ying L, Peng N 2011 Dynamic effect analysis on Chinese fiscal policy in post-crisis era – numerical simulation based on dsge model *The Economist* **115**(4) 53-62 (in Chinese)
- [5] Min G, Wenbo L 2007 Analysis on the sock effect of aggregate supply and aggregate demand in chinese economic fluctuation *Economic Research Journal* **86**(11) 32-44 (in Chinese)
- [6] Jian G, Nian Y 2007 Chinese aggregate supply-aggregate demand model: analytical framework of fiscal and money policies *The Journal of Quantitative & Technical Economics* **106**(5) 3-11 (in Chinese)
- [7] Wenfu W, Juan M 2009 Dynamic effect analysis on aggregate demand, aggregate supply and macroeconomic policy – matching of ad-as model and Chinese data *The Journal of Quantitative & Technical Economics* **126**(11) 39-50 (in Chinese)
- [8] Jing W, Xianhua W 2012 Empirical analysis on endogenous problems of chinese money supply *Contemporary Finance & Economics* **78**(6) 61-8 (in Chinese)
- [9] Jun L, Dan Z 2003 Con-integration test of Taylor rule in china *Economic Research Journal* **112**(8) 76-93 (in Chinese)
- [10] Yisheng L, Yilin L 2010 Research on Taylor rule and its adaptability to china *Contemporary Economic Research* **98**(10) 48-53 (in Chinese)
- [11] Gang G 2003 Primary exploration on the goal and management of Chinese money policy *South China Financial Research* **112**(5) 98-104 (in Chinese)
- [12] Tiemei G M 2009 Econometric Analysis Methods and Modelling (2nd edition) Press: Tsinghua University Press (in Chinese)

Next, basing on the variance decomposition analysis, the paper makes recommendations on Chinese macroeconomic policy adjustment after the financial crisis. Main conclusions are as follows: 1) expanding domestic demand should be used as an important approach to keep Chinese economic growth currently; 2) the proactive fiscal policies can effectively curb Chinese economic decline, while easy monetary policy may have opposite effects on macro-economic adjustment; 3) interest rate reduction has limited effects on economic regulation & control. Therefore, in the post-crisis era, the proactive fiscal policy adopted by Chinese government plays an important role in keeping sustained and stable economic growth. However, the effects of monetary policies are still controversial among scholars. According to the empirical study in the paper, currently Chinese government should focus on the stability and continuity of fiscal policies and use monetary policies and interest rate adjustment policies flexibly as supplementary tools.

The follow-up study of the empirical analysis in the paper should be improved in three aspects. First, the paper only selects five variables of price level, output, government expenditure, money supply and interest rate as the endogenous variables of SVAR model, but does not take other macroeconomic variables, such as credit, revenue and labour employment into consideration. Second, the paper uses monthly data for the empirical analysis, but Chinese macroeconomic data are only quarterly and some variables even do not have monthly data, which affected the empirical analysis results to some extent, so the selection of monthly-data variables should be further discussed in later period. Third, the modification of AD-AS model under Chinese economic environment should be further studied.

Acknowledgements

In the process of writing paper, thanks for the help and guidance of my colleagues and supervisor.

- [13] Qicai H 2011 Theoretical and Empirical Study on Chinese Interest Rate Change and Operation Rule *Western Economics* 286(6) 127-8 (in Chinese)
- [14] Gali J 1992 How well does the is-lm model fit post-war U.S. data *The Quarterly Journal of Economics* 107(2) 709-38
- [15] Gali J, Gertler M 1999 Inflation dynamics, a structural econometric analysis *Journal of Monetary Economics* 89(4) 195-222
- [16] McCallum B T M 1989 *Monetary Economics: Theory and Policy* Press: New York
- [17] Beyer A, Farmer R, Henry J, Marcellino M 2005 Factor Analysis in a New Keynesian Model Working Paper Series: No. 510 European Central Bank
- [18] Taylor 1993 Discretion versus policy in practice *Carnegie-Rochester Conference Series on Public Policy* 39(7) 195-214
- [19] Blanchard O J 1989 A traditional interpretation of macroeconomic fluctuations *American Economic Review* 165(5) 1146-64
- [20] Bernanke B S, Blinder A S 1992 The federal funds rate and the channels of monetary transmission *American Economic Review* 156(82) 901-21
- [21] Blanchard O J, Quah D 1989 The dynamic effects of aggregate demand and supply disturbances *American Economic Review* 109(4) 654-73
- [22] Bernanke B M 1986 Alternative explanations of the money-income correlation *Carnegie-Rochester Conference Series on Public Policy*: North-Holland
- [23] Shapiro M, Watson M 1988 Sources of business cycle fluctuations *NBER Macroeconomics Annual* 108(6) 111-48
- [24] Sims C A 1980 Macroeconomics and reality *Econometrica* 206(48) 1-48
- [25] Sims C A 1986 Are forecasting models usable for policy analysis *Quarterly Review Federal Reserve Bank of Minneapolis* 211(10) 2-6
- [26] Jianxin B, Lianghai L 2014 *Computer Modelling & New Technologies* 18(2) 115-23

Authors	
	<p>Jianxin Bi, born in May, 1974, Jiutai City, Jilin Province, P.R. China</p> <p>Current position, grades: the lecturer of Faculty of Computer and Information, Zhejiang Wanli University. University studies: doctoral candidate of Faculty of Management, the University of Shanghai for Science and Technology. Scientific interest: Financial and fiscal Management, data mining. Publications: Presided over 3 scientific research projects the completion of provincial; more than 10 papers published in various journals. Experience: Graduated from Fu zhou University in 2005, has completed 3 scientific research projects; more than 10 papers published in various journals.</p>
	<p>Lianghai Lei, born in February, 1962, Qianjiang City, Hubei Province, P.R. China</p> <p>Current position, grades: professor, doctoral supervisor and Doctor of Economics of Faculty of Management, the University of Shanghai for Science and Technology. University studies: Graduated from Shanghai University of Finance and Economics in 1996, received a doctor's degree in Economics. Scientific interest: Financial and fiscal Management, Company financial, The theory and methods of enterprise internal control. Publications: Presided over the completion of national and provincial, than 20 scientific research projects; more than 140 papers published in various journals. Experience: Graduated from Shanghai University of Finance and Economics in 1996, received a doctor's degree in economics, was approved as a tutor of doctoral students in 2002; has completed including the national and provincial and ministerial level, than 20 scientific research projects; more than 140 papers published in various journals published monographs, 7 teaching materials.</p>

Study on Xiangyang's population and aging trend prediction based on discrete population development equation model

Qiong Gu^{1,2}, Wu Zheng¹, Xianming Wang^{3*}

¹ College of Mathematics and Computer Science, Hubei University of Arts and Science, No.296 Long Zhong Road Xiangyang, Hubei 441053, China

² Centre for the Study of Logic and Intelligence, Southwest University, No.2 Tiansheng Road, BeiBei District, Chongqing 400715, China

³ Oujian College, Wenzhou University, Chashan University Town, Wenzhou 325035, China

Received 1 March 2014, www.tsi.lv

Abstract

Population problem is an important factor that influences economy and social development of China. This paper takes the statistic data of 6th census in 2010 in Xiangyang as the accordance to establish a discrete model of population development equation, to analyse the population aging trend in the future in Xiangyang from a short period, and further to predict the long-term population development trend and aging population change condition in Xiangyang in the case of different total fertility rate to provide reference accordance for the government to make relevant social and economic decisions.

Keywords: population aging, population development equation, discrete model, total fertility rate

1 Introduction

Population problem is one of the important factors that influence social and economic development, and correct forecast on the development trend of future population has important guiding significance to the overall planning of the country and local government. Many foreign and domestic experts and scholars have been paying more and more attention to population prediction and control issue and propose many different prediction methods and models, such as population index growth model [1] established by British demographer T.Malthus, logistic model raised by Netherlands mathematician Verhulst [2], and also time-space regression model [3], grey model[4], ARMA model [5], neural network model [6] and gene expression model [7] and so on. In order to predict and analyse the population development of China, Scholars like Song Jian and Yu Jingyuan and others [8] raised population development equation model, and Jiang Yingyuan and others [9] applied this model to predict the population of China.

According to the international convention, the population of 65 years and above people exceeds 7% of the total population or the population of 60 years and above exceeds 10% means that the population aging is coming. From the 6th census data of Xiangyang, we can see that the population of 65 years and above is 467,159 in 2010 in Xiangyang, accounting for 8.49% of the total population, and have exceeded the definition boundary of international population aging proportion as 7%. At present, the population of 40 years to 59 years accounts for 16.8% of the total population, so it can be predicted that the aged people in Xiangyang will become more after

10 to 20 years. It is found from the population structure pyramid of Xiangyang that the population aging speed of Xiangyang will still accelerate and various social problems brought by aging will become more serious. Facing to the inevitable population aging problems, what kind of policy should the government to take so as to be able to reduce the influences from population aging and shorten the continuing time of irrational population structure? What kind of population and family planning policy is rational and can make the population of Xiangyang develop stably and sustainably? This paper takes the 6th census data in Xiangyang as the accordance, uses the discrete model of population development equation to predict the population development in Xiangyang, to analyse the population aging trend in the future in Xiangyang from a short period and further predict the long-term population development trend and aging population change condition in Xiangyang in the case of different total fertility rate.

2 Discrete Model of Population Development Equation

In the condition of not considering social factors like population migration and so on, the factors that influence population structure are birth rate and mortality rate. Suppose that $x_r(t)$ expresses the population of r years (within r years but not within $r+1$ years) in the year of t , the maximum age is m , and $m, r=0,1..m$. In order to study the population of different ages at any time, suppose $b_r(t)$ expresses the fertility rate of childbearing age women at years old in the year of t , namely, the

* Corresponding author e-mail: xmwung@ustc.edu

number of babies born by each r years old female on average in the year of t . Suppose $[r_1, r_2]$ is the reproductive age interval and $k_r(t)$ expresses the ratio of the population of r years old female in the total population in the year of t , then the population of people born in the year of t is:

$$b(t) = \sum_{r=r_1}^{r_2} b_r(t)k_r(t)x_r(t), \tag{1}$$

$d_r(t)$ expresses the mortality rate of the r years old population in the year of t , and let $s_r(t) = 1 - d_r(t)$, then $s_r(t)$ is the survival rate of years old population in the year of r years old population in the year of t . $d_0(t)$ is infant mortality rate, then the infant survival rate is $s_0(t) = 1 - d_0(t)$. We can get:

$$\begin{cases} x_0(t) = s_0(t)b(t) \\ x_{r+1}(t+1) = s_r(t)x_r(t), r = 0, 1, 2, \dots, m-1 \end{cases} \tag{2}$$

Equation (1) and (2) describe the population development process. $s_0(t)$, $s_j(t)$ and $k_j(t)$ can be got by the statistic data calculation of population census. We find that population policy can control the population development by controlling $b_i(t)$. Let us make further analysis on $b_i(t)$. Suppose $\beta(t) = \sum_{r=r_1}^{r_2} b_r(t)$, and $\beta(t)$ is the total fertility rate, namely, the number of babies born by each woman on average in the year of t . If let $b_r(t) = \beta(t)h_r(t)$, thereinto, $\sum_{r=r_1}^{r_2} h_r(t) = 1$, $h_r(t)$ is called women's fertility mode, which is r years old female' fertility weighted factor. The current population policy in China is to realize less birth by changing $\beta(t)$ and get late childbirth by controlling $h_r(t)$, and finally to slow down population growth.

Introduce the vector and matrix, let

$$X(t) = \begin{bmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_m(t) \end{bmatrix},$$

$$A(t) = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ s_1(t) & 0 & \dots & 0 & 0 \\ 0 & s_2(t) & \dots & 0 & 0 \\ \vdots & \vdots & \dots & 0 & 0 \\ 0 & 0 & 0 & s_{m-1}(t) & 0 \end{bmatrix},$$

$$B(t) = \begin{bmatrix} 0 & \dots & 0 & b_{r_1}^*(t) & \dots & b_{r_2}^*(t) & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix}_{m \times m+1}.$$

Thereinto, $b_r^*(t) = s_0(t)s_r(t)k_r(t)h_r(t)$, then the population development equation expressed by Equation (1) and (2) can be described as:

$$X(t+1) = A(t)X(t) + \beta(t)B(t)X(t). \tag{3}$$

$A(t)$, $B(t)$ and $\beta(t)$ can be got by calculating the population census data, and then it is not hard to work out the equation (3).

3 Model Establishment and Solution

As Xiangyang has come into aging society, in order to make clearly the future population aging trend, we use population development equation to predict the population in a short period and analyse the change trend of future population aging speed. At the same time, suppose predicting the long-term population development process of Xiangyang in the case of different total fertility rate, and try to seek for a relatively good total fertility rate to make that the fertility rate can control too fast population growth but also ease the population aging condition.

3.1 SHORT-TERM POPULATION AGING TREND ANALYSIS

As the data information provided by the 6th census in Xiangyang is limited, and there are many factors that influence population development, it is relatively difficult to get the rule that mortality rate function and fertility mode changes as the time changes. The population fertility policy will not change in a relatively short period; in the condition of stable society, we usually suppose that mortality rate $d_r(t)$, fertility mode $h_r(t)$ and gender ratio $k_r(t)$ relatively do not change. The data published from the 6th census in Xiangyang shows the total fertility rate of women $\beta(t) = 1.452$ at present. Under the supposition above, we use the data of the 6th census in 2010 in Xiangyang as the cardinal number to predict the population in the future five years in Xiangyang. The prediction result is shown as Table 1.

TABLE 1 Prediction on the Future Five Years' Population in Xiangyang (Person)

Age	2011	2012	2013	2014	2015
0-4	319562	323160.2	326806.4	329764	331130.2
5-9	250614	264097	279564.7	294580	306350.6
10-14	262343	246321	241095.2	241843.1	242771.8
15-19	438958	410846	378771.8	334236	293393.6
20-24	491697	493552	477170.5	461832.6	457817.9
25-29	395696	412710	437961.1	465298	475771.9
30-34	388379	381586	380444	387224.4	393617.3
35-39	485979	457145	422285.5	391049.8	382742.4
40-44	530538	536472	550352.8	551842.5	515717.6
45-49	489769	543013	536654.6	508658.2	514476.1
50-54	364486	336106	359195.7	397285.4	439060.3
55-59	364572	374832	380700.9	389189.1	384041.7
60-64	247700	263389	280934.4	305755.1	328583.5
65-69	177346	182812	190082.6	201436.9	214760.1
70-74	121284	126869	135469.9	142383.8	150189.2
75-79	89102	95630	102164.1	105792.7	101310.5
80-84	42087	46416	49371.2	52850.42	59597.86
85-89	19450	20568	21133.56	22086.06	23254.98
90-94	4380	5219	6098.094	7085.156	8162.724
95-99	1286	1428	1471.349	1499.016	1556.674
100 years	133	85	137.3737	178.8721	180.093
Total	5485361	5522256	5557866	5591871	5624487

From the prediction result, we can see the population in Xiangyang is still growing. Although the data of the 6th census shows the population fertility rate in Xiangyang is relatively low, the total population still continuously grows because the population cardinal number is very big.

TABLE 2 Proportion of the Population at Each Age Section in Total Population

Age	2011	2012	2013	2014	2015
0-14	0.150949	0.152481	0.154901	0.156504	0.15839
15-64	0.762306	0.75649	0.749726	0.744107	0.737264
65 years and above	0.086745	0.091029	0.095373	0.099389	0.104346

We can see from the calculation result of population proportion at each age section that the proportions of aged people and children still continuously increase, but the proportion of adults continuously decreases. This is mainly because there were three times' child-bearing peak after the establishment of the People's Republic of China, the baby climax caused the population to grow rapidly and these persons born in the child-bearing peak period will come into aged period and thus to cause the growth of aging population; moreover, according to the baby climax occurrence time and statistic data of Xiangyang, we can deduce that the population aging speed in the future in Xiangyang will intensify further.

In the condition of not considering population migration, in short time, it is very difficult for us to ease population aging trend through the fertility policy regulation. The population mortality rate is the key that decides the population aging development trend; by the comparison between the urban population and rural population mortality rate of Xiangyang, as shown in

Figure 1, we can predict the change direction of population mortality rate of Xiangyang.

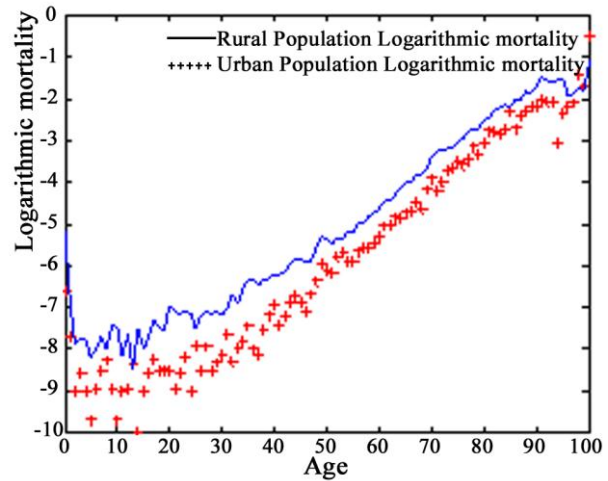


FIGURE 1 Comparison between Urban Population and Rural Population Rate of Xiangyang

From Figure 1, we can see that the mortality rate of urban population is much lower than that of the rural population. Xiangyang is in "Four Xiangyang" construction at present, and the future urbanization degree, people's living environment and living condition all will be changed a lot, these will further decrease the population mortality rate, and it will definitely cause the population aging further to accelerate. We can see from the prediction and analysis above that the population aging of Xiangyang is an inevitable population problem, and the population aging speed is still continuously intensifying.

3.2 LONG-TERM POPULATION DEVELOPMENT TREND ANALYSIS

As the population in China comes into aging society, some experts and scholars take opposite attitudes towards the original only-child policy, and corresponding adjustments on family planning policy are also made by the country and each local government. For example, the couple are both rural residents, and one of the couple is in the family which is only-child family for two generations; one of the couple is a level B, grade two above disabled veteran; the man goes to the woman whose family has only one child to get married and settle; the couple who only has a girl; the couple both are ethnic minorities. The couple are both urban residents, whose first child is disabled and can't grow up as a normal labour; the couple does not have child and requires to bear child after adopting a child legally; the couple both are from only-child families and other conditions; the couple meets the conditions above all can apply for bearing the second child. The report of the 18th CPC National Congress did not refer to stabilizing low child-bearing level but strengthened completing family planning policy step by step to promote population to develop in a long-term balanced mode; it shows that China has not only

controlled the population quantity but also paid attention to the population quality and the sustainable population development.

From the short term, the family planning policy will not cause big problems to the population of Xiangyang temporarily, however, from the long term, it will have certain influences on the population structure. In order to seek for a kind of relatively good population fertility policy, we suppose that the other factors do not change in the dozens of years in the future, and predict the population development condition of Xiangyang in the case of different total fertility rate. With the improvement of people's living quality, the baby cultivation cost in the future will increase gradually; even if in the condition of opening only-child bearing policy, a relative part of people who meet the conditions of bearing the second child do not choose to bear more children, so the total fertility rate at present in China basically is below 2. Following, we will suppose that the total fertility rate $\beta(t)$ is respectively 1.452, 2.0 and 2.5, and analyse and predict the total population and aging population proportion in the future 50 years in Xiangyang in these conditions; the result is shown as Figure 2.

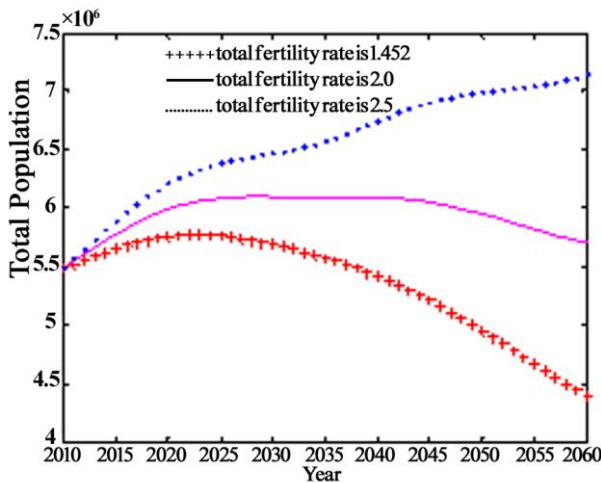


FIGURE 2 Total Population Development Trend in the Future 50 Years in Different Total Fertility Rate Conditions

From Figure 2. we can find that the total population quantity in the future will increase continuously when the total fertility rate is 2.5, and it obviously bring more pressure to the population quantity control. When the total fertility rate maintains the current condition, namely, $\beta(t)=1.452$, the total population quantity gets to peak value in very short time, and soon the total population will further decrease. Under the trend of continuous acceleration of population aging speed of Xiangyang, this total population development trend will aggravate the aging and thus to cause great pressure to the economy and society and further to result in rapid decrease of labour proportion. When $\beta(t)=2.0$, the total population development will be relatively stable in the dozens of years in the future. However, there will be a decrease trend after 2050; it is because a relatively part of families

are only-child families at present, and when the parents of these children from only-child families pass away, it will cause the total population to decrease. Meanwhile, we also can see the change trend of aging population proportion in the future 50 years from the prediction result; the result is shown as Figure 3.

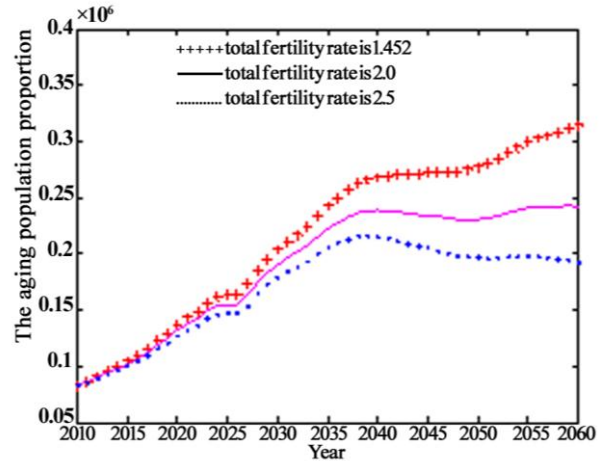


FIGURE 3 Change Trend of Aging Population Proportion in Different Total Fertility Rate Conditions

From Figure 3, we can see when $\beta(t)=1.452$, the aging population proportion will continuously increase and finally exceed 30% in 2060. When $\beta(t)=2.0$, the aging population increase trend is relatively slow, and will be stably at about 24% after 20 years. When the total fertility rate is 2.5, the aging population proportion is relatively slow, and the aging population proportion will maintain within 20% basically in the future 50 years.

From the analysis result, we can find that it is hard to get a win-win situation between population quantity control and population aging release. The relatively high fertility rate will make the population aging keep at a relatively low level, but will cause the rapid population growth; only when the total fertility rate keeps at two can a relatively good result appear.

4 Summary

This paper predicts the population development process in the future 5 years in Xiangyang by establishing the discrete model of population development equation; the total population in the future 5 years in Xiangyang will continuously increase and the aging speed will further accelerate. The population aging is an inevitable population problem, and several generations of us need to endure and face population aging; we can not only solve it through population policy but should release the economic pressure and social problems caused by population aging by making corresponding old-age supporting policy and social security policy, and regulating economic structure and other ways. In order to seek for a relatively rational population fertility policy, we predict the population development in the future 50 years in Xiangyang in different fertility rate conditions.

We find that the existing population policy will further speed up population aging and increase the pressure caused by population aging. When the total fertility rate of every woman is 2.0, the population development in the future in Xiangyang will be relatively stable, and the population-aging trend will get slow and finally keep at the level of 24%.

This paper predicts the population development process in Xiangyang and analyses the population aging trend in the future. It makes us clearly know the population situation in the face of Xiangyang, and it has important significance to achieving sustainable

population development goal and making a rational population policy. In future, we will also analyse the change trend of factors influencing population development, such as population mortality rate, fertility mode and gender ratio and so on, so as to get more an accurate conclusion.

Acknowledge

This work was supported by Educational Commission of Hubei Province of China (D20132601).

References

[1] Thomas R M 2008 *An essay on the principle of population* Shaanxi Normal University Press: Shaanxi (in Chinese)

[2] Hu X S, Fan H L, Song P, Hong W, et al 2008 Application of Urban Population Prediction Based on Modified Logistic Model *Journal of Beihua University (Natural Science)* 9(4) 370-3 (in Chinese)

[3] Liu Q.P 2009 Application of Spatio –Temporal Regression Model to the Population Prediction of Each Province in China *Journal of Nanjing Normal University (Natural Science Edition)* 32(3) 119-124 (in Chinese)

[4] Wang Y X, Wang H, Xiao J 2010 Forecast on population distribution of Shanghai pension system based on the grey GM(1,1) model *Systems Engineering- Theory & Practice* 30(12) 2244-53 (in Chinese)




[5] Ren Q , Hou D D 2011 Stochastic Model for Population Forecast: Based on Leslie Matrix and ARMA Model *Population Research* 35(2) 28-42 (in Chinese)

[6] Shan R, Wang S H, Li L L, Gao D L 2012 Combination Model Based On ARIMA, BP Neural Network And GM *Journal of Liaoning Technical University(Natural Science)* 31(1) 118-22 (in Chinese)

[7] Liu M W, Li X, Liu T 2010 A Gene Expression Programming Algorithm for Population Prediction Problems *Acta Scientiarum Naturalium Universitatis Sunyatseni* 49(6) 115-120 (in Chinese)

[8] Song J, Yu J Y 1985 *Population Control Theory* Science Press: Beijing (in Chinese)

[9] Jiang Y Y, Wang X 2011 Application on CHINA’s Population Prediction Based Population Development Equation Model *Statistics and Decision* 27(15) 52-4 (in Chinese)

Authors	
	<p>Qiong Gu, born on March 31, 1973, Jingmen, Hubei Province</p> <p>Current position, grades: associate professor University studies: received the B.Sc. in Accounting from Beijing Institute of Technology in 2002, She received her M.Sc. in Computer Science and Technology and PhD degree in Geosciences Information Engineering from China University of Geosciences, Scientific interest: She research interest fields include data mining, Machine Learning, and internet of things and nonparametric statistics Publications: more than 20 papers published in various journals. Experience: She has teaching experience of 5 years, has completed four scientific research projects.</p>
	<p>Wu Zheng, born on March 14, 1990, Enshi, Hubei Province</p> <p>Current position, grades: Student University studies: received the B.Sc. in Mathematics from Hubei University of arts and Science Scientific interest: His research interest fields include: Statistics Experience: He has teaching experience of 1 years.</p>
	<p>Xianming Wang, born on November 11, 1979, Huanggang, Hubei Province</p> <p>Current position, grades: the lecturer of School of Oujiang, Wenzhou University, China. University studies: received his B.Sc. in Physics from Yunnan University in China. He received his M.Sc. from Chinese Academy of Sciences in China. Scientific interest: His research interest fields include Web data processing, Web data mining Publications: 5 papers published in various journals. Experience: He has teaching experience of 5 years, has completed two scientific research projects.</p>

A study on fast assessment of medium and small earthquake

Dong-ping Li*, Xiangsheng Kong

Earthquake Administration of Zhejiang Province, Hangzhou, China, 310013

Received 1 March, 2014, www.tsi.lv

Abstract

The key to rapid assessment of earthquake losses is to identify the seismic intensity area. The information about the scale of earthquake may help the government and the relative department to make countermeasures, dispose disaster rescue action and strive for foreign aid. In this paper, the data of history earthquakes of Zhejiang Province and its surrounding areas, after being processed by GIS system, are used to access the length of the earthquake axis parameters. Then the data are compared with the tectonic structure of the area to determine the classification. After that, the rapid assessment of earthquake model is applying to the axis parameters of earthquakes, which have impact in Zhejiang Province. The model can provide references of earthquake rapid assessment.

Keywords: Fast Earthquake Loss Assessment, GIS, Earthquake disaster emergency

1 Introduction

After the destructive earthquake, using the existing technical conditions may quickly draw the longitude, the latitude and the magnitude of earthquake. However, the rapid assessment requires more detailed information, for example, the Intensity distribution. Based on the intensity distribution and the emergency database, the loss of earthquake, which contains the casualties and the property loss, may be estimated. This evaluation mode is more meaningful to the government.

In order to get the Earthquake Intensity Attenuation Model of Zhejiang province and the surrounding area, we collect great references of historical earthquake intensity material. After analysing these data statistically, we get the Earthquake Intensity Attenuation Model of Zhejiang province and the surrounding areas. Firstly, we adopt the Earthquake Isoseismal Line Drawing Diagram Evaluation Model to get the Long and Short Axis Polynomial Relation Model under different intensity influence field. Secondly, we statistically analysed the example of earthquake in Zhejiang province and the surrounding areas, and get the Seismic Influence Field Model. Through the use of mode for rapid earthquake assessment, may calculation future earthquake damage prediction [1].

2 The design of Fast Evaluation Mode

After the earthquake, making the Snap judgments needs to identify the scope the range of earthquake. introduced the Isoseismal Line Model . We will use this model in our research. After earthquake, the earthquake quick determine system can be used to measure the epicentre, the seismic intensity ring or isoseismal line. Actually, the situation of the seismic intensity ring is very complex. In

most cases, it is irregular oval shape. In order to build this model, we assume that all the isoseismal lines are regular ellipse. The VII degree circle isoseismal line will appear multiple times. It means there will be many epicentres in this earthquake. We assume that there are only one intensity ring, which is in the same level, to simplify the process. If there are several different heart intensity rings, we will select the largest as the modelling data sources. Based on the above hypothesis, according to the structure fracture and tectonic division condition of local area, we may measure the direction of the minor axis of ellipse. According to the minor axis, major axis of an ellipse to measure the magnitude of earthquake may get the parameter of the minor axis and major axis. It may quickly get the intensity ring and provide reference for the earthquake evaluation [2].

2.1 DATA REDUCTION

Zhejiang province is an Eldorado, the records of local area is widespread. Although earthquakes are not common in this part of the world, but the record of the few earthquakes is very detailed. This research is concerned to the destructive earthquake. In fact, the records about the destructive earthquake is very detailed. The 43 isoseismal maps of this research comes from the record of earthquake of magnitude 4 or over, which happened in Zhejiang province. Firstly, we use the ArcGIS to calibrate the isoseismal maps into the Xian 1980 coordinate system. Then vectorise the epicentre and the isoseismal line of different intensities. Secondly, we mark the name and earthquake magnitude of the epicentre, and the name and the intensity of isoseismal line. Thirdly, after vectorising the isoseismal line and the epicentre, we combine the epicentres into one map. Superposing this map with the boundary and earthquake structure-zoning

* *Corresponding author* e-mail: llgis@163.com

map of Zhejiang province. Then look the tendency of aggregations of earthquake in three blocks: 1) the north of Zhejiang province and the south of Jiangsu province; 2) the south of the Yellow Sea, the north of Jiangsu province; 3) the Yangtze estuary. In other blocks, the number of earthquake is too little to build the model. Actually, the earthquakes, which happened in Zhejiang province lack the record of isoseismal line. We mark these points to explain the limiting factor of modelling. The application of historical earthquake records have two limiting factors in modelling. Firstly, due to the lack of earthquake records in Zhejiang province, Jiangsu province, and Shanghai where earthquake are infrequent, the research result becomes uncertainty; Secondly, the magnitudes of earthquakes in Zhejiang province are below the standard of example which can be choose in research (magnitude lower than 6). The magnitude of earthquakes are often concentrated in 4 to 6.

After determining the research object, we add up the major and axis of isoseismal line in three blocks. In most cases, the isoseismal line is closer to elliptical. With the help of ranging function of ArcGIS, we may read the major and axis of isoseismal line. Following is the outcome.

2.2 THE MODEL AND RESULTS

The index model is suitable for the earthquake statistics. We adopt the classic index model.

$$y = ae^{bx} \tag{1}$$

It uses CurveExpert 1.38 as processing program, which match automatically by default. Although the earthquake magnitudes are distribute in the narrow region, the result is desirable [3, 4]. It is worthwhile to note that there is unusual data; these data should be deleted, for the cause of earthquakes, which happened in these areas are different from the other earthquakes. For example, in the V degree zone of the adjustable shaft model of northern Zhejiang and southern Jiangsu, we can a point, which is isolated from the other points. In addition, it is not because the large gap which between the lateral axis and earthquake magnitude. The result of curve fitting is $y_1=52.999\exp(0.293x)$. After deleting the data of this point, the result of curve fitting is $y_2=26.786\exp(0.432x)$. The shape of the curve is improved obviously. If the prediction of the magnitude of the earthquake is 6.1, then $y_1=316.57$ km, $y_2=373.56$ km, the difference obviously. The results is selected according to the above case. Following is the result of curve fitting (Fig. 1, Fig. 2, Fig. 3, Fig. 4).

TABLE 1 The Earthquake Cases Statistics

IV degree area statistics					
<i>magnitude</i>	<i>Statistical number</i>	<i>magnitude</i>	<i>Statistical number</i>	<i>magnitude</i>	<i>Statistical number</i>
the north of Zhejiang province and the south of Jiangsu province		the south of the Huanghai Sea, the north of Jiangsu province		Yangtze estuary area	
4¼	1	4.8	1	5.0	3
4.9	1	6.2	1	5½	1
5¼	1	6½	2	6.1	1
5½	2	6¾	1		
5.8	1	7.0	1		
6.0	1				
V degree area statistics					
<i>magnitude</i>	<i>Statistical number</i>	<i>magnitude</i>	<i>Statistical number</i>	<i>magnitude</i>	<i>Statistical number</i>
the north of Zhejiang province and the south of Jiangsu province		the south of the yellow sea, the north of Jiangsu province		Yangtze estuary area	
4¼	2	4.8	1		
4½	2	5½	1		
4¾	4	6.0	2		
4.9	1	6.2	1		
5.0	2	6½	3		
5¼	1	6¾	2		
5½	2	7.0	1		
5.8	1				
6.0	1				

TABLE 2 The Fitting results Model of the Seismic Intensity Attenuation

the area	Long axis	Short axis
the north of Zhejiang province and the south of Jiangsu province V degree	$y=26.786\exp(0.432x)$	$y=12.806\exp(0.458x)$
the north of Zhejiang province and the south of Jiangsu province IV degree	$y=1.720\exp(0.968x)$	$y=0.357\exp(1.187x)$
the south of the Yellow Sea, the north of Jiangsu province V degree	$y=17.225\exp(0.536x)$	$y=1.497\exp(0.856x)$
the south of the Yellow Sea, the north of Jiangsu province IV degree	$y=4.809\exp(0.808x)$	$y=6.753\exp(0.724x)$
Yangtze estuary IV degree	$y=38.004\exp(0.459x)$	$y=59.933\exp(0.298x)$

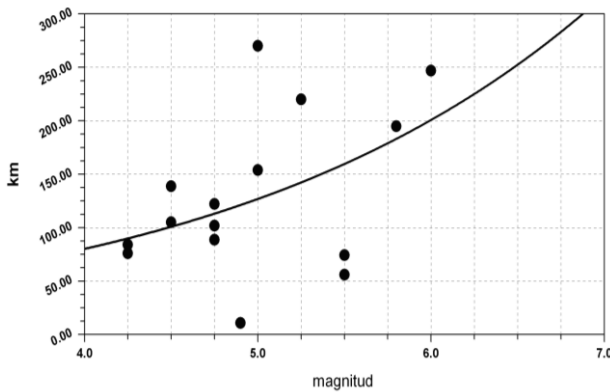


FIGURE 1 The Fitting results Model of the area of the Southern Zhejiang and the south of Jiangsu province V degree short axis

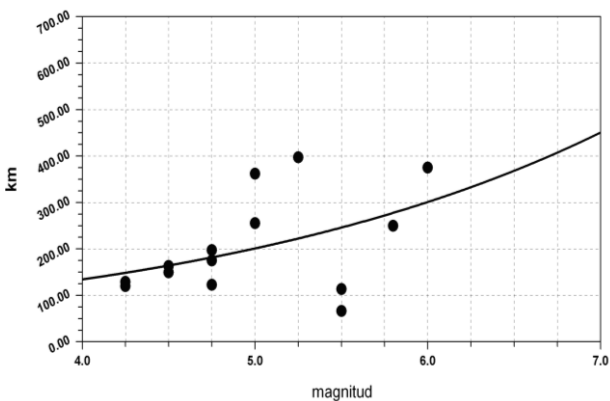


FIGURE 2 The Fitting results Model of the Southern Zhejiang and the south of Jiangsu province V degree Long axis

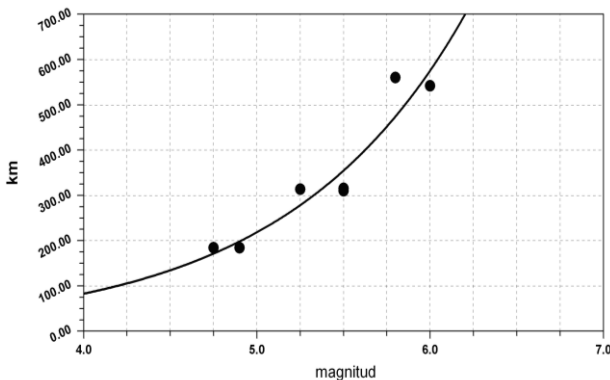


FIGURE 3 The Fitting results Model of the area of the Southern Zhejiang and the south of Jiangsu province IV degree short axis

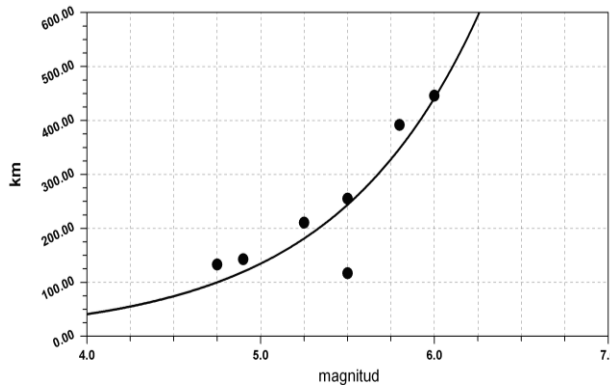


FIGURE 4 The Fitting results Model of the area of the Southern Zhejiang and the south of Jiangsu province IV degree Long axis

3 The Model of Seismic Hazard Evaluation

We can use the data and the divivable model of earthquake to make the rapid evaluation of seismic hazard. In the calculation, we use GIS to read the acreage of residential building in different villages and the proportion of residential building in each village. Then we use the Earthquake Intensity Attenuation Model of Zhejiang province and the surrounding area to calculate the infection of earthquake, and judge the distribution of population and the Gross Indices of Macroeconomic Operations which may belong to corresponding area of seismic intensity [5, 6].

3.1 THE CALCULATION OF ECONOMIC LOSSES

We think that the economic loss caused by damage mainly because building direct economic losses, with the following formula said

$$L(I) = \sum_{s=1} \sum_{j=1} b_s(j)\beta_s(j) + \sum_{s=1} \sum_{j=1} Q_s(j)W_s \quad (2)$$

In this formula, “j” represents the earthquake damage level. It consisted of 5 levels (sound, mild, moderate, severe and destroy); “S” represents the type of building. It consisted of 3 levels (steel, multi-layer, single); $b_s(j)$ represents the loss ratio of “s” class building in the “j” kind of damage (intensity 5, 6, 7, 8, 9, 10), it is the value ratio of the cost of rebuilding and the total value of building; $Q_s(j)$ represents the total value of the “s” class building in the “j” kind of damage; $Q_s(j)$ represents the loss ratio of the “s” kind of building damage in the “j” kind of damage. It is the value ratio of the loss of interior assets and the total value of indoor assets; $Q_s(j)$ represents the value of indoor assets in “s” kind of building.

The Fast Earthquake Loss Assessment. Actually, the distribution of building is not average. That using 1:50000 construction layer to correct. The database, which is used for the rapid evaluation of seismic hazard, is based on the GDP and the vital statistics of villages and towns. In order to use the vital statistic data adequately, we depend on the residential distribution chart and allot the information of census and the GDP data according to the proportion of the occupied acreage of residential buildings. At last, we can use the data and the divivable model of earthquake to make the rapid evaluation of seismic hazard [7, 8].

3.2 THE CASUALTIES CALCULATION FORMULA

Casualties prediction include: calculate number of the casualties and number of the injured in the affected areas in a given earthquake condition.

$$M_d(I) = c\eta(A_1r_{d1} + A_2r_{d2} + A_3r_{d3}), \quad (3)$$

$$M_h(I) = c\eta(A_1r_{h1} + A_2r_{h2} + A_3r_{h3}). \quad (4)$$

In the formulas (3) is number of the casualties and (4) is number of the injured, C is for the personnel percentage indoors in the earthquake, A_1 is for the destroyed houses' area, A_2 is for severely damaged houses' area, A_3 is for medium damaged houses' area, η is for personnel density indoors. Units: people/m²; r_{d1} , r_{h1} are respectively the mortality rate and the seriously injured rate in the destroyed house; r_{d2} , r_{h2} are the mortality rate and seriously injured rate in the severely damaged house; r_{d3} , r_{h3} are respectively the mortality rate and seriously injured rate in medium damaged houses. In the forecast, time can be divided into the day and the night, now suppose during the day in the earthquake indoors personnel percentage is for 40%, and during the night in the earthquake indoors personnel percentage is for 100%. The daytime is from 8:00 to 18:00, and the night time from 18:00 to 8:00 [9].

In Blind estimate system casualties calculating is similar to property losses calculating, using space superposition analysis method to calculate the casualties. First, the overlying town's clustered houses' area and separate house's area can be obtained by the space superposition method, then put the result into the

casualties calculation formula to get the number of the dead, and the injured and the homeless.

4 Conclusion

The rapid evaluation of earthquake loss uses experiential models to evaluate the loss of earthquake in a short time. The information about the scale of earthquake may help the government and the relative department to make countermeasures, dispose disaster rescue action and strive for foreign aid. If we can estimate the distribution and degree of the earthquake in a short time, the government may make a scientific decision to rescue disasters, then more people who lived in the earthquake area may be rescued, the loss of economy may be reduced. In the future, more researches can be made on the new methods of seismic hazard evaluation [10].



5 Acknowledgement

This study is supported by Science and Technology Projects of Zhejiang (No.2011C23060).

References

- [1] CHEN Zhi-rong, LI Zhao, LIU Ting 2010 Spatial information grid service model for mobile devices *Journal of Zhejiang University (Science Edition)* 5(2) 577-82 (In Chinese)
- [2] HE Xiong, GAO Yi-yang 2010 WSRF-Based Spatial Information Services State Management *Journal of Geomatics Science and Technology* 5(2) 375-8 (In Chinese)
- [3] Souhaib Besrou, Imran Ghani 2012 Measuring Security in Requirements Engineering *International Journal of Informatics and Communication Technology* 2(1) 72-81
- [4] ZHAO Gang 2009 Study of Normal Geothermal Dynamics *Earthquake* 29(3) 109-16 (In Chinese)
- [5] ZHANG Jing-fa, XIE Li-li, TAO Xia-xin 2001 Pattern analysis of remote sensing imagery for some typical earthquake damages *Journal of Nature Disasters* 10(2) 89-95 (In Chinese)
- [6] Chengjiang Lu 2012 Research on Near-Fault Problems in Earthquake Engineering *TELKOMNIKA Indonesian Journal of Electrical Engineering* 10(5) 1033-9 (In Chinese)
- [7] Mardiyono Mardiyono, Reni Suryanita, Azlan Adnan 2012 Intelligent Monitoring System on Prediction of Building Damage Index using Artificial Neural Network *TELKOMNIKA Indonesian Journal of Electrical Engineering* 10(1) 155-64
- [8] CHEN Qiang 2008 The Application of GoogleEarth in Earthquake Emergency *Earthquake* 28(1) 121-8 (In Chinese)
- [9] Keshuai XU, Weiguo CUI 2012 A GIS-based Framework for Improving the Rural Settlement System in Dancheng County of Henan Province *Agricultural Science & Technology* 13(7) 1598-602 (In Chinese)
- [10] QIAN Yong-tao 2012 Emergency GIS System Applying By RIA *Science Technology and Engineering* 20(23) 5903-7 (In Chinese)

Authors

	<p>Li Dongping, born in June 22, 1976, Hangzhou, China</p> <p>Current position, grades: Senior engineer in Earthquake Administration of Zhejiang Province University studies: Geographic Information System in Northwest Normal University Scientific interest: Geographic Information System and Earthquake Emergency Command System Publications: 1 Patents, 15 Papers Experience: In 2003, graduated from institute of geography and the environment of northwest normal university and Get a master of science degree. In the same year entered the Earthquake Administration of Zhejiang Province in the 10 years earthquake emergency command work.</p>
	<p>Kong Xiangsheng, born in July 18, 1978, Beijing, China</p> <p>Current position, grades: Research Scholar in the Department of Computer Engineering, Xinxiang University supervisor in Xi'an Jiaotong University University studies: Mechanical and Electronic Engineering in Shandong University Scientific interest: system analysis & design and software testing Publications: 3 Patents, 9 Papers Experience: He received B.E. Degree in 2003 from PLA Information Engineering University, China. He has teaching experience of 11 years. He has published one PHP book, one JAVA book, two MySQL books and five international papers.</p>

Deep neural network based load forecast

Wan He*

State Grid Energy Research Institute, State Grid Corporation of China

Received 1 March, 2014, www.tsi.lv

Abstract

Accurate electrical load forecast has great economic and social value. In this paper, we study deep neural networks based load forecast approaches. We first analyse the critical features related to load forecast. Then we present details of deep neural networks and pre-training technologies, including RBM pre-training and discriminative pre-training. We compare the performances of different neural network models and show the advantages of the proposed methods using a rather large data set of loads.

Keywords: Load Forecast, Deep Neural Networks, Pre-training, RBM

1 Introduction

In modern society, electrical power plays a key role in supporting national economy. Moreover, with the rapid development of China's economy and society, the demand for electricity is continuously growing. However, within a certain period, the electricity demand may be affected by a number of factors and varies greatly. Moreover, an important feature of electricity is that it is difficult to store it. Thus, when the electricity supply is greater than consumption, it will bring wastes and cause losses to power companies. Therefore, we need to forecast the load of electrical power, to guide electricity production and scheduling.

Because of the important economic and social significance, great attention has been paid on electrical load prediction research. The EUNITE load forecast completion held in 2001 attracted 56 registered competitors from 21 countries. In addition, 105 teams from well-known universities and companies, such as Cambridge University and Petrobras, joined the load forecast subtask in Global Energy Forecasting Competition 2012 (GEFCom2012).

From a broad perspective, load forecasting is a time series prediction problem, with many similar problems, such as stock price prediction, oil risk prediction, etc. According to different planning horizon lengths, load predictions can be roughly classified into three categories: one hour to one week for short-term forecast, a week to a year for medium forecasts, and longer than a year for long-term forecasts [1]. We focus on short-time prediction in this paper.

To improve the performance of load forecasts, various approaches have been proposed. Mbamalu and El-Hawary [2] used regression analysis and they represented load as a function of factors such as time and weather conditions, while Kiartzis and Bakirtzis [3] introduced fuzzy logic to predict daily peak load. Since the sequence

of load data can be described as a time series, prediction methods for time series, which are widely used in economics can be applied. Taylor considered several recently developed exponentially weighted methods for load forecasting and compared their performances [4]. Huang and Shih [5] proposed an autoregressive moving average (ARMA) model including non-Gaussian process considerations and applied it on a practical power system.

Machine learning methods have also been widely used in load forecast. The most commonly used machine learning methods include Support Vector Machines [6], Gaussian process [7] and artificial neural networks [8]. Load forecast is a very complicated problem, which is highly nonlinear and has no simple analytical formulas. Neural networks are quite suitable for forecasting loads for they can also be highly nonlinear, and can approximate any complex function when they have enough nodes in hidden layer or have enough number of hidden layers. Since Peng et al. first proposed a neural network based approach to tackle the influence of holidays and drastic changes in weather patterns [8], a lot of researchers employed neural networks to handle various problems in load forecast [9, 10]. The method proposed in this paper is also based on neural network.

Although a neural network with only one hidden layer can represent arbitrarily complex functions when the number of hidden layer neurons is large enough, a network with multiple hidden layers not only has many theoretical advantages, but also brings practical benefits [11, 12]. Recently, deep neural networks have achieved great success in image processing and speech recognition, and led to the rise of deep learning and the "renaissance" of neural network research. In this paper, we propose a method that is based on deep neural networks and combines a rich set of electrical load related features to predict one day ahead hourly load. We demonstrate the advantages of our method through carefully designed experiments using load data in a city of north China.

* *Corresponding author* e-mail: streamlethe@126.com

The rest of this paper is organized as follows. In Section 2, we discuss the characteristics of electrical load series and various factors related to load forecast. After that, we detail our methods in Section 3. Then we provide experiments and evaluation results in Section 4. Finally, we conclude our work in Section 5.

2 Factors for Load Forecast

The actual electrical loads are influenced by a variety of factors. In this section, we analyse some of the most important factors. On the basis of these analyses, we consider extracting representative features which are used as input of our deep neural network model for load prediction. We focus on load periodicity, time dependency, holiday effect and weather influence in the following.

Periodicity

Short-term electrical load usually exhibits remarkable periodicity. Intuitively, electricity consumption is directly associated with daily work and rest patterns of people. Indeed, by plotting the hourly loads of 7 consecutive days in figure 1, it is rather obvious that hourly loads in adjacent days have similar patterns.

Considering the daily periodicity of load sequences, three types of features can be extracted. Firstly, the average load of the previous day is a significant feature since it is often close to that of current day. Secondly, loads from the same hour of the previous days are good indicators for the current hour's load. Finally, loads from the previous hours of yesterday (might extend to hours of even earlier days) are also good indicators for predicting the current hour's load. As an example, loads at 3:00 and 4:00 AM on yesterday can help forecast load at 5:00 AM today.

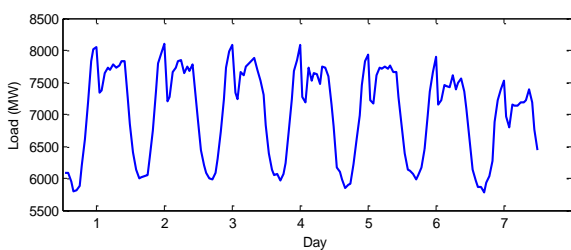


FIGURE 1 Daily periodicity of loads

In addition, load data in adjacent weeks show strong positive correlations. Figure 2 shows how load sequences of three adjacent weeks match one another. The data used in figure 2 come from hourly loads from April 4th to April 24th 2011 in our data set. So load from the same hour of same day of the previous week is used a type of feature.

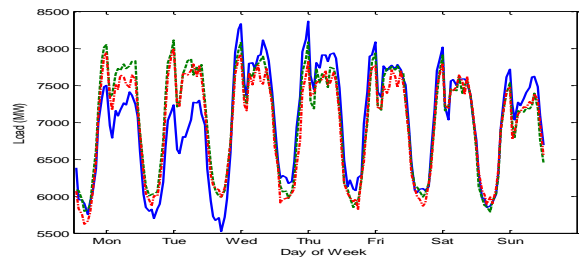


FIGURE 2 Weekly periodicity of loads

Time dependency

Figure 1 and figure 2 demonstrate how the load values change over time. So load can be expressed as a stochastic function of time. Based on this, hour of day and day of week are two features in our methods.

Holiday effect

Earlier works have pointed out that holiday is a type of factor for load. It is easy to find out that load on holidays is lower than usual. For example, in our data the average load of May 1st 2011 is 12.6% lower than that of April 25th. We add a feature to indicate whether the current day is holiday or not.

Weather influence

It is common sense that weather factors have a large impact on electrical load. On one hand, some production works are arranged according to weather conditions. On the other hand, people consume electricity to regulate temperature if it is too high or too low. Thus, daily maximum temperature and minimum temperature are used as features. Hourly temperature may be more informative, but unfortunately we cannot get them. Figure 3 show the correlation between daily average load and maximum temperatures in our 2011 data.

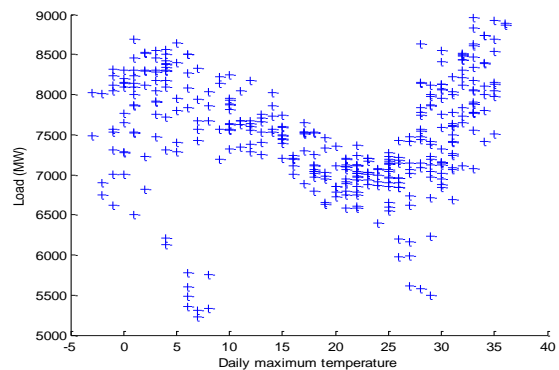


FIGURE 3 Correlation between daily average load and temperature

3 Deep Neural Network Based Load Forecast

Neural networks have the ability to learn complex nonlinear relationship between input patterns and the target to predict. The method proposed in this paper is based on deep neural networks (DNN). In general, a DNN is a hybrid model that combines traditional multilayer perceptron and recently developed pre-training

technologies. Although many papers on load forecast employ neural networks, the research of neural networks became much less hot in 1990s [12]. The situation continued until Hinton et al. proposed Deep Belief Network (DBN) and greedy layer-wise pre-training methods using Restricted Boltzmann Machine (RBM) in 2006 [11]. Deep neural networks have received extensive attention and achieved remarkable success in industry since then. In the following, we will first briefly introduce neural network (NN) and then introduce the two pre-training methods used in this paper: RBM pre-training and discriminative pre-training.

The neural network in this paper is standard feed forward multilayer perceptron structured with one input layer, one or multiple hidden layers and one output layer. The input to the network is a vector of various features extracted from the load sequence and meta-information such as weather and holiday. Each hidden layer receives a vector of input from the previous layer and converts it to its output vector through a linear transformation followed by a nonlinear activation.

Assume the set of patterns and corresponding target values are $D = \{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(T)}, y^{(T)})\}$. For each pair $(x^{(p)}, y^{(p)})$, $x^{(p)}$ is the input to the network. Since the input layer in our network is an identity mapping, its output is also $x^{(p)}$. For neuron j in hidden layer l , supposing the inputs from the previous layer are $\mathbf{a}^{(l-1)} = (a_1^{(l-1)}, a_2^{(l-1)}, \dots, a_n^{(l-1)})^T$, it outputs

$$a_j^{(l)} = f\left(\sum_{i=1}^n w_{ji}^{(l)} a_i^{(l-1)} + b_j^{(l)}\right), \tag{1}$$

where $w_{ji}^{(l)}$ and $b_j^{(l)}$ are parameters that represent weights and bias in layer l , and f is the activation function which is the sigmoid function in this paper

$$f(z) = \sigma(z) = \frac{1}{1 + e^{-z}}. \tag{2}$$

Note that in our work the output layer has a single neuron because the target to predict is hourly load, which is a scalar value. Moreover, the output layer applies only the linear transformation and uses no nonlinear activation function.

For each input feature vector $x^{(p)}$, let $o^{(p)}$ denote the corresponding output of the neural network. The expected load value $o^{(p)}$ should be close to the actual load $y^{(p)}$. The network adjusts its parameters (weights and biases) to learn good estimation of $y^{(p)}$ by minimizing the square error function $E = 1/2 \cdot \sum_{t=1}^T (o^{(t)} - y^{(t)})^2$. In the optimization process, Back Propagation (BP) algorithm is used to compute the derivatives of parameters.

The standard BP algorithm is slow and tends to get trapped in local minima. In order to find better parameters and accelerate the optimization process, we use Conjugate Gradient method, which is a kind of second order optimization algorithm [13]. We refer to the implementation of the Polack-Ribiere conjugate gradient method as in [11].

For neural networks with multiple hidden layers, the BP training with random initialization of parameters frequently gets trapped in bad parameters. In addition, in such cases, the performance of the network is usually worse than networks with a single hidden layer. To solve this problem, Hinton et al. first introduced a greedy layer-wise pre-training method [11, 12] by use of Restricted Boltzmann Machines (RBM), to find better initialization of neural network parameters. Their method has achieved huge success and lead to the rise of “deep learning”.

Restricted Boltzmann Machine is a special type of energy-based models whose energy function is bilinear [11, 12]. The energy function of an RBM is defined as:

$$E(\mathbf{v}, \mathbf{h}) = -\mathbf{b}^T \mathbf{v} - \mathbf{c}^T \mathbf{h} - \mathbf{h}^T \mathbf{W} \mathbf{x}, \tag{3}$$

where \mathbf{v} and \mathbf{h} are visible and hidden variables respectively, and \mathbf{b} , \mathbf{c} and \mathbf{W} are parameters. Because of the particular structure of RBMs, visible and hidden units are conditionally independent given one another. According to this, the conditional probabilities $p(\mathbf{h} | \mathbf{v})$ and $p(\mathbf{v} | \mathbf{h})$ can be written as

$$\begin{aligned} p(\mathbf{h} | \mathbf{v}) &= \prod_i p(h_i | \mathbf{v}) \\ p(\mathbf{v} | \mathbf{h}) &= \prod_j p(v_j | \mathbf{h}) \end{aligned} \tag{4}$$

If the values of \mathbf{v} and \mathbf{h} are further limited to set $\{0,1\}$, the conditional probabilities of variables can be expressed as,

$$\begin{aligned} p(h_i = 1 | \mathbf{v}) &= \sigma(c_i + \mathbf{W}_i \mathbf{v}) \\ p(v_j = 1 | \mathbf{h}) &= \sigma(b_j + \mathbf{W}_j^T \mathbf{h}) \end{aligned} \tag{5}$$

where \mathbf{W}_i and \mathbf{W}_j are the i -th and j -th row of \mathbf{W} , respectively. Based on formula 5, an RBM can be trained effectively using 1-step Contrastive Divergence. For more details on RBM and Contrastive Divergence, please refer to [12, 11].

RMBs are building blocks for DBNs because they share parameters with neural networks. The parameters \mathbf{W} and \mathbf{c} are also weights and biases in the corresponding layer of the neural network. The initial values of weights and biases of each hidden layer can be obtained by layer-wise learning of RBMs consists of the current hidden layer and the previous layer. In this situation, the current hidden layer corresponds to the

hidden variable of the RBM, while the previous layer corresponds to visible layer. Please refer to [12] and [11] for details on the construction of DBN using RBM and greedy layer-wise pre-training methods.

Besides RBM, we employ another type of pre-training, discriminative pre-training. Hinton et al. have pointed out that discriminative pre-training may perform better [14]. The RBM does not require the target value of each sample (pattern). Therefore, it is good for unsupervised learning. However, in our load data, all samples have the corresponding targeting load values. Therefore, we can pre-train the neural network discriminatively. We do this by starting from a network with a single hidden layer and increasing hidden layers gradually. The following pseudo-code illustrates this process.

```

Input: training data set  $D$ , hidden layer size vector  $L$ ;
for  $l$  from 1 to  $\text{length}(L) - 1$ 
    construct a neural network with  $l$  hidden layers according
    to  $L$ ;
    if  $l > 1$  then
        use  $\mathbf{W}^{(k)}$  and  $\mathbf{b}^{(k)}$  from the last run to initialize  $\mathbf{W}^{(k)}$ 
        and  $\mathbf{b}^{(k)}$ ,  $k=1, \dots, l-1$ ;
    end
    randomly initialize other parameters;
    train the network using data set  $D$ ;
    save parameters  $\mathbf{W}^{(k)}$  and  $\mathbf{b}^{(k)}$ ,  $k=1, \dots, l$ ;
end
Output: values of parameters  $\mathbf{W}^{(l)}$  and  $\mathbf{b}^{(l)}$ ,  $l=1, \dots,$ 
 $\text{length}(L) - 1$ ;
    
```

4 Experimental Results

In this section, we describe the evaluation of the proposed DNN based load forecast methods. First, we briefly introduce how we create the data set. Then we present our experimental results with quantitative evaluation of our methods on this data set and some qualitative discussions.

We prepared a data set by collecting hourly load data of a city in north China. Then we collected weather information about this city on the Internet. The final data set contains load and weather from February 10st 2000 to November 30th 2012. Hourly temperature is of great significance for predicting load of that hour. However, we can only get the highest and lowest daily temperature. The data in November 2012 are used as testing data, while data in October 2012 are used as validation data. All other data are used as training data for the neural networks. Thus, we have 23136 training samples (hourly load to predict), 744 validation samples and 720 testing samples.

Our task is to provide one day ahead prediction of hourly loads. In other words, we are to forecast the load at an hour using weather information of the targeting hour and historical load information of 24 hours before. For example, to predict the load at 12:00 on July 7th, we can only use historical loads up to 12:00 on July 6th.

In section 2, we have described the types of features used in our system. The final feature representation of each sample is a vector of 10 dimensions: 3 dimensions for history load at the same hour in previous days, 1 for average load of previous day, 1 for load at the same hour of the same day in previous week, 2 for daily highest and lowest temperatures respectively, 1 indicates day of week and 1 indicates hour of day, and 1 indicates whether it is a holiday or not. So the input layers of the neural networks have 10 neurons. We have tried different number of hidden layers and different size for each hidden layer. The best performed network has three hidden layers and each hidden layer has 30 neurons.

The common way to train a DNN contains two steps, an optional pre-training step and a final tuning step. For pre-training, we have tried RBM pre-training and discriminative pre-training, and a network with no pre-training. The performances of different methods are presented in the following table, where performance of a very strong baseline with only one hidden and 100 hidden neurons is also provided.

TABLE 1 Performance of different methods

Method	MAPE (%)	MAE (MW)
NN single hidden layer	2.07	179.0
DNN no pre-train	2.08	178.1
DNN RBM pre-train	1.98	169.7
DNN disc pre-train	1.90	164.2

The performance metrics are mean absolute percentage error (MAPE) and mean average error (MAE), which are computed according to formula (6) and formula (7), respectively,

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y^{(i)} - o^{(i)}}{y^{(i)}} \right| \times 100, \tag{6}$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y^{(i)} - o^{(i)}|, \tag{7}$$

where $y^{(i)}$ is the actual value of load, $o^{(i)}$ is the prediction given by neural network, and N is the number of testing samples. We can see from table 1 that either the DNN using RBM pre-training or the one using discriminative pre-training outperforms other methods. In addition, the DNN using discriminative pre-training performs the best. Moreover, even without temperature of each hour, the DNN based methods can produce very promising prediction results.

In the following figure, we presented the actual load values and the forecasted values for the testing data, so as to get an intuitive vision. We can see that the forecasted load is very close to the actual value in most of the cases. By further analysing the data, we find that our methods perform much worse for weekend loads. The MAPE for is 2.93% for weekends in the testing set. This suggests that we need to put more efforts on improving weekend load forecast in the future.

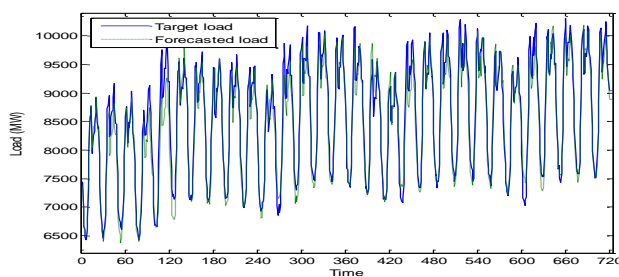


FIGURE 4 Actual load versus forecasted load

5 Conclusions

In this paper, we proposed to forecast electrical load using deep neural networks, which have attracted much attention in research and achieved great success in

References

- [1] Feinberg E A, Genethliou D 2005 *Applied Mathematics for Restructured Electric Power Systems: Optimization, Control, and Computational Intelligence* Springer 269-85
- [2] Mbamalu G A N, El-Hawary M E 1993 Load forecasting via suboptimal seasonal autoregressive models and iteratively reweighted least squares estimation *IEEE Transactions on Power Systems* 8(1) 343-8
- [3] Kiartzis S J, Bakirtzis A G 2000 A fuzzy expert system for peak load forecasting: application to the Greek power system *Proceedings of the 10th Mediterranean Electrotechnical Conference* 3 1097-100
- [4] Taylor J W 2012 Short-term load forecasting with exponentially weighted methods *IEEE Transactions on Power Systems* 27(1) 458-64
- [5] Huang S J, Shih K R 2003 Short-term load forecasting via ARMA model identification including non-Gaussian process considerations *IEEE Transactions on Power Systems* 18(2) 673-9
- [6] Chang M W, Chen B J, Lin C J 2001 EUNITE network competition: electricity load forecasting
- [7] Lloyd J R 2012 Hierarchical load forecasting using gradient boosting machines and Gaussian processes *GECom2012*
- [8] Peng T M, Hubele N F, Karady G G 1992 Advancement in the application of neural networks for short-term load forecasting *IEEE Transactions on Power Systems* 7(1) 250-7
- [9] Zhu Y, Fang G 2012 Short-term load forecasting based on dynamic adaptive artificial neural network *Power System Protection and Control* 40(1)
- [10] Shi H, Niu X, Lu Y 2012 The short-term load forecasting model based on bayesian neural network *Chinese Journal of Management Science* 20(4) 118-24
- [11] Hinton G E, Osindero S, Teh Y 2006 A fast learning algorithm for deep belief nets *Neural Computation* 18 1527-54
- [12] Bengio Y 2009 *Learning Deep Architectures For AI: Foundations and Trends in Machine Learning* 1-127
- [13] Nocedal J, Wright S J 2006 *Numerical Optimization* Second Edition
- [14] Mohamed A, Dahl G E, Hinton G E 2012 Acoustic modeling using deep belief networks *IEEE Transactions on Audio, Speech, and Language Processing* 20 14-22

Authors



Wan He, born in December 2, 1981, Beijing, China

Current position, grades: PhD of economics, senior researcher in State Grid Energy Research Institute of State Grid Corporation of China

University studies: Quantitative Economics in Renmin University of China

Scientific interest: Electricity Markets and Power System Economics, Country Risk Analysis

Publications: 10 papers

Experience: She is a senior researcher in State Grid Energy Research Institute of State Grid Corporation of China. She got her doctor's degree from Renmin University of China in 2010. Her research interests include energy economics, power system economics and country risk control and analysis.

Retrieving product information of collaborative enterprises based on Bayesian network

Junxiang Tu*

College of Mech. Eng., Fuzhou University, Fuzhou, P. R. China,

Fujian Haiyuan Automatic Equipments Co., Ltd., Minhou Fujian, P. R. China

Received 1 March, 2014, www.tsi.lv

Abstract

There exist many differences in nomenclature and descriptions of products and parts in collaborative enterprises, which greatly hinder the retrieval and sharing of web-based product information. In this paper, we present an extended Bayesian network for retrieving and integrating the product information of collaborative enterprises based on product ontology. This approach not only reduces the complexity of existing ontology mapping methods, but also increases the efficiency of product information integration.

Keywords: Product Information Retrieval, Bayesian Network, Ontology, Collaborative Enterprises

1 Introduction

With economic globalization and the development of network technology, the demand for collaborative manufacturing is increasing rapidly. Web-based product information sharing is the cornerstone for the implementation of collaborative manufacturing. However, due to the different enterprise cultures, there are many differences in nomenclature and descriptions of products and parts in different enterprises. The differences greatly hinder retrieving and sharing of enterprise products information and become a bottleneck in collaborative manufacturing development.

The methods of retrieving and integrating web-based product information of different enterprises can be summarized as name-based matching methods [1] and rule-based matching methods [2]. Name-based matching methods have been the most widely used solutions to retrieving product information of collaborative enterprises. Li [3] proposes the use of "term matching" method to explore a database for publishing content on the Web. Wang and Zhang [4] introduce a similar approach that allows querying databases through keywords. However, these methods are difficult to identify synonyms or semantically similar terms, so that there is the problem of poor accuracy with name-based matching methods. Differently, the rule-based matching methods aim to analyse the semantics of keywords based on language grammar rules, which are more targeted and accurate. Greiff [5] proposes an inference network model that allows for structured queries via a rich set of probabilistic operators. Models based on language analysis have been applied to information retrieval [6, 7]. However, the rule-based matching methods have not been

widely used in practice because of complexity of algorithms and lack of scalability.

Bayesian network is used to dealing with uncertainty in artificial intelligence [8, 9]. This paper combines it with the product ontology to construct a product information retrieval model. The model utilizes fuzzy matching of ontology elements to automate product information retrieval and integration without building complex matching rules for semantic analysis.

2 Product Ontology and Product Information Retrieval

2.1 PRODUCT ONTOLOGY

For retrieving and integrating product information between collaborative enterprises, the respective product ontology of different enterprises should be built. According to Gruber [10], ontology is the specification of conceptualizations, used to help programs and humans share knowledge. The product ontology can be expressed as the following four-tuple of the form: $O = (C, R, I, A)$, where O represents the ontology to be defined, C is the set of concepts in the ontology, R is the relationship between concepts, I represents the set of instances and A is the axiom asserted.

The function mapping from product ontology O_1 to product ontology O_2 can be defined as follows:

$F_m = (\{n_1\}, \{a_1\}, \{r_1\}) \rightarrow (\{n_2\}, \{a_2\}, \{r_2\})$, where

$\{n_1\}$: the collection of product object names of O_1

$\{a_1\}$: the collection of product object attributes of O_1

* Corresponding author e-mail: junxtu@163.com

$\{r_1\}$: the collection of relationships between product objects in product ontology O_1
 $\{n_2\}$: the collection of product object names of O_2
 $\{a_2\}$: the collection of product object attributes of O_2
 $\{r_2\}$: the collection of relationships between product objects in product ontology O_2

2.2 Product Information Retrieval and Sharing

As shown in Fig. 1, the process of collaborative enterprises product information retrieval and sharing include three steps: importing target product ontology of collaborative enterprises, ontology mapping and product information integration.

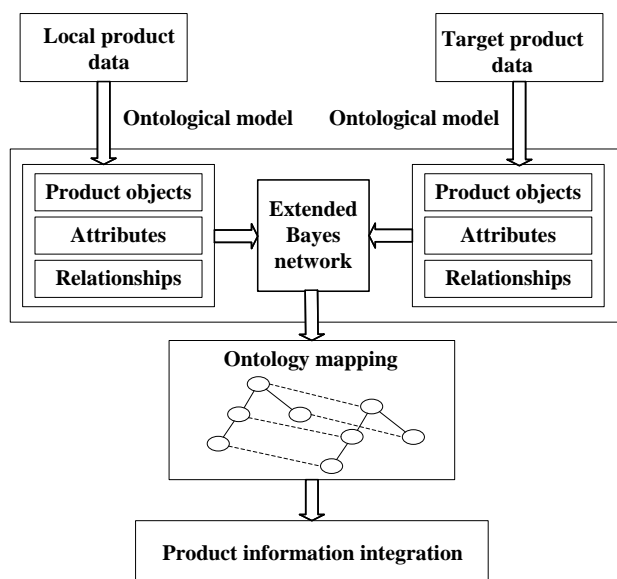


FIGURE 1 Retrieval and sharing of product information

Ontology mapping plays a key role in product information retrieval and integration [11]. Its main aim is to find semantic relationships between local product ontology and target product ontology. In this paper, product ontology mapping is implemented through an extended Bayesian network model, which can automatically discover the elements of target product ontology matching with the ones of the local ontology.

Product information integration is to integrate the product information of collaborative enterprises based on the establishment of ontology mapping after searching web-based product data.

3 Ontology Mapping and Extended Bayesian Network

3.1 BAYESIAN NETWORK

A Bayesian network is a graphical model that encodes probabilistic relationships among variables of interest [12]. The relevance of associated variables is characterized by joint probability that quantifies the interdependencies of the variables [13]. A typical

Bayesian network is demonstrated in Fig. 2, where nodes represent random variables, directed edges between nodes represent causal relationships between variables [14].

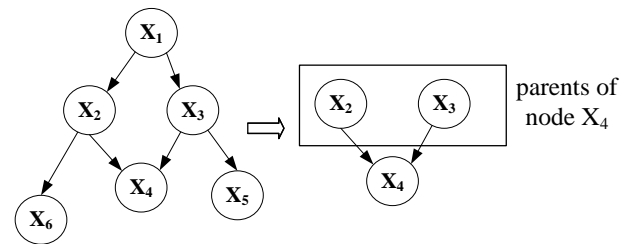


FIGURE 2 Example of a Bayesian network

We use P_{ai} to denote the parents of any node x_i in the network, and $P(x_i | P_{ai})$ to denote the conditional probability at the node x_i . For a set of variables $X = (x_1, x_2, \dots, x_n)$, the joint probability for X is given by $P(X) = \prod_{i=1}^n P(x_i | P_{ai})$.

3.2 EXTENDED BAYESIAN NETWORK

To match the elements of target product ontology of collaborative enterprises with the ones of the local product ontology efficiently, we design an extended Bayesian network composed of the probabilistic inference and the semantic inference. The model shown in Fig. 3 is divided into three layers: the local product ontology layer, the root layer and the target product ontology layer.

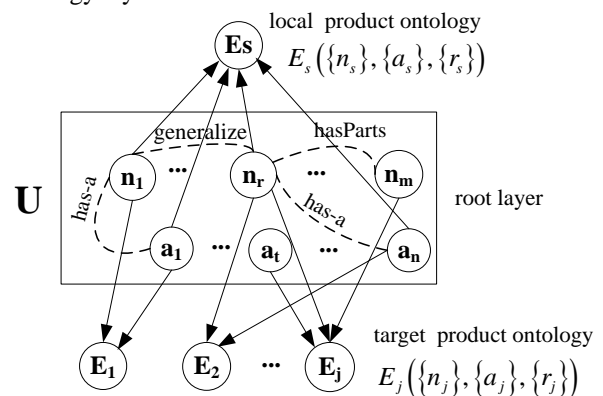


FIGURE 3 Extended Bayesian network

The root layer consists of terms of product objects, product object attributes and their semantic relationships. The nodes $n_r \in U$ ($r = 1, \dots, m$) in the root layer are terms that represent names of product objects in collaborative enterprises. The nodes $a_t \in U$ ($t = 1, \dots, n$) in the root layer represent attributes of product objects. Semantic relationships between these nodes signified by dotted lines includes synonym relationship, generalization relationship, composition relationship, and the relationship between product objects and their attributes (has-a). A generalization relationship is a relationship in which one term (the child) is based on another term (the

parent). A composition relationship represents whole-part relationship among terms.

3.3 ONTOLOGY MAPPING PROCESS

Ontology matching can be seen as a process finding the matching elements of different product ontologies, which is the process of calculating the conditional probability $P(E_k | E_s)$, as follows:

$$P(E_k | E_s) = \eta \sum_u P(E_k | u)P(E_s | u)P(u),$$

where the set u is used to refer to any of the 2^{m+n} possible states of the root nodes. If $P(E_k | E_s)$ is greater than the specified threshold, the element E_k is considered to match with E_s .

First, we define the value of $P(u)$ as a constant $(1/2^{m+n})$ because there is no a priori preference for any set of terms and attributes (subsets of U), as follows:

$$P(u) = \left(\frac{1}{2}\right)^{m+n}.$$

Second, according to vector space model, we can calculate $P(E_k | u)$, as follows:

$$P(E_k | u) = \frac{\sum_{i=1}^{m+n} w_{ik} \times w_{is}}{\sqrt{\sum_{i=1}^{m+n} w_{ik}^2} \times \sqrt{\sum_{i=1}^{m+n} w_{is}^2}},$$

where w_{ik} is the weight associated with the root node in the target product ontology and w_{is} is the weight associated with the root node in the local product ontology. It should be noted that different weight calculation methods would lead to different mapping strategies.

Third, in order to referring to the state of the variable k_i in u (subsets of root layer U), we define an indicator function as follows: $g_u(k_i) = \begin{cases} 1 & \text{if } k_i \in u \\ 0 & \text{otherwise} \end{cases}$.

Given the definition of $g_u(k_i)$, the probability of $P(E_s | u)$ is now defined as:

$$P(E_s | u) = \begin{cases} 1 & \text{if } \forall n_r, \forall a_i, g_{\{n_r\}}(n_r) = g_u(n_r) \\ & \text{and } g_{\{a_i\}}(a_i) = g_u(a_i) \\ 0 & \text{otherwise} \end{cases}$$

4 Experiments

Fig. 4 shows the production structure tree of a drilling machine produced by a machine tool enterprise. The drilling machine is mainly composed of drill head, gearbox, electric motor, column and base unit. Most of parts of the product are produced by the enterprise itself, some parts such as electric box are produced by collaborative enterprises.

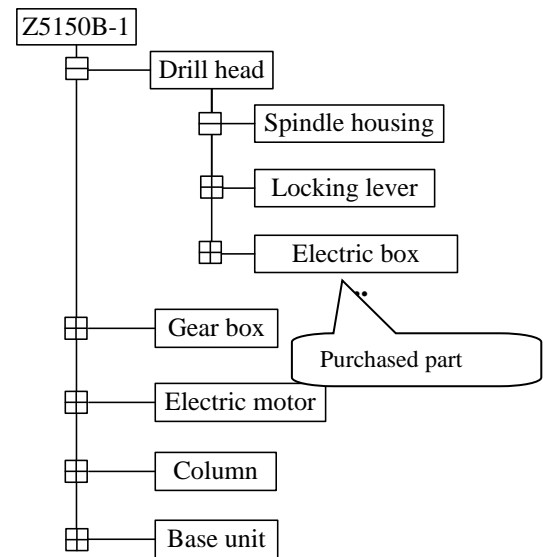


FIGURE 4 Production structure tree of a drilling machine

After importing target product ontology of a collaborative enterprise, we applied our extended Bayesian network aforementioned to ontology mapping. Table 1 lists matching elements of two product ontologies. The last column of the table indicates the credibility of the matching.

TABLE 1 Matching elements of two product ontologies

Element of local product ontology	Element of target product ontology	Credibility
Locking lever	Locking handle	1.00
Front cover	Front lid	0.91
Electric Box	Electric control cabinet	1.00
Engine	Electric motor	0.99
Ball bearing	Bearing	0.86
Table	Workbench	0.80
Gear box	Main spindle box	0.95
Base plate	Base	0.83
Spacer	Bushing	0.88

The adjusted product structure tree is shown in Fig. 5 according to the ontology mapping, where “Electric Box” and “Electric control cabinet” has established a matching relationship.

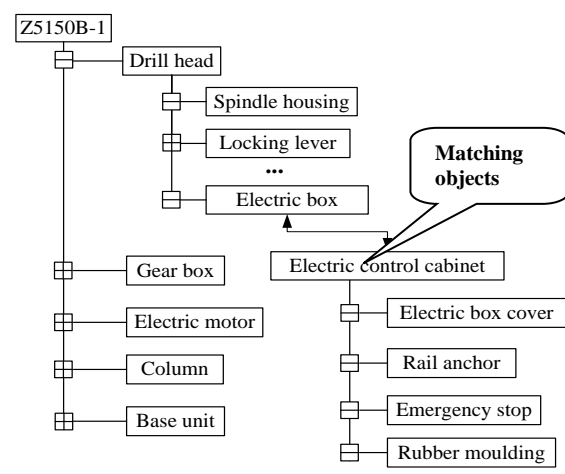


FIGURE 5 Adjusted product structure tree of the drilling machine

5 Conclusions

In this paper, we have proposed an extended Bayesian network model for collaborative enterprises product information retrieval and integration, which can efficiently deal with the problem of semantic differences of product descriptions in different enterprises. The model provides a novel approach to integrate the product

information of collaborative enterprises based on ontology fuzzy matching without building complex grammar rules.

As future work, the research on the evaluation of the approach for product information retrieval remains necessary, which helps to further improve the proposed method.

References

- [1] Calado P, Ribeiro-Neto B, Ziviani N, Moura E, Silva I 2003 Local versus global link information in the web *ACM Transactions on Information Systems* **21**(1) 42-63
- [2] Sudarsan R, Fenves S J, Sriram R D, Wang F 2005 A product information modeling framework for product lifecycle management *Computer Aided Design* **37**(13) 1399-411
- [3] Li W, Zhao T, Wang X 2010 Context-sensitive query expansion *Journal of Computer Research and Development* **47**(2) 300-4
- [4] Wang S, Zhang K L 2005 Searching databases with keywords *Journal of Computer Science and Technology* **20**(1) 52-62
- [5] Greiff W R, Croft W B, Turtle H 1999 PIC matrices: A computationally tractable class of probabilistic query operators *ACM Transactions on Information Systems* **17**(4) 367-405
- [6] Revuelta-Martinez A, Rodriguez L, Garcia-Varea I, Montero F 2013 Multimodal interaction for information retrieval using natural language *Computer Standards & Interfaces* **35**(5) 428-41
- [7] Kolomiyets O, Moens M F 2011 A survey on question answering technology from an information retrieval perspective *Information Sciences* **181**(24) 5412-34
- [8] Boughanem M, Brini A, Dubois D 2009 Possibilistic networks for information retrieval *International Journal of Approximate Reasoning* **50**(7) 957-68
- [9] Liao W, Ji Q 2009 Learning Bayesian network parameters under incomplete data with domain knowledge *Pattern Recognition* **42**(11) 3046-56
- [10] Gruber T R 1995 Toward principles for the design of ontologies used for knowledge sharing *IEEE Transactions on Power Electronics* **43**(5/6) 907-28
- [11] Panetto H, Dassisti M, Tursi A 2012 ONTO-PDM: Product-driven ontology for product data Management interoperability within manufacturing process environment *Advanced Engineering Informatics* **26**(2) 334-48
- [12] Heckerman D, Mamdani A, Wellman P M 1995 Real-world applications of Bayesian networks *Communications of the ACM* **38**(3) 24-6
- [13] Dogan I 2012 Analysis of facility location model using Bayesian networks *Expert Systems with Applications* **39**(1) 1092-104
- [14] Hemmecke R, Lindner S, Studeny M 2012 Characteristic imsets for learning Bayesian network structure *Characteristic imsets for learning Bayesian network structure* **53**(9) 1336-49

Authors



Junxiang Tu, born in November 12, 1971, Jiangxi, China

Current position, grades: Doctor of Mechanical Engineering, lecturer in Fuzhou University
University studies: Mechanical Engineering in Huazhong University of Science & Technology
Scientific interest: Computer integrated manufacturing system

Publications: 1 Patent, 11 Papers

Experience: Lecturer of Fuzhou University, China, 2010-Present; Ph.D. Mechanical Engineering, Huazhong University of Science & Technology, China, 2010; Product engineer in Diamond Company, Guangzhou, China, 1997-2003; Assistant Engineer in Nanchang Zerowatt Electric Appliance Co., Nanchang, China, 1993-1997; B.S. Chemical and Machinery, Nanchang University, China, 1993

Geoinformatics-based study on the regionalization of ecological function in the Chaohu Lake Basin, East China

Li Wu¹, Xinyuan Wang^{2*}

¹ College of Territorial Resources and Tourism, Anhui Normal University, Wuhu, 241002 Anhui, China

² Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Haidian, 100094 Beijing, China

Received 1 March 2014, www.tsi.lv

Abstract

Ecological function regionalization is a kind of geographic spatial division, which is based on the spatial differentiation of ecosystem functions. Based on an analysis of the primary features of the ecological environment of Chaohu Lake Basin in Anhui Province, the principles, bases, methodology and nomenclature of ecological function regionalization were determined. As the sub-valley is an independent geographical unit within the lake basin, its ecosystem sustains ecological integrity from the upstream through to the downstream. Therefore, ensuring the monitoring and management of the regional ecological environment in the sub-valley unit is of great importance to the conservation and ecological restoration of the regional ecosystem. Through extraction of land use information from remote sensing data, and sub-valley division from DEM analysis, this paper discusses the methodology of sub-valley ecological function regionalization in the research area based on the application of geoinformatics technology (e.g. RS and GIS technology). The ecological function regionalization of the Chaohu Lake Basin is then calculated, and the five ecological function regions and twelve sub-regions are subdivided. This study has an important practical relevance for the integrated management of the ecological environment of the Chaohu Lake Basin, and provides scientific grounds for the improved industrial distribution, ecological hazard prevention and reduction, environmental protection and construction planning in this area.

Keywords: RS and GIS, ecological function regionalization, ecological environment, Chaohu Lake Basin

1 Introduction

The ecological function regionalization is one of the most significant methods in conducting the geographical spatial division, which is based on regional environmental resources, the ecological sensitivity, and spatial differentiation of the ecological service functions [1, 2]. By using this method, the whole area can be divided into several different ecological function regions. Therefore, its purpose is to clarify ecological and environmental problems and fragile areas, and to provide scientific grounds for improved environmental protection and construction planning, ecological security, rational use of resources, and industrial distribution in the research area. Management information and resources can also be obtained more easily by decision-making departments [3]. This is the basis and premise of partition management in the regional ecological environment [4]. Therefore, ecological function regionalization plays an important role in the government management, ecological protection guidance, and standard ecological construction regulation.

The Chaohu Lake Basin is located in the central Jiang-Huai region of the Anhui Province (30°58'00"~32°06'00"N, 116°24'30"~118°00'00"E), which is adjacent to the fluvial plain along the Anhui section of the Yangtze River Valley in the southeast, and

Huaihe River Basin in the northwest. It is bounded by Mt. Fucha, Huangshan, Fenghuang, Yinping, Yefu, Dabie, and the Jianghuai watershed. Administratively, it belongs to the Hefei City, the Shucheng County of Lu'an City, and the Yuexi County of Anqing City, with an area of 9131 km² [5]. This paper's research area is limited to the upstream of Chaohu sluice. The downstream river systems of the Chaohu sluice are not considered in this study.

2 Materials and Methods

Based on the theories of landscape ecology and the natural-social-economic composite system, the ecological function regionalization should be calculated in accordance with the following principles [1-10]: the principle of sustainable development, the principle of similarity and differentiation, the principle of comprehensiveness, the principle of holistic integrality, the principle of genealogy, and the principle of adjustability. There is a long history of human activity in the Chaohu Lake Basin. Therefore, along with the natural sub-area, social and economic factors should also be considered. Since ecological function regions are constantly changing, the ecological function regionalization should also consider time factors. Thus, historical factors require the adjustment of ecological

* Corresponding author e-mail: xywang@ceode.ac.cn

function regions, so as to allow for changing situations in the ecological environments. Since the sub-valley is an independent geographical unit within the lake basin, its ecosystem sustains ecological integrity from the upstream through to downstream. Ensuring the monitoring and management of regional ecological environment in sub-valley unit is therefore of great importance to the ecological restoration and conservation of regional ecosystems [11, 12]. However, most sub-valleys cover multiple administrative or economic units, making it difficult to obtain the thematic information of ecological environments within each sub-valley. Therefore, based on the information extraction of land use from remote sensing data, and sub-valley division from DEM spatial analysis, thematic information extraction of the ecological environment within each sub-valley can be collected. The basic ecological features of the Chaohu Lake Basin and the ecological function regionalization of Anhui Province should also be clearly considered [4, 13]. The main premise of further division includes the climate-geographic features, ecosystem types, ecosystem service function types, and ecological environment problems within each sub-valley.

The data for this paper includes: 1:50,000 digital basic geographic data of the Chaohu Lake Basin, such as the data layer of river system, residential areas, railway networks, administrative divisions, and contour lines, etc.; digital data layers for regional resources and environment derived from vectorization of 1:250,000 soil type cartography, land type cartography, and land use cartography; 1:100,000 land use map interpreted from the Landsat TM images in January 11, 2009; and the above data layers tested and updated from the Landsat TM images (Figure 1).

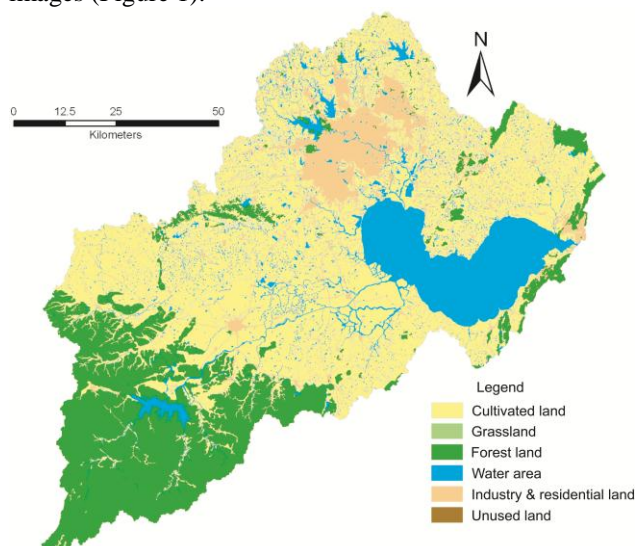


FIGURE 1 Land utilization map for the Chaohu Lake Basin

In the ArcGIS 9.2 software platform, existing vector data is edited and modified, in order to establish a unified coordinate system. Altimetric data is used for the TIN's construction, and then interpolated to generate DEM with 25 m resolution (Figure 2). Based on the above altimetric

data and the DEM, sub-valleys are extracted and divided using the gully auto-extract method [14] in the research area. The major sub-valleys include the Hangbu-Fengle-Baishishan River valley, the Nanfei-Dianbu River valley, the Pai River valley, and the Zhegao River valley (Figure 3). There are other shorter gullies developing from lake-swamp plains surrounding the Chaohu Lake (Figure 3).

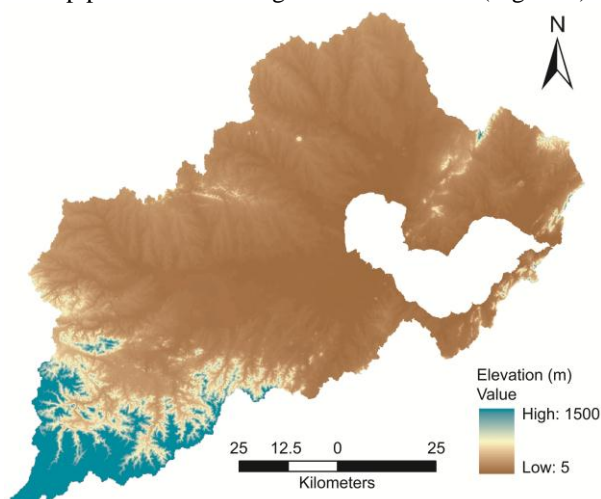


FIGURE 2 25 m grid DEM for the Chaohu Lake Basin

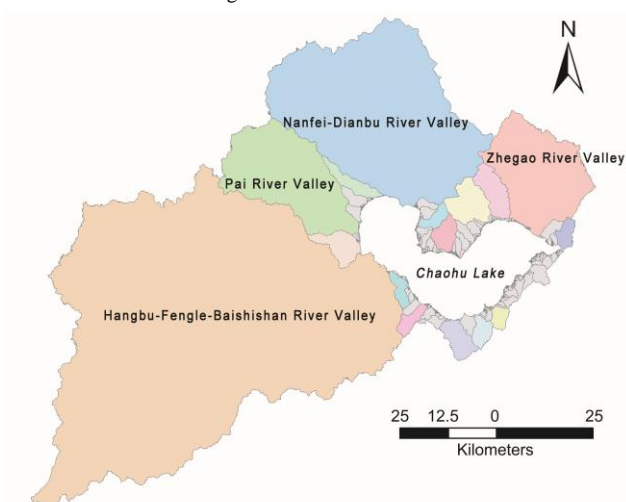


FIGURE 3 Sub-valley extraction for the study area based on DEM

The ecological function regionalization is performed using the superposition method on the basis of sub-valley division [15]. First, by using terrain information and DEM analysis, the regional gully system is extracted in an ArcGIS operating environment based on the D8 algorithm [16]. Sub-valleys are divided automatically, and then used as a main reference for the first-level division boundary. On this basis, first-level land use type's layer is overlaid. By using the multilayer superposition function of ArcGIS, resources and environmental elements are overlaid, and the administrative boundaries are separated. Most coincidence lines can be taken for boundaries. As for areas with fewer coincidence lines, key ecological elements are used for demarcation and necessary correction. Therefore, the second-level division

boundaries of the ecological function are determined. Finally, through the field investigations and confirmation, combined with the ecological function regionalization of Anhui [4], regionalization principles and rationality of the first-level and second-level ecological function regions are considered comprehensively. Necessary corrections are then processed, so as to determine the last schematic of ecological function regionalization within the whole basin. Detailed processes are shown in Figure 4.

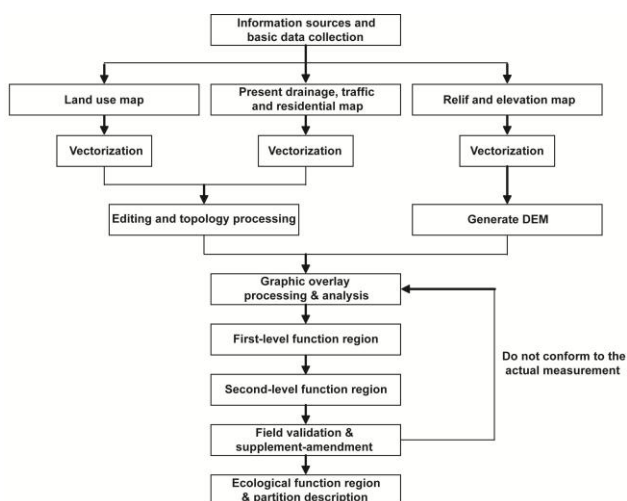


FIGURE 4 Follow chart of ecological function regionalization

The ecological function regionalization of the Chaohu Lake Basin can be divided into two levels. The dominant factor of the first-level ecological function region is the boundaries of the sub-valleys, reflecting the general pattern of ecological functions with each sub-valley as a unit. Its nomenclature can be expressed as *sub-valley name + ecological function region*. The dominant factor of the second-level ecological function region is the land use type. At the same time, characteristics of landform or climate, ecosystem type, and ecosystem service function should also be represented. The second-level ecological function region is the basic unit of ecological function type, reflecting the differentiation of ecological function within the first-level division zone. The nomenclature can be expressed as *place name + geomorphic feature or ecosystem type + ecological function feature + ecological function sub-region*.

3 Results and Discussion

In accordance with the above principles, the five ecological function regions and twelve sub-regions are subdivided in the Chaohu Lake Basin, as shown in Figure 5 and Table 1.

Through the ecological function regionalization of the Chaohu Lake Basin and its map compilation, the successful application of RS and GIS is fully demonstrated here. The advantages found [17, 18] in data collection and multilayer superposition are made full use of, so as to solve the difficult problem of ecological information extraction. This technique saves costs greatly

and improves the accuracy and efficiency of graphics, realizing classic graphic principles excellently.

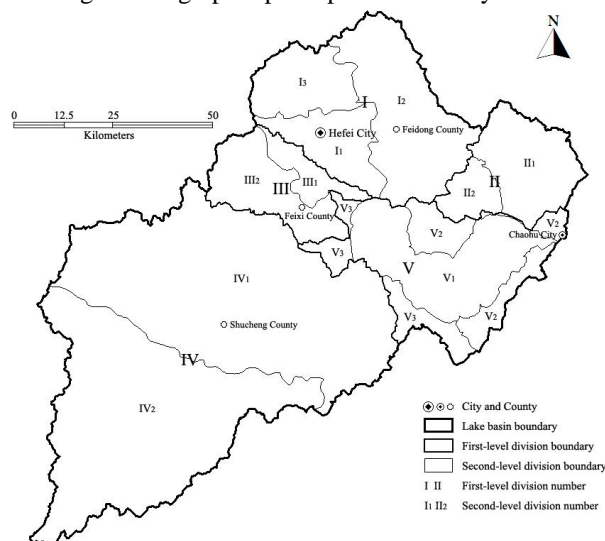


FIGURE 5 Ecological function regionalization of the study area

TABLE 1 Ecological function regions and their features in study area

First-level ecological function regions	Second-level ecological function sub-regions	Area (km ²)	Proportion (%)
I Nanfei-Dianbu River Valley	I ₁ Hefei city and suburban agriculture	366	4.0
	I ₂ Eastern hilly-plain agriculture	919	10.1
	I ₃ Southern Jianghuai Watershed hilly agriculture and soil-water conservation	390	4.3
II Zhegao-Tongyang River Valley	II ₁ Zhegao River Valley hilly and polder agriculture	530	5.8
	II ₂ Tongyang River Valley hillock agriculture	178	1.9
III Pai River Valley	III ₁ Northern town and suburban agriculture	158	1.7
	III ₂ Southern hillock-plain irrigation farming and soil erosion control	478	5.2
IV Hangbu-Fengle-Baishishan River Valley	IV ₁ Western Chaohu plain polder agriculture and nonpoint source pollution control	2664	29.2
	IV ₂ Northern Mt. Dabie mid-lower mountain forest soil-water and biodiversity conservation	2138	23.4
V Chaohu Lake and lakeside	V ₁ Chaohu Lake wetland flood water storage and agriculture	760	8.3
	V ₂ Eastern lakeside hillock ground stone mining, collapse recovery and ecological conservation	329	3.6
	V ₃ Western lakeside plain water channel wetland and polder agriculture	220	2.4

The ecological function regionalization result of the Chaohu Lake Basin indicates many vulnerable and key protection areas of the regional ecosystem, and has an important practical significance for the comprehensive

control of the basin's ecological environment, ecological hazard prevention and reduction, industrial distribution, environment protection, and the planning of construction in this area.

Acknowledgments

Authors wish to acknowledge the field investigations and other assistances from Prof. Xianfu Cheng, Dr. Chao Gao,

Mr. Guanyong Wang from the Provincial Key Laboratory of Natural Disaster Process and Prevention Research, Anhui Normal University. This research was financially supported by the National Natural Science Foundation of China (41271516), the Major Program of Natural Science Research at University of Anhui Province (ZD200908), the Doctoral Scientific Research Foundation (161070143) and Cultivating Foundation of Anhui Normal University.

References

[1] Zhou H R, Xiao D N 2010 *Journal of Arid Land* 2(2) 123-132
 [2] Ministry of Environmental Protection of the People's Republic of China 2002 *The Tentative Specifications of Ecological Function Regionalization* Environment Press: Beijing (in Chinese)
 [3] Xu X B, Zhang J M, Qi Y A, Nian Y Y 2005 *Arid Zone Research* 22(1) 41-44 (in Chinese)
 [4] Jia L Q, Ouyang Z Y, Zhao T Q, Wang X K, Xiao Y, Xiao R B, Zheng H 2005 *Acta Ecologica Sinica* 25(2) 254-260 (in Chinese)
 [5] Liu J, Li C 2003 *Chinese Science Bulletin* 48(2) 181-183 (in Chinese)
 [6] Liu G, Fu B 1998 *Advances in Environmental Science* 6(6) 67-72 (in Chinese)
 [7] Fu B J, Liu G H, Chen L X, Ma K M, Li J R 2001 *Acta Ecologica Sinica* 21(1) 1-6 (in Chinese)
 [8] Wang Z, Li P, Wang Y, Hu T, Gong Z, Sun T, Wan Z, Chen D 2005 *Chinese Journal of Ecology* 24(11) 1339-1342 (in Chinese)
 [9] Hu M C, Ma R H 2003 *Journal of Arid Land Resources and Environment* 17(1) 49-53 (in Chinese)
 [10] Research Team of "Hebei Ecology and Disaster" 2003 *Geography and Geo-Information Science* 19(5) 82-85 (in Chinese)
 [11] Hu L W, Wang X J, Luo D G, Jiang Y 2007 *Advances in Water Science* 18(2) 235-240 (in Chinese)
 [12] Tian F, Han S M, Hu Y K 2009 *Chinese Journal of Agrometeorology* 30(1) 60-65 (in Chinese)
 [13] Xie P 2009 *Reading about the Histories of Cyanobacteria, Eutrophication and Geological Evolution in Lake Chaohu* Science Press: Beijing (in Chinese)
 [14] Wang G Y 2009 *Chaohu Lake Basin Gully Auto-Extract and Gully Spatial Characteristics Study* MSC Thesis, Anhui Normal University: Wuhu (in Chinese)
 [15] Chang K T 2006 *Introduction to Geographic Information Systems* McGraw-Hill: New York
 [16] Sun C, Wang J 2008 *Progress in Geography* 27(1) 118-124 (in Chinese)
 [17] Wu L, Wang X Y, Zhou K S, Mo D W, Zhu C, Gao C, Zhang G S, Li L, Liu L, Han W G 2010 *Journal of Geographical Sciences* 20(5) 687-700
 [18] Wu L, Wang X Y, Zhu C, Zhang G S, Li F, Li L, Li S Y 2012 *Quaternary International* 275 23-29 (in Chinese)

Authors	
	<p>Li Wu, born on February 17, 1985, Hefei, China</p> <p>Current position, grades: Ph.D., lecturer in Anhui Normal University University studies: Geographic and oceanographic sciences in Nanjing University Scientific interest: Application of remote sensing and GIS in ecological environmental change Publications: 15 SCI indexed papers</p>
	<p>Xinyuan Wang, born on June 14, 1964, Lu'an, China</p> <p>Current position, grades: Ph.D., research fellow in Institute of Remote Sensing and Digital Earth, CAS University studies: Earth Sciences and Engineering in Nanjing University Scientific interest: Processing and analysis of remote sensing image, space technologies for ecosystem Publications: 3 books, 60 papers</p>

A new method of digital manufacturing of orthoses

Yuewei Ai¹, Yan He^{1*}, Zhijian Wang², Yang Wang

¹ State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing-China

² Ningbo Institute of Material Technology & Engineering, CAS, Ningbo-China

³ Faculty of Science and Engineering, University of Nottingham, Ningbo-China

Received 1 March, 2014, www.tsi.lv

Abstract

The proportion of disabled people is rising and now represents 1 billion people–15% of the global population, which leads to increasingly demand for orthotic device. However, moulds for orthoses manufacturing through traditionally manual technique are often dedicated, and this causes problems such as long lead time, lack of flexibility, low-efficiency and material waste, further leading to serious financial burdens and environmental pollution as well. In this paper, an innovative method is proposed to replace traditionally dedicated moulds with reconfigurable moulds utilizing screw-pins that are directly transferred to the vacuum forming of thermoplastic material at low cost for the fabrication of orthoses. In the developed system, the fast reconstruction of human body anatomy based on the 3D digital scanning, is introduced firstly, the reconfigurable mould utilizing screw-pins is then generated and machined based on the reconstructed human body anatomy. After this, vacuum forming is performed on the reconfigurable mould, which could be reused for different anatomical shape variations by adjusting screw-pins. Additionally, an intelligent database is developed and a lot of reconstructed anatomies, the best practices of experienced orthotists, optimal parameters for 3D digital scanning, reconfigurable mould generation and machining and vacuum forming are stored, which will allow rapid recall of the stored information to reduce too much man-machine interaction and expertise dramatically.

Keywords: Orthoses, Digital manufacturing, Reconfigurable moulds, Fast reconstruction

1 Introduction

For the disabled and elderly, orthoses are important tools to help them recover in the clinical environment and live independently, especially for the population ageing and the rapid spread of chronic diseases. According to the United States Census Bureau report [1], global population will reach 9.3 billion in 2040, of which over 65-year-old population will be twice from 7% to 14%. Moreover, the proportion of disabled people is rising and now represents 1 billion people–15% of the global population [2]. All these demographic changing are expected to underpin a rise in the incidence of orthopaedic problems triggering significant demands for orthoses. Nowadays the conventional fabrication technique of orthoses by repairing plaster cast manually plays a dominant role in orthotics industry, which means that the quality of orthoses mainly depends on operators' experience and skills [3]. Since orthoses are generally personalized products for individual patients, thus, the manufacturing process needs one or more dedicated moulds, which can only be used once. Taking the traditional fabrication of foot orthoses as an example, the whole process is time and labour intensive, dirty and messy [4], with the making of the plaster cast (the negative cast and the positive cast), as shown in Fig.. Moreover, the natural plaster as the material of cast can only be used once and then thrown away, leading to natural resource depletion

and environmental pollution. Recent advances in computer technology, especially Reverse Engineering (RE), Computer Aided Design and Manufacture (CAD/CAM), Motion Simulation and Rapid Prototyping have been introduced into orthotics and have significantly reduced the manual work and shortened production time recently [5, 6, 7, 8, 9, 10, 11, 12]. However, the problems associated with dedicated moulds including lack of flexibility, long lead-time, material waste and low-efficiency, still remain.



FIGURE 1 Traditional manufacturing of foot orthoses

In order to improve the efficiency of moulds, the reconfigurable moulds were presented in several patents recently [13, 14, 15, 16]. Nevertheless, most of the reconfigurable moulds are only applied in forming some simple geometries and high-end products, such as the shell of aviation aircraft and the panels for a high-speed train [17, 18]. This research aims at replacing traditionally dedicated moulds with reconfigurable moulds utilizing screw-pins [19, 20, 21, 22, 23] that are directly transferred to the vacuum forming of thermoplastic materials at low cost for the manufacturing

* Corresponding author e-mail: heyanc@cq.edu.cn

of low-end mass customization products by using orthoses as examples, which eliminates the plaster cast making process. An example of the screw-pin reconfigurable mould applied in other field is shown in Fig.2 [19, 20, 21, 22, 23], in which mesh screw pins were adjusted in vertical direction to represent the component geometry and were then machined to make the contact surface of screw-pins be conformable for the vacuum forming purpose.

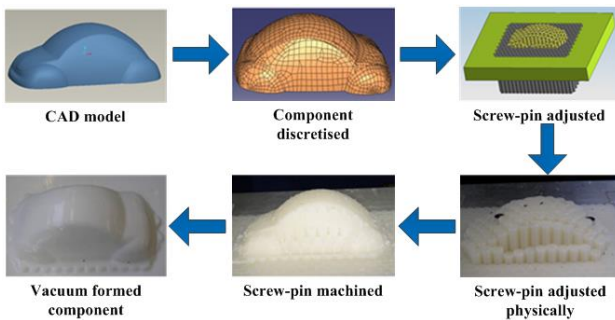


FIGURE 2 Concept of reconfigurable moulds using screw pins

The purpose of this paper is to propose a novel method of orthoses manufacturing based on 3D digital scanning, CAD/CAM/RE and the screw pin reconfigurable moulding technology. The whole process of orthoses manufacturing using reconfigurable moulds method is clean, fast and convenient, and environmental problems caused by plaster cast are minimized to realize the digital manufacturing.

2 The Research Framework

In order to reduce man-machine interaction and expertise demanding, a digital manufacturing system for orthoses is developed based on PowerSHAPE, one of the famous reverse engineering software. The system is comprised of four parts: 3D digital scanning and fast reconstruction, generation and machining of reconfigurable moulds, vacuum forming of orthoses, and intelligent database of human anatomy. The overall framework of the methodology is depicted in Fig. 3 and each part will be explained in detail in this paper.

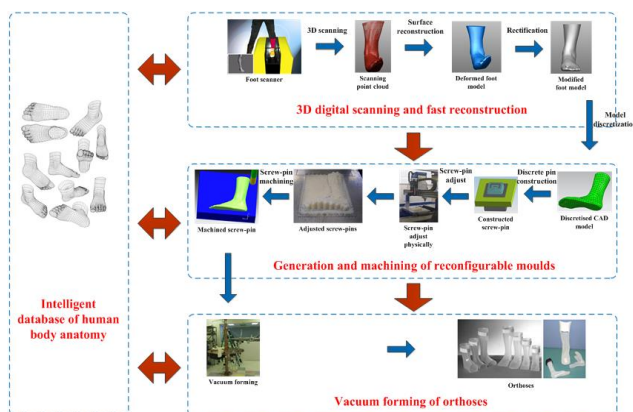


FIGURE 3 Research framework

2.1 3D DIGITAL SCANNING AND FAST RECONSTRUCTION

A specialized 3D scanner for the human anatomy, which can be available on the market, is used to scan patient's body region to be rectified by using orthopedics, and the contour is recorded by the 3D image in form of scanning point cloud. The scanned data is then processed to reconstruct the human anatomy CAD model after point cloud pre-processing within the reverse engineering environment. The reconstruction of the CAD model is a complex process, which involves much man-machine interaction and expertise. Since the system is operated by prosthetists/orthotists/technicians who has only medical or technical background, a friendly user interface for the rapid reconstruction of the human anatomy is provided. After that, the CAD model of human body anatomy should be rectified before generation of reconfigurable moulds, which is achieved by manual adding or removing material on the plaster in traditional manufacturing of orthoses. In this system, the obtained 3D geometric model, which is used as the positive cast is rectified through the developed tools in the principle, which is the same as that of the conventional manual method. Optimized scanning parameters for the generation of high quality point cloud are also embedded into the system, which is important for the reconstruction and rectification.

2.2 GENERATION AND MACHINING OF RECONFIGURABLE MOULDS

After the rectified CAD model of human body anatomy is done, it will be used to generate the corresponding reconfigurable moulds using screw-pin. Generation and machining of reconfigurable moulds consist of three steps: discrete screw-pin construction, adjustment of screw pins and machining of the adjusted screw-pins.

Since the size and the number of screw-pins have an influence on the process of pin adjustment, the tooling time, and representing accuracy of the human anatomy model, they should be set and optimized according to the discretized CAD model. The discrete screw-pin matrix is then constructed to be adjusted based on the input parameters. The amount of adjustment from the existed position of each screw-pin to the position representing the contour surface of human anatomy model is calculated firstly. In order to generate adjusting every screw-pin into certain position, the amount of adjustment is converted into the number of required adjusting turns of every screw-pin in constructed screw-pin matrix according to the screw-pins' parameters (diameter and pitch, etc.). The position and the required adjusting turns are recorded in G code [19, 20, 21, 22, 23], which is same as the NC code to control the cutter. The obtained G codes are input into the CNC system in the machine tool, which uses the screwdriver as the ordinary cutter to adjust every screw-pin into required position. The adjusted screw-pin matrix is representation for the rectified human anatomy model.

Due to steps existed between the screw-pins in the adjusted screw-pin matrix, screw-pins are required to be machined so that the contact surface of screw-pins is conformable with that of the part. In NC programming, the machining parameters are similar for similar geometry, such as human body anatomy model, which provides convenience for the exploration of optimized machining parameters. The optimized machining parameters such as size of the cutter, cutting speed, cutting depth, feed rate, type of tool path etc. for each of the human anatomy are explored and embedded in the system. Moreover, it allows minor correction for the users to generate NC codes rapidly. The screw-pins are machined in the CNC machining tool after input the generated NC codes.

2.3 VACUUM FORMING OF ORTHOSES

The reconfigurable moulds representing the CAD model of human body anatomy can be used as the positive plaster in conventional method for vacuum forming of thermoplastic material to manufacture orthoses. Vacuum forming parameters (heating temperature, time, vacuum pressure setting, etc.) are related with the shape and size of human body anatomy, matrix materials and thickness of the material. The optimal material properties and vacuum forming parameters are identified through Finite Element Analysis and experiments. After that, the orthosis is cut from the formed thermoplastic material. And then it is trimmed, polished and attached accessories for trying on.

2.4 INTELLIGENT DATABASE OF HUMAN BODY ANATOMY

As illustrated in Fig.3, the fabrication of orthoses can be completed through the above three parts. However, this system is developed for prosthetists/orthotists/technicians who has only medical or technical background. Moreover, the main purpose of the system is to render the manufacturing process of orthoses as simple and automated as possible. Therefore, the intelligent database is required.

In the intelligent database, many samples of human body anatomy are stored according to the gender, height and weight, and each of these is associated with the optimal parameters for digital scanning, reconstruction and rectification of the human anatomy, generation and machining reconfigurable moulds and vacuum forming.

After input the patient individual information and symptom, the most similar samples will be matched and displayed in the system associated with the optimal parameters as reference, which greatly simplify the process and reduce the manual repeatable work.

By using this intelligent database, the 3D digital scanning to reconstruct human body anatomy can be replaced by manual measurement at clinics where there is no scanner available. Based on the measurement data of

cross section of the human body anatomy, the similar model stored in the database is searched out and then modified to the personalized shape which can represent patient's human body anatomy for fabrication of orthoses. Particularly, the intelligent database is open, and when more people use the system, more samples will be stored, and a better chance that a match will be found. So, the operation of the system becomes more and more simple and intelligent and the quality of the manufactured orthoses improves over time. The whole process of orthoses manufacturing using reconfigurable moulds method is clean, fast convenient and without demanding expertise, and environmental problems caused by plaster cast are minimized to realize the digital manufacturing.

3 Case study

In this part, the valgus was taken to as a case study, as shown in Fig. 4. The point cloud of the deformed crus model was obtained through the scanning of a foot scanner and saved as the common file format (*.asc), and then input into the system which is developed based on PowerSHAPE. The crus 3D CAD model is reconstructed rapidly only through inputting several simple parameters and rectified later. After that, the corresponding reconfigurable mould is generated and machined to be the representation of the rectified model. It is used for vacuum forming of orthosis. The orthosis will be gained and to be tried on by the patient. The 3D digital scanning and fast reconstruction, and the generation and machining of reconfigurable moulds have been done. The vacuum forming of orthoses and the intelligent database of human anatomy are being researched. They will be finished soon.

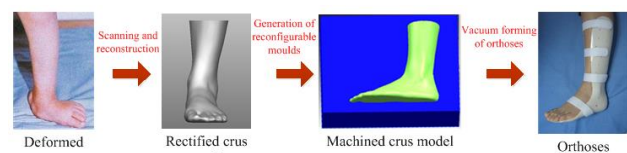


FIGURE 4 The digital manufacturing of orthoses

4 Conclusions

Over a billion people are estimated to live with some form of disability according to the first official global report on disability, which corresponds to about 15% of the world's population. The demand for orthotic devices is growing increasingly. However, the conventional fabrication technique of orthoses plays a dominant role in orthotics industry and the problems associated with which include lack of flexibility, long lead time, low-efficiency and material waste are still existed. An innovative method is proposed in this paper to replace traditionally dedicated moulds with reconfigurable moulds utilising screw-pins that is directly transferred to the vacuum forming of thermoplastic material at low cost for the digital manufacturing of orthosis, and an associated support system is developed. The whole

process of orthoses manufacturing using this method is clean, fast convenient and without much man-machine interaction and demanding expertise. The cost and cycle of orthoses manufacturing and environmental pollution are also minimized to realize the digital manufacturing.

References

- [1] People United States Census Bureau report: Global population aging faster <http://nmpmc.people.com.cn/GB/9753153.html>, September 2009
- [2] World Health Organization Disability and health <http://www.who.int/mediacentre/factsheets/fs352/en/> November 2012
- [3] Jerome C, Remi K, Christophe G 2005 Orthoses for mild scoliosis: a prospective study comparing traditional plaster mold manufacturing with fast, noncontact, 3-dimensional acquisition *Clinical Case Series* 30(4) 339-405
- [4] Lochner S J, Huissoon J P, Bedi S S 2012 Parametric design of custom foot orthotic model *Computer-Aided Design & Applications* 9(1) 1-11
- [5] Jones D 1998 Impact of advanced manufacturing technology on prosthetic and orthotic practice *J. Biomed. Eng* 10 179-83
- [6] Weinberg B, Nikitczuk J, Paritti B 2007 Design, control and human testing of an active knee rehabilitation orthotic device *IEEE International Conference on Robotics and Automation, Roma, Italy, April 2007*, pp. 4126-33
- [7] Milusheva S M, Tosheva E Y, Toshev Y E 2007 Ankle foot orthosis with exchangeable elastic elements *Series on Biomechanics* 23(1) 91-5
- [8] Sinha A N, Udai A D 2007 Development of artificial limb surface using reverse engineering method *Proceedings of International Conference on Advances in Machine Design & Industry Automation, Pune, India, January 2007* 519-22
- [9] Condie D N 2008 The modern era of orthotics *Prosthetics and orthotics international* 32(3) 313-23
- [10] Paul S K, Sivarasu S, Mathew L 2012 Customized foot orthosis development by 3D reconstruction of the CT images *Engineering* 4 692-5
- [11] Telfer S, Gibson K S, Hennessy K, Steultjens M P, Woodburn J 2012 Computer-Aided Design of customized foot orthoses: reproducibility and effect of method used to obtain foot shape *Archives of Physical Medicine and Rehabilitation* 93(5) 863-70
- [12] Jumani S, Memon Z I, Larik J 2013 Process Modelling of Rapid Manufacturing Based Mass Customisation System for Fabrication of Custom Foot Orthoses: Review Paper *Mehran University Research Journal of Engineering & Technology* 32(2) 183-96
- [13] Pinson G T 1980 Apparatus for Forming Sheet Metal *U.S. Patent 4212188 July 1980*
- [14] Fuchs A, Radolfzellerstr L L, Konstanz D E 1982 Method and Arrangement for Producing a Curved Sail *U.S. Patent 4309824*
- [15] Kommineni P, Hollandsworth P E, Jones J W 1988 Method of shaping an antenna panel *U.S. Patent 4731144 March 1988*
- [16] Sherrill D E, Young K G 2001 Apparatus for constructing a composite structure *U.S. Patent 6298896 B1 October 2001*
- [17] Liu C G, Li M Z, Sui Z 2008 Multi-point forming technology applied in the manufacturing of aircraft sheet parts *Journal of Plasticity Engineering* 15(2) 109-14
- [18] Sky control net *Flexible and reconfigurable tooling for manufacturing aircraft panels* <http://www.skycontrol.net/industry/dataform-flexible-and-reconfigurable-tooling-for-manufacturing-aircraft-panels> March 2011
- [19] Wang Y, Wang Z, Gindy N 2010 Automated discrete-pin adjustment for reconfigurable moulding machine *International Journal of Computer Integrated Manufacturing* 23(3) 229-36
- [20] Wang Z, Wang Y, Gindy N 2010 Planning and control of a hybrid vacuum-forming system based on screw-pin tooling *WSEAS Transactions on Systems and Control* 5(7) 557-66
- [21] Wang Z, Wang Y, Gindy N 2008 Design and construction of reconfigurable screw-pin tooling for vacuum forming system *Proceeding of the 18th international conference on flexible automation and intelligent manufacturing (FAIM 2008), June to - July, Sweden, 2008* 4027-31
- [22] Wang Z 2009 *Rapid Manufacturing of Vacuum Forming Components Utilising Reconfigurable Screw-Pin Tooling* PhD thesis, The University of Nottingham
- [23] Wang Y, Wang Z J, Gindy N 2011 A method for representation of component geometry using discrete pin for reconfigurable moulds *Advances in Engineering Software* 42(7) 409-18

Acknowledgements

The authors would like to thank the support from the National Natural Science Foundation of China (Grant No. 51105212), Ningbo International Collaboration Funds (Grant ID: 2009D10016 and 2011D10003) and the Ningbo Natural Science Foundation (Grant ID: No.2011A610152).

Authors



Yuewei Ai, born in September 15, 1989, Shaoyang, Hunan

Current position, grades: Master of Mechanical Manufacturing and Automation, postgraduate in Chongqing University

University studies: Mechanical Manufacturing and Automation in Chongqing University

Scientific interest: Green Manufacturing and Reversing Engineering in Mechanical Engineering

Publications: 1 paper

Experience: 2007-2011: Studying for Bachelor's Degree of Mechanical Design and Manufacturing; 2011-: Studying for Master's Degree of Mechanical Manufacturing and Automation, as the main researcher undertaking the National Natural Science Foundation of China—A method of green design and manufacturing of orthoses based on reconfigurable moulds utilizing the screw-pin.



Yan He, born in January 30, 1981, Chongqing, China


Current position, grades: Doctor of Mechanical Engineering, associate professor and doctoral supervisor in Chongqing University

University studies: Mechanical Engineering in Chongqing University

Scientific interest: Green Manufacturing, High Efficiency Low Carbon Manufacturing Equipment and System

Publications: 6 Patents, 20 Papers

Experience: Undertaking the National Natural Science Foundation of China, the 12th Five-year Plan 863 Project, British Engineering and Natural Sciences Research Council (ERSRC) Project, Doctoral Fund Project and other state-level scientific research projects, published more than 20 academic theses in "Journal of Engineering Manufacture", "Journal of Advanced Mechanical Design, Systems, and Manufacturing", "Journal of Cleaner Production" and "Chinese Journal of Mechanical Engineering" and other SCI journals, serving as the deputy director of Green Manufacturing Technology Research Institute of Chongqing University, the member of SINO-UK Low Carbon Manufacturing Consortium and other academic positions and the reviewer of many SCI journals.

	<p>Zhijian Wang, born in June 27, 1973, Kunshan, Jiansu</p> <p>Current position, grades: Doctor of Mechanical Engineering, associate professor and Master supervisor in Ningbo Institute of Material Technology & Engineering, CAS, Ningbo, China</p> <p>University studies: Mechanical Engineering in the University of Nottingham, UK</p> <p>Scientific interest: Low Carbon Manufacturing Technology and Equipment</p> <p>Publications: 1 Patent, 20 papers</p> <p>Experience: As the main member complete 2 Sino-British science and technology innovation plan projects, 2 UK EPSRC projects, 2 British Rolls-Royce projects, 1 item of 863 / CIMS project of Guangdong province</p>
	<p>Yang Wang, born in September 8, 1973, Yiyang, Hunan</p> <p>Current position, grades: Doctor of Mechanical Engineering, associate professor and Doctoral supervisor in the University of Nottingham, Ningbo, China</p> <p>University studies: Mechanical Engineering in the University of Nottingham, UK</p> <p>Scientific interest: Green Manufacturing, Reconfigurable moulds</p> <p>Publications: 6 Patents, 20 papers</p> <p>Experience: Undertaking the project-The development of reconfigurable fixture and commercialization supported by the science and technology innovation plan China and Britain, as the reviewer of International Journal of Advanced Manufacturing Technology, Computer-Aided Design and so on.</p>

Evolutionary game and simulation of organizational information security investment

Wei Sun¹, Yang Yu^{2*}

¹ Business School, Liaoning University, Shenyang, 110136, P.R.China

² Institute of Systems Engineering, Northeastern University, Shenyang, 110819, P.R.China

Received 1 March 2014, www.tsi.lv

Abstract

To investigate the evolution law of organizational information security investment, this paper analyses evolutionary stable strategies of organizational information security investments using evolutionary game theory and verifies the evolutionary stable strategies through the simulation based on Repast, a multi-agent simulation platform. First, according to the bounded rationality of actual organizations, this paper sets up the evolutionary game model of organizational information security investment. And then, we investigate the evolutionary stable strategies by replicator dynamics. Finally, we simulate the evolutionary game by Repast based on Java programming language, and the experimental results verify the evolutionary stable strategies obtained from the theoretical analysis. The research results can be used to predict the long-term stable trend of organizational information security investment, state that investment cost is the key for organizations to choose the strategy, and provide decision support for organizational information security investment.

Keywords: evolutionary game, information security investment, multi-agent simulation, evolutionary stable strategy

1 Introduction

Information security investment is an inevitable investment decision of organization in digital economy. As the result of global information security survey reveals [1], over half of enterprises spent more money on information security than before, but they do not feel safer. However, information security investment has the characteristics of concealment and confrontation and is different from other investments, which get effect instantly. In addition, the effect of information security investment is inconspicuous. Therefore, ordinary research method is not suitable for the study of information security investment.

The research of information security concentrates on technology field [2-3], however, the management study of information security is more important than technology study to some extent. In fact, the abstract characteristic of information security is the strategy interdependence, while strategy interdependence is just the focus of Game Theory. Thus, we can explore the information security investment by game theory. Samuel N.H. pointed out directly that applying game theory to information security study is a promising research direction.

Due to strategy dependency of inter-organizational information security investment, game theory can be used to research information security investment. Researches on information security game theory are currently still in the exploratory stage, only a few studies [3-7] are based

on the assumption of game party complete rationality, which is a great difference with game party bounded rationality in the real world.

Evolutionary game theory is the scientific method to study game party bounded rationality. Evolutionary game theory is based on Darwin's evolution theory and Lamarck's genetic theory, and takes game party bounded rationality as the study object. Therefore, Evolutionary game theory enhances the game analysis of scientific and credibility. John Nash's "herd behaviour explains" [8] can be thought as the earliest evolutionary game theory; Weibull summarized the evolutionary game theory [9] to form the systematic evolutionary game theory. Evolutionary game theory is a new field of game theory, but there are many gaps in the applications of economics, biology and other social science [10].

Considering information security investment entities in the real world have bounded rationality, the paper uses evolutionary game theory to analyse information security investments and simulates the theoretical result through the Repast, i.e. a multi-agent simulation platform.

The organization of the paper is as following. Section 2 establishes information security investments evolutionary game model. In section 3, we investigate the evolutionary stable strategy of information security investment by replicator dynamics. Section 4 simulates the evolutionary game by Repast simulation platform and verifies the evolutionary stable strategy of the third section by the experimental results. In addition, the results

* Corresponding author e-mail: yuyang@ise.neu.edu.cn.

predict the long-term stability trends of information security investment and state that the investment cost is the key to organizational strategic choice. In the last section conclusions are given.

TABLE 1 Evolutionary game model of information security investment of two organizations

Organization 1	Organization 2	
	Investment	No investment
Investment	$E-C+I, E-C+I$	$E-C-qL+I, E-pL$
No investment	$E-pL, E-C-qL+I$	$E-pL-(1-p)qL, E-pL-(1-p)qL$

Game model, including the proceeds of a comprehensive information security function value of the investment benefits, not only to consider the prevention of information security incidents resulting direct value benefits, but also to consider the organization of information security investment brings enhance brand value, increased goodwill and other organizations the indirect value of intangible benefits.

2 Evolutionary game model of information security investment

In real world, information security investment entities have bounded rationality, and so investment decision often depends on instinctive intuition, or the simple imitation of the successful strategies. That means the strategic choice is rarely decided according to precise mathematical analysis, and thus rational level in investors decision-making is low. In view of this situation, this paper establishes information security investment evolutionary game model considering game party bounded rationality.

In the game model, the two players are assumed the same type of risk-neutral organization. The parameters in the model are defined as follows. E is the normal income of organization, which is independent of the cost-effectiveness of information security investment. C stands for costs of organization investing in information security. Since the information security investment can impact the intangible asset of the organization, we choose I to reflect the intangible asset from the investment of information security. L is the loss caused by information security failure. The disaster probability caused by the organization itself is p , and the disaster probability caused by other organizations is q , which represents negative externalities among the organizations.

Information security investments evolutionary game model is established with benefit matrix. The benefits of two organizations are $E-C+I$, when they are investing in information security. When one organization invests in information security but another organization doesn't invest, the benefit of organization investing is $E-C-qL+I$ because negative externalities cause the loss qL , and the benefit of organization not investing is $E-pL$ because its own information security incidents cause the loss pL .

When both two organizations are not investing in information security, each organization has its own security risks caused by p and negative externalities caused by other organizational impact q , and so the benefit of two organization are $E-pL-(1-p)qL$. Therefore, we establish information security investment evolutionary game model, as shown in Table 1.

3 Replicator dynamics of evolutionary game of information security investment

Replicator dynamics is the basic dynamics of evolutionary game theory, and can reasonably characterize the stability of bounded rationality game party in dynamic process of learning and adjustment, predict changing trends of game party groups [11]. Compared to other dynamic equations, replicator dynamics have a broader application prospects. The decision-making body of information security investment is less rational, and so it is suitable to investigate information security investments evolutionary game using replicator dynamics.

Replicator dynamics is the most valuable function in evolutionary game theory for application. Replicator dynamics can imitate the studying and regulating processes of the players with bounded rationality well, forecast the change trend of the colony behaviour, and provide the scientific reference for macro control. Compared with other dynamics function, replicator dynamics is the most widely used dynamics in evolutionary game theory.

In the analysis of Replicator Dynamics of information security evolutionary game, suppose x be the proportion of the organizations with the strategy of investing in information security, and the proportion of no investment strategy is $1-x$. Let $\frac{dx}{dt}$ represent the change ratio of the proportion of investment strategy. Then, replicator dynamics of information security evolutionary game is expressed:

$$\frac{dx}{dt} = x(u_i - \bar{u}). \tag{1}$$

In the above information security evolutionary game, u_i is the expectation benefit of game party adopting investment information security strategy:

$$u_i = x(E - C + I) + (1 - x)(E - C - qL + I). \tag{2}$$

u_n is the expectation benefit of game party not adopting investment information security strategy:

$$u_n = x(E - pL) + (1 - x)[E - pL - (1 - p)qL]. \tag{3}$$

Furthermore, the overall average expected benefit \bar{u} can be calculated:

$$\bar{u} = xu_i + (1-x)u_n. \tag{4}$$

Based on equations (2)–(4), we can obtain replicator dynamics of information security investment evolutionary game:

$$\frac{dx}{dt} = x(1-x)[pqLx + (pL + I - C - pqL)]. \tag{5}$$

Let replicator dynamics of information security investment evolutionary game equals to zero, and we can get three stable states as follows by solving equation (5):

$$x_a^* = \frac{C + pqL - pL - I}{pqL}, x_b^* = 0 \text{ and } x_c^* = 1. \tag{6}$$

According to the value of $x_a^* = \frac{C + pqL - pL - I}{pqL}$, we

investigate three evolutionary stable strategies of information security investment evolutionary game respectively and verify them based on the Repast simulation platform.

Evolutionary stable strategy (ESS) can be analysed on the basis of replicator dynamics. ESS is the equilibrium with the true stability and strong forecast ability, and can be reached by the regulation of the player. Also, ESS can resist the disturbance of wrong departure. Replicator dynamics and ESS are the two basic concepts in evolutionary game theory. The stable condition is the condition that the proportion of investment strategy and the proportion of no investment strategy are fixed.

4 Evolutionary stable strategy and simulation of evolutionary game of information security

In this paper, we use Repast simulation tools, JBuilder platform to verify the evolutionary stable strategy. Repast (Recursive Porous Agent Simulation Toolkit) is a multi-agent modelling and simulation toolset, which developed by the University of Chicago and Argonne National Laboratory [12]. Repast platform can control and reproduce simulation experiment, and the output results of the simulation experiment can be used to analyse the game's behaviour and trends well. Repast includes simulation scheduling Class, agent Class and space Class, and a series of input and output Class library, to provide the corresponding programming framework for the realization of the simulation model. Repast simulation modelling process and corresponding functions is shown in Figure 1.

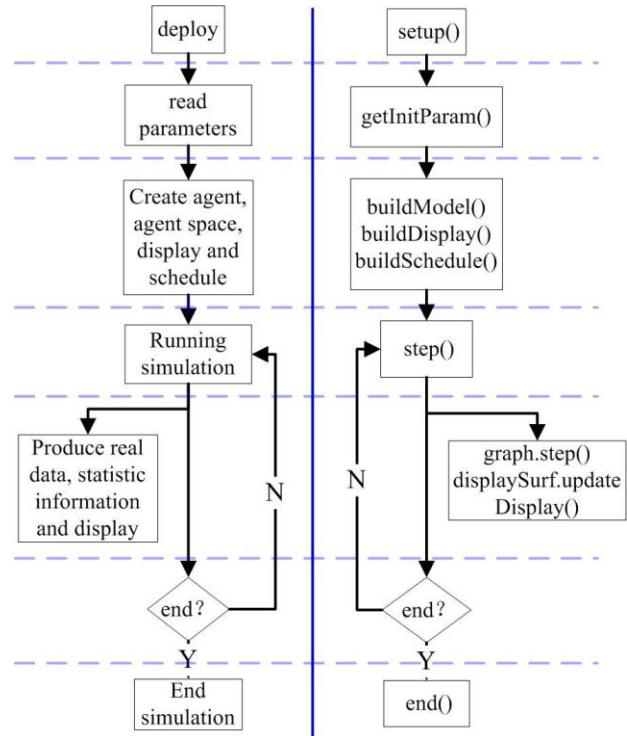


FIGURE 1 Repast simulation process and corresponding functions

4.1 EVOLUTIONARY STABLE STRATEGY AND SIMULATION WITH $x_a^* \leq 0$

Evolutionary stable strategy (ESS) is real equilibrium stability in evolutionary game theory, and can be used to predict long-term stable trend of game party group behaviour. $x_a^* \leq 0$ means $C \leq pL + I - pqL$, in addition x ranges from 0 to 1, so the replicator dynamics of information security investments evolutionary game have only two stable states, i.e. $x_b^* = 0$ and $x_c^* = 1$. In the two stable states, the evolutionary stable strategy equilibrium must be able to withstand a slight deviation from the disturbance. The explanation of mathematical analysis is as follows: when the disturbance makes x less than evolutionary stable strategy, $\frac{dx}{dt}$ must be greater than 0; when the disturbance makes x greater than evolutionary stable strategy, $\frac{dx}{dt}$ must be less than 0. In the phase diagram of replicator dynamics equation, the one with negative tangent slope in intersection points of replicator dynamics curve intersects and the horizontal axis is evolutionary stable strategy of evolutionary game replicator dynamics. Figure 2 shows the phase diagram of replicator dynamics equation of information security investment evolutionary game with $x_a^* \leq 0$. In Figure2, we can easily observe that when $x_a^* \leq 0$, $x_c^* = 1$ is the evolutionary stable strategy of the information security evolutionary game.

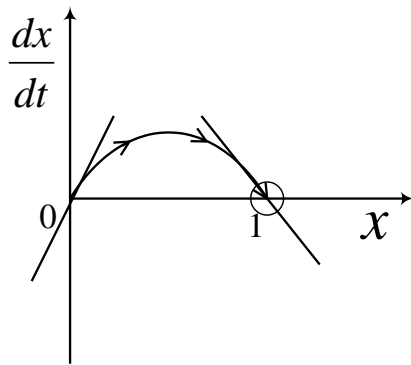


FIGURE 2 The phase diagram of replicator dynamics equation with $x_a^* \leq 0$

We use Java-based Repast simulation platform to verify the evolutionary stable strategy of information security investment with $x_a^* \leq 0$. The detailed simulation input parameters are shown in Table 2.

TABLE 2 The detailed simulation input parameters with $x_a^* \leq 0$

parameters	<i>E</i>	<i>I</i>	<i>L</i>	<i>p</i>	<i>q</i>	<i>C</i>	<i>x</i>
value	10	6	9	0.6	0.4	5	0.1

The overall strategy selection of the simulation initial stage is shown in Figure 3, where an agent represents an organization. The red means the investment strategy which the organization selected is information security, while the blue represents to not invest in information security. In initially stage, only 10% of organizations choose to invest in information security. With the simulation running of the evolution mechanism, the proportion of organization adopting investment in information security gradually raises. Figure 4 shows the phase diagram at which the proportion of organization adopting investment in information security is 50%. Figure 5 shows the final evolution result of the simulation, where all organizations choose to invest in information security.

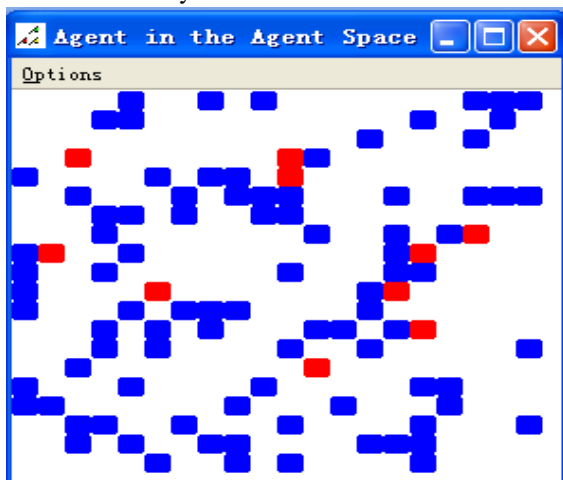


FIGURE 3 The overall strategy selection of the simulation initial stage with $x_a^* \leq 0$

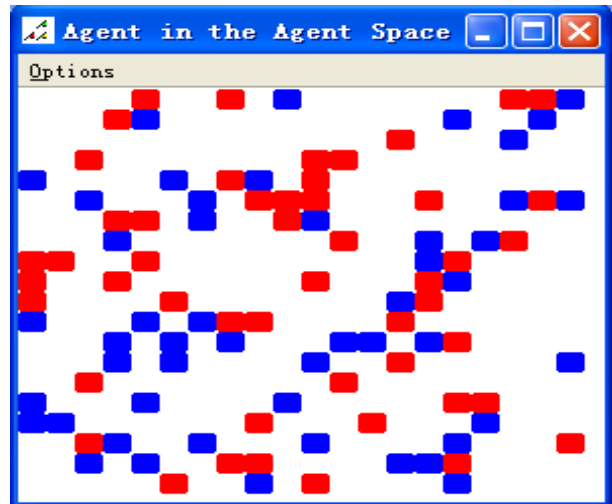


FIGURE 4 The overall strategy selection of the simulation phase diagram $x_a^* \leq 0$

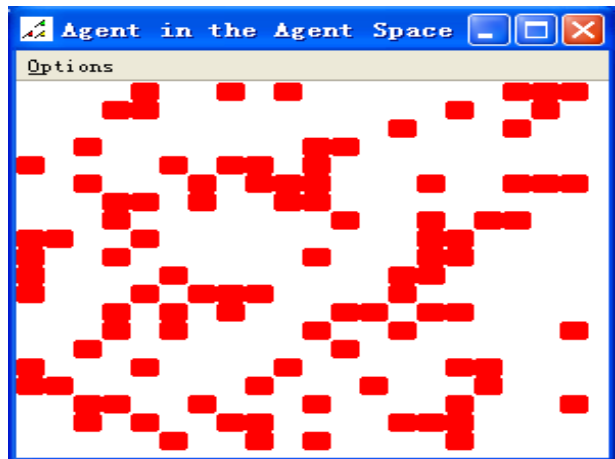


FIGURE 5 The overall strategy selection of stable trend with $x_a^* \leq 0$

Figure 6 shows the trend of the proportion of information security investment increasing from 10% to 50%. Figure 7 shows the whole trend, i.e., the proportion of information security investment increasing from 10% to 100%.

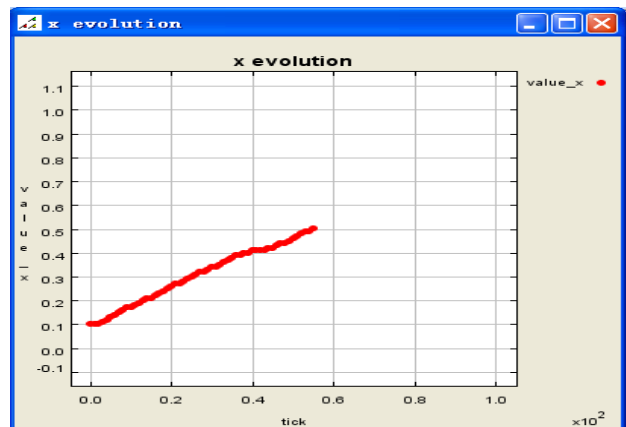


FIGURE 6 Trend of the proportion increasing from 10% to 50% under information security investment with $x_a^* \leq 0$

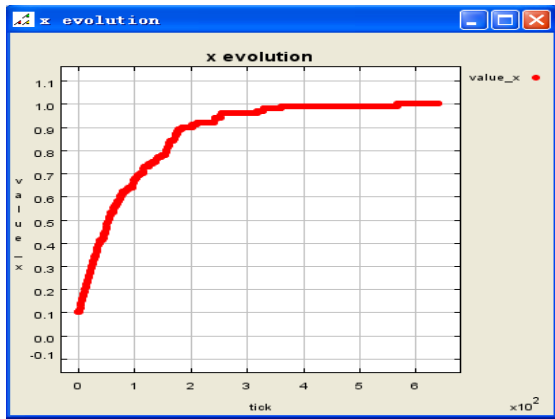


FIGURE 7 Whole trend of the proportion increasing from 10% to 100% under information security investment with $x_a^* \leq 0$

Therefore, the evolutionary game simulation results verify the replication dynamics evolutionary stable strategy with $x_a^* \leq 0$, i.e. when x is closing to 1, all organizations choose to invest in information security through long-term evolution. This evolutionary result interprets the actual situation well. $x_a^* \leq 0$ means $C \leq pL + I - pqL$, and so the cost of information security investment is lower, which prompts all organizations invest in information security. For the organizations choosing non-investment strategy, they will mimic the strategy of the successful enterprise to invest in information security when they find the benefit of information security investment, i.e. the expected revenue of investment is higher than that of no investment. Therefore, finally all organizations will adopt the strategy of investing information security.

4.2 EVOLUTIONARY STABLE STRATEGY AND SIMULATION WITH $x_a^* \geq 1$

$x_a^* \geq 1$ means $C \geq pL + I$, in addition x ranges from 0 to 1, so at the same time the replicator dynamics of information security investment evolutionary game also have only two stable states, i.e. $x_b^* = 0$ and $x_c^* = 1$.

Figure 8 is the phase diagram of replicator dynamics equation of information security investment evolutionary game with $x_a^* \geq 1$. From Figure 8, we can observe that $x_b^* = 0$ is the evolutionary stable strategy.

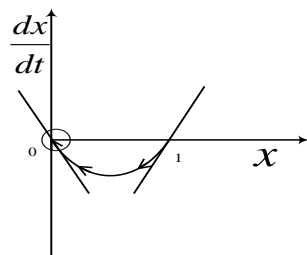


FIGURE 8 The phase diagram of replicator dynamics equation with $x_a^* \geq 1$

The evolutionary stable strategy of information security investment with $x_a^* \geq 1$ will be verified by simulation. The detailed simulation input parameters are shown in Table 3.

TABLE 3 The detailed simulation input parameters with $x_a^* \geq 1$

E	I	L	p	q	C	x
10	6	9	0.6	0.4	15	0.9

The simulation process is similar to Figures 4-6, and Figure 9 shows the completely evolutionary trend of the proportion of information security investment with $x_a^* \geq 1$ increasing from 90% to 0.

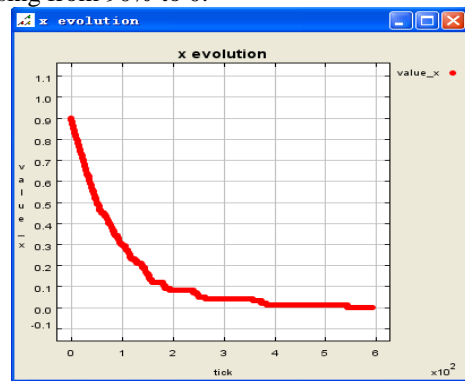


FIGURE 9 Whole evolutionary trend of the proportion decreasing from 90% to 0 under information security investment with $x_a^* \geq 1$

The simulation result proves that when $x_a^* \geq 1$, the evolutionary stable strategy is $x_b^* = 0$. The evolutionary result is consistent with the actual situation. $x_a^* \geq 1$ means $C \geq pL + I$, and so the cost of information security investment is higher, therefore, organizations will eventually not invest in information security in the long evolution process. To avoid this dangerous situation, information security technology must be improved to effectively reduce the cost of information security investment.

4.3 EVOLUTIONARY STABLE STRATEGY AND SIMULATION WITH $0 < x_a^* < 1$

$0 < x_a^* < 1$ means $pL + I - pqL < C < pL + I$. At this time, replicator dynamics includes three stable strategies, i.e. $x_a^* = \frac{C + pqL - pL - I}{pqL}$, $x_b^* = 0$, and $x_c^* = 1$. According to this phase diagram of replicator dynamics equation with $0 < x_a^* < 1$, as shown in Figure 10, we can get the evolutionary stable strategies under $0 < x_a^* < 1$ are $x_b^* = 0$ and $x_c^* = 1$.

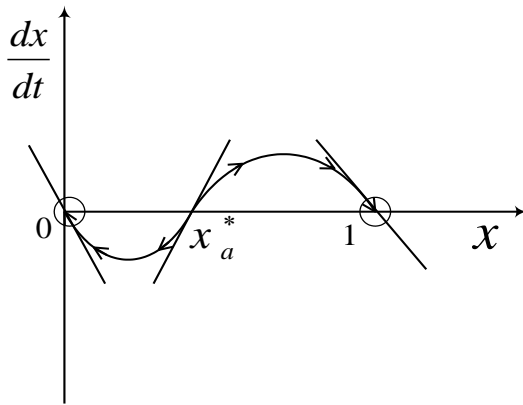


FIGURE 10 The phase diagram of replicator dynamics equation with $0 < x_a^* < 1$

According to the phase diagram of replicator dynamics, when the proportion of initial adoption information security investment (x) falls in interval $[0, x_a^*]$, the replicator dynamics will tend to evolutionary stable strategy of $x_b^* = 0$; when x falls in interval $[x_a^*, 1]$, the replicator dynamics will tend to be evolutionary stable strategy of $x_c^* = 1$.

The evolutionary stable strategy of information security investment with $0 < x_a^* < 1$ will be verified by simulation. For the situations with x in intervals $[0, x_a^*]$ and $[x_a^*, 1]$, the detailed simulation input parameters are shown in Table 4 and Table 5, respectively.

TABLE 4 The detailed simulation input parameters with $0 < x_a^* < 1$ and $0 < x < x_a^*$

E	I	L	p	q	C	x
10	6	9	0.6	0.4	10	0.3

TABLE 5 The detailed simulation input parameters with $0 < x_a^* < 1$ and $x_a^* < x < 1$

E	I	L	p	q	C	x
10	6	9	0.6	0.4	10	0.4

Figure 11 shows the whole evolutionary trend of the proportion decreasing from 30% to 0 under information security investment with $0 < x_a^* < 1$ and $0 < x < x_a^*$. Figure 12 shows the whole evolutionary trend of the proportion decreasing from 40% to 100% under information security investment with $0 < x_a^* < 1$ and $x_a^* < x < 1$.

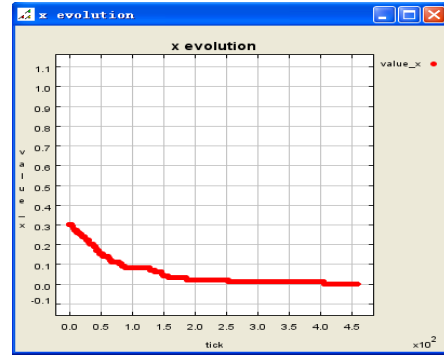


FIGURE 11 Whole evolutionary trend of the proportion decreasing from 30% to 0 under information security investment with $0 < x_a^* < 1$ and $0 < x < x_a^*$

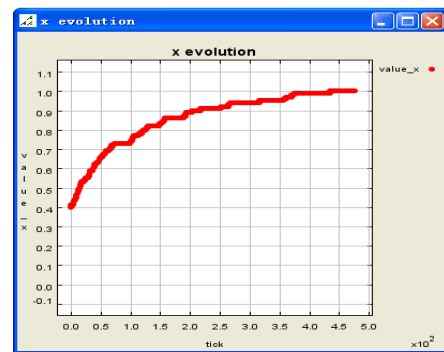


FIGURE 12 The whole evolutionary trend of the proportion decreasing from 40% to 100% under information security investment with $0 < x_a^* < 1$ and $x_a^* < x < 1$

The simulation results verify the evolutionary stable strategies under $0 < x_a^* < 1$, and explains the actual situation well. In practice, only when the proportional of organizations investing in information security is higher than x_a^* , the negative externality among organizations is relatively small. At this time, the organizations that have invested in information security will produce a positive demonstration effect in other organizations, which causes the number of organizations that will invest in information security increases. To achieve this ideal equilibrium result, x_a^* should close to the origin of the coordinate axes as possible, that is to say the information security investment cost C should close to $pL + I - pqL$ when $pL + I - pqL < C < pL + I$. Therefore, effectively reducing the cost of information security is the key to promote the organizations to invest in information security.

5 Conclusion

This paper analyses organizational information security investments with Evolutionary Game Theory and simulates the evolutionary game through Repast. Considering the bounded rationality of organizations in reality, we set up the evolutionary game model of information security investment. We investigate the evolutionary stable strategy by replicator dynamics and simulate the evolutionary game by Repast based on Java,

and the experimental results verify the evolutionary stable strategy obtained from the theoretical analysis. The research result reveals that investment cost is the key for organizations to choose the strategy, predicts the long-term stable trend of organizational information security investment, and provides good direction for organizational information security investment.

References

[1] Qiang T 2007 The global survey of information security *Inform. Weekly*, 5

[2] Daniel M, Eduardo F M, Mario P 2007 A common criteria based security requirements engineering process for the development of secure information systems. *Comp. Stand. Inter.* 29(2) 244-53

[3] Ilan O, Julia K, Corey H 2007 Information security in networkable Windows-based operating system devices: Challenges and solutions. *Comput. Secur.* 26(2) 177-82

[4] Cavusoglu H, Mishra B, Raghunathan S 2004 A model for evaluating IT security investments. *Commun. ACM* 47(7) 87-92

[5] Varian H 2004 System reliability and free riding. Working paper, *Economics of Information Security* Univ. California: Berkeley

[6] Kunreuther H, McNulty P, Kang Y 2002 Improving environmental safety through third party inspection *Risk Anal.* 22(2) 309-18

[7] Kunreuther H, Heal G 2003 Interdependent security *J. Risk Uncertainty* 26(2-3) 231-49

[8] Nash J 1950 *Non-cooperative Games*. Princeton Univ: New Jersey

[9] Weibull J 1995 *Evolutionary Game Theory*. MIT Press: Cambridge

[10] Xie S Y 2001 Evolutionary game theory under bounded rationality *J. Shanghai Univ. Financ. Econ.* 3(5) 3-9

[11] Taylor P, Jonker L 1978 Evolutionarily stable strategies and game dynamics *Math. Biosci.* 16 76-83

[12] Tobias R, Hofmann C 2004 Evaluation of free Java-libraries for social-scientific agent based simulation *J. Artif. Soc. & Soc. Simul.* 7(1) 15-28

Acknowledgments

This research is supported by China Postdoctoral Science Foundation under Grant 2012M510828.

Authors	
	<p>Wei Sun, born on January 1, 1980, Fushun, Liaoning, P.R.China</p> <p>Current position, grades: Liaoning University, Associate Professor University study: Dalian University of Technology (Scholl of Management Science and Engineering), 2004; Ph.D. degree in management science and engineering, Dalian University of Technology, Dalian, China, 2008 Scientific interest: Game Theory, Operations Research, Management Science</p>
	<p>Yang Yu, born on November 3, 1980, Dalian, Liaoning, P.R.China</p> <p>Current position, grades: Northeastern University, Ph.D Current position, grades: Dalian University of Technology (Scholl of Management Science and Engineering), 2005; Ph.D. degree in management science and engineering, Dalian University of Technology, Dalian, China, 2009 Scientific interest: Operations Research, Management Science</p>

The application of fuzzy association rules in the employment data mining of a higher vocational college

Laiquan Liu¹, Li Lei², Yanrui Lei^{1*}

¹ Hainan College of Software Technology, Fuhai Road No.128, Qionghai, Hainan, China

² Chongqing University of Arts and Sciences, Huachuang Road No.598, Yongchuan, Chongqing, China

Received 1 March 2014, www.tsi.lv

Abstract

Data mining is able to extract potentially useful information from plentiful seemingly unrelated data. A high efficiency is therefore obtained using these useful data in work or study. Association rules mining is a significant branch in data mining. It mirrors the implicit relations among transactions in mass data. In addition, association rules can intuitively reflect the associations among item sets in data, and the relations are established according to the frequencies of the item sets appearing in data. This method, which explains its rules clearly and is easily to understand, therefore is different from the traditional statistical method. This research introduced and applied the mining algorithms of fuzzy association rules to the employment data analysis of a higher vocational college, in order to find significant association rules from numerous data and provide guidance for the education and employment in the future, therefore improving the employment rate further.

Keywords: Association rules, Data mining, Research, Application

1 Introduction

Data mining is an effective method to solve the problem of data rich but information poor currently. By using this method, potentially useful information can be discovered in mass data. Moreover, the relevant predication and discovering etc. of neglected information can be carried out using the discovered information. Data mining therefore presents broad application prospects.

Association rules mining is a significant branch in data mining. It is to discover the potential associated information in mass data. It was first reported in the data mining process of customers' transaction records in shopping malls [1]. There are no causal relationships in the results of association rules mining, and the results cannot be described using these relationships. Association rules mining mirrors the implicit relations among transactions in mass data. In addition, association rules can intuitively reflect the associations among item sets in data, and the relations are established according to the frequencies of the item sets appearing in data. This method, which explains its rules clearly and is easily to understand, therefore is different from the traditional statistical method. Association rules can intuitively express the relations among item sets (different values of variables) in data. The relations are not based on certain distributions and obtained using repeatedly iteration fittings of data in certain models. However, they are established according to the probabilities of the item sets appearing in data.

2 The definition of association rules

Let $D=\{t_1, t_2, \dots, t_n\}$ be a transactional database, T be any transaction set with an unique mark in D , and $I=\{i_1, i_2, \dots, i_m\}$ be a set composed of different items of number m . Each transaction t_i ($i=1, 2, \dots, n$) corresponds to an unique subset in I [2].

Definition 1. Item: Any element i in the set of $I=\{i_1, i_2, \dots, i_m\}$ is defined as a item.

Definition 2. Item set: In association analysis, a set containing none or multi-item is an item set. If an item set contains items of number k , it is called a k -item set. For example, {notebook computer, printer} is a 2-itemset. An empty set is an item set does not contain any items. If an item set X is a subset of a transaction T , that means the transaction T includes the item set X , which is denoted as $X \subseteq T$.

Association rules are implication expressions in the form of $X \Rightarrow Y$, where $X \subseteq T$, $Y \subseteq T$, and $X \cap Y = \varnothing$. The antecedent and the consequent of association rules are X and Y , respectively. Association rules mining is designed to find the implications which meet the set minimum support and confidence in mass data.

The support is applied to express the percentage of a rule in all the transactions in a database, and is a criterion to measure the importance of an association rule as well. The larger the support, the more important the association rule is in the whole database. The confidence is a measurement to determine the accuracy of an association

*Corresponding author e-mail: leiyarui@139.com

rule. Generally, the association rules meeting the minimum support and confidence in the meantime are considered. If a rule is high support and low confidence, the rule presents low reliability; on the contrary, if a rule is low support and high confidence, the rule is seldom used.

Therefore, two threshold values, namely the minimum support and the minimum confidence, are set to guarantee the discovered association rules can be used in practice. These two values are employed to abandon the redundant and invalid rules.

Definition 3. Count. The number of the transactions containing item set X in the database D is called the count.

Definition 4. Support.: It is the ratio of the count of the item set X divided by the number of all the transactions in the database D. The support of item set X is denoted as $sup(X)$, and the support of $X \Rightarrow Y$ is denoted as $sup(X \Rightarrow Y)$:

$$sup(X \Rightarrow Y) = P(X \cap Y) \tag{1}$$

According to Definitions 3 and 4, in the case that the number of the transactions contained in the transaction database D is marked by |D|, the relation of the count and the support of item set X is expressed in the following formula:

$$count = sup \times D \tag{2}$$

Definition 5. Confidence. The confidence of $X \Rightarrow Y$ is the specific value of the number of the transactions containing item sets X and Y at the same time and the number of the transactions merely containing item set X in the database D. The confidence of $X \Rightarrow Y$ is denoted as $conf(X \Rightarrow Y)$:

$$conf(X \Rightarrow Y) = \frac{sup(X \cup Y)}{sup(X)} \times 100\% \tag{3}$$

Definition 6. Minimum support. According to the requirements, a threshold value is set which represents the minimum importance of the acceptable association rules, is denoted as minsup.

Definition 7. Minimum confidence. According to requirements, another threshold value is set. It displays the minimum confidence of the acceptable association rules and is denoted as minconf.

Definition 8. Strong association rule. The minsup and the minconf are set. If $sup(X \Rightarrow Y) \geq min\ sup$ and $conf(X \Rightarrow Y) \geq min\ conf$, the $X \Rightarrow Y$ is a strong association rule; otherwise, it is a weak one.

Association rules mining is aimed to discover all the strong association rules in D, the item sets corresponding to which must be frequent item sets. The process of association rules mining therefore begins with the discovering frequent item sets, then strong association rules are generated, and finally the rules are explained.

3 The process of association rules mining

Association rules mining is generally divided into and carried out as two subproblems [3].

3.1 DISCOVERING FREQUENT ITEMSETS

According to the minsup set by the decision maker, all the frequent item sets in the database D are found out. The frequent item sets refer to the item sets satisfying support as well as minsup. Since possibly there are inclusive relations among these frequent item sets, users need to find out the set of those frequent large item sets, which cannot be included in other frequent item sets. This is the basis of finally generating association rules as well.

3.2 DISCOVERING RULES

According to the set minconf, the rules with confidences not less than minconf were discovered in every maximum frequent item sets. The model of association rules mining is shown in Fig. 1.

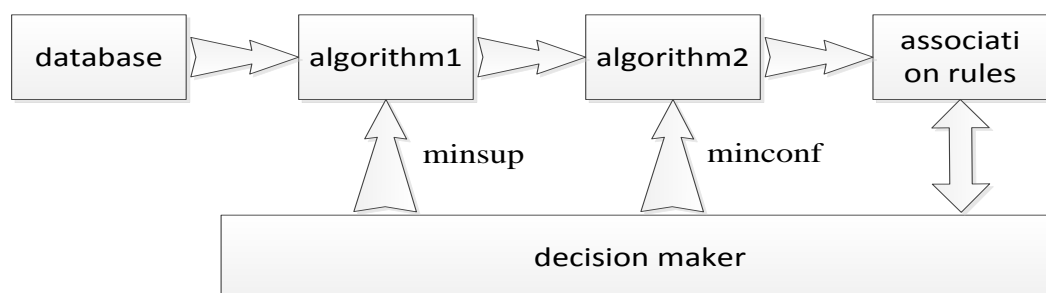


FIGURE 1 The model of association rules mining

4 Introduction of fuzzy association rules

The association rules are used most widely in market basket analysis. In a database, except items, there is

numerical information relating to these items as well, such as the quantity and the price etc. of commodities. When first proposed, the association rules just considered the membership information, while ignored the numerical

one. Researchers therefore focus on whether the numerical information is useful in further mining or not.

Generally, in reality, when carry out association rules mining of data which are not Boolean or categorical, researchers transform them to Boolean one. However, the transformation damages the edge data seriously, therefore influencing the mining results. Afterwards, the fuzzy concept was introduced in the association rules, and the problem of edge data was readily solved. The fuzzy association rules request the fuzzy concept is fuzzy as well as the membership functions. In the transformation, using different membership functions can result in different results, which influencing the mining results significantly [4].

Let $I = \{I_1, I_2, \dots, I_m\}$ be the attribute set of the database D. For any attribute $I_i (1 \leq i \leq m)$, it can be divided into fuzzy attributes of number q_j using fuzzy membership functions. After the original numeric attributes being divided into fuzzy ones, the database D was transformed to fuzzy database D_f , the attribute set of which is $I_f = \{I_1^1, I_1^2, \dots, I_1^{q_1}, I_2^1, I_2^2, \dots, I_2^{q_2}, \dots, I_m^1, I_m^2, \dots, I_m^{q_m}\}$, and the ranges of all the new attributes expand to $[0, 1]$.

Definition 9. The support of record i in fuzzy item sets $X = \{x_1, x_2, \dots, x_p\} \subset I_f$ to the fuzzy item set X is defined by the following formula:

$$\text{Sup}T_i(X) = x_{1i} \wedge \dots \wedge x_{pi} \text{ or } \text{Sup}T_i(X) = x_{1i} \times \dots \times x_{pi}, \quad (4)$$

where x_{ji} represents the value of fuzzy item x_j on the record i , and $x_{ji} \in [0, 1] (i=1, 2, \dots, n \quad j=1, 2, \dots, p)$.

Definition 10. The support of the whole fuzzy item set $X = \{x_1, x_2, \dots, x_p\} \subset I_f$ to X is defined by the following formula:

$$\text{sup}(X) = \frac{\sum_{i=1}^n \text{sup}T_i(X)}{|D_f|}, \quad (5)$$

where $|D_f|$ represents the number of the transactions in the database. If the support of a fuzzy item set is not less than the set fuzzy minsup, X is a fuzzy frequent item set.

Definition 11. Similar to Boolean association rules, in the implication $X \Rightarrow Y$ in fuzzy association rules, X and Y indicate the antecedent and the consequent of the fuzzy association rules, respectively. Similarly, $X \subseteq I_f, Y \subseteq I_f$, and there are no relevant items from the same attributes in, $X \neq \emptyset, Y \neq \emptyset, X \cap Y \neq \emptyset$, and $I = X \cup Y$.

Definition 12. The $\text{sup}(X \Rightarrow Y)$ and $\text{conf}(X \Rightarrow Y)$ of implication $X \Rightarrow Y$ in the fuzzy association rules are defined by the following formulas:

$$\text{sup}(X \Rightarrow Y) = \text{sup}(X \cup Y), \quad (6)$$

$$\text{conf}(X \Rightarrow Y) = \frac{\text{sup}(X \cup Y)}{\text{sup}(X)}, \quad (7)$$

where $X \subseteq I_f, Y \subseteq I_f$, and there are no relevant items from the same attributes in $X \neq \emptyset, Y \neq \emptyset, X \cap Y \neq \emptyset$, and $I = X \cup Y$.

Similar to Boolean association rules, the minsup and the minconf are set by the decision maker prior to discovering fuzzy association rules. The fuzzy association rules mining is carried out in the following process: determining membership functions, establishing transactional database, discovering fuzzy association rules, and finally explaining discovered rules.

5 The application of fuzzy association rules in the employment data analysis of a higher vocational college

This research was on the basis of the accumulated relevant data of graduates of the vocational college in Hainan, China and analysed the employment trend and rules using data mining. The results are able to provide suggestions on vocational counsel and educational reform for the management and decision-making sections of the college, thus promoting the sustainable development of the college [5].

In this research, it is undoubtedly of great practical significance to apply data mining to the analysis of employment and educational reform, to discover the internal and hidden information from plenty of historical data using fuzzy association rules, and to employ the information to the decision-making of the college.

5.1 DATA PREPARATION

The collected employment information of 379 graduates from different majors of the vocational college in Hainan from 2012 to 2013 was employed in the research, and the relevant information such as name, gender, and birthday etc. was omitted. Then the following attributes were carried out association rules mining, including the average scores of common required courses, professional basic courses, and professional courses, majors, industries, income, and business natures etc.. Since the data size is large, this research merely shows partial data below, as shown in Table 1.

TABLE 1 The employment data of the vocational college in Hainan

Serial number	Average score of common required courses (ASCRC)	Average score of professional basic courses (ASPBC)	Average score of professional courses (ASPC)	Major	Industry	Income/month	Business nature
001	86	82	84	software technology	IT	3800	private operated company
002	71	67	71	software technology	marketing	2700	private operated company
003	81	72	87	software technology	IT	4200	foreign company
004	81	87	68	software technology	education	2700	state-owned enterprise
...
376	65	72	69	electronic commerce	finance	2200	private operated company
377	71	67	71	electronic commerce	marketing	2100	private operated company
378	83	72	87	electronic commerce	education	3300	state-owned enterprise
379	81	83	89	electronic commerce	marketing	4100	foreign company

5.2 DISCOVERING ASSOCIATION RULES

1) The minsup and minconf are set to be 0.3 and 0.5, respectively. Then the clustering centres of the data were calculated using C means clustering algorithm. The centres are displayed in Table 2 [6]. The fuzzy database is not demonstrated due to its large size.

TABLE 2 Clustering centre of different attributes

Attribute	Clustering centre		
	1	2	3
ASCRC	58	71	81
ASPBC	51	73	83
ASPC	57	76	83
Income	2100	3300	3900

2) The Counts of all the fuzzy items were calculated and the fuzzy items with Counts not less than 35 were classified into frequent 1-itemsets L_1 . Then C_2 was generated by connecting the item sets L_1 , and the fuzzy items corresponding to the same attributes were not connected. Afterwards L_2 was generated by C_2 and C_3 was generated by connecting L_2 . In the case that $C_4 = \phi$, the mining is therefore finished. The generated association rules are indicated in Table 3 [7].

TABLE 3 The discovered association rules

Association Rules	Support (%)	Confidence (%)
{ASCRC.high, ASPC.medium} \Rightarrow not related	33.4	72.6
{ASCRC.high, ASPC.high} \Rightarrow related	34.7	64.1
{ASCRC.high, ASBPC.high, ASPC.high} \Rightarrow high income	37.5	73.6

6 Conclusions

The research employed the fuzzy association rules algorithms in the analysis of the employment of the higher vocational college and discovered the relationships among the scores and the employment attributes. According to this method, the fuzzy C means clustering algorithm was used to cluster the quantitative attributes,


The association rules in Table 3 indicate that most of the graduates with medium scores of professional courses and high scores of professional basic ones work in the industries not related to their majors; most of the graduates with high scores of professional and professional basic courses work in the industries related to their majors; and the incomes of the graduates with high scores of common required, professional and professional basic courses are generally at a high level.

The above data indicate that, in order to work in the industries related to their majors after graduation, students require to concentrate their attentions on all the following courses, including common required, professional basic, and professional courses, so that they can develop in an all-around way; meanwhile, if some students are interested in other majors, proper suggestions need to be given to improve their relevant professional quality, so that they are ready for the employment in the future. The information provides guidance and references for the training scheme of the higher vocational college. And it is useful in the training of the applied talents needed for society and improving the employment rate of students.

and then the clustering centres were mined using the association rules mining. However, the research needs to be perfected in some aspects as well. For example, the research discovered all the strong association rules, of which some with high support and confidence are not valuable in practical applications. In addition, whether the discovered quantitative association rules are valuable in application or not needs to be verified as well.

References

- [1] Agrawal R, Imielinski T, Swami A 1993 Mining association rules between sets of items in large databases, in *Proc. ACM SIGMOD Int. Conf. Management of Data* Washington DC 5 207-16
- [2] Han J W 2008 *Data Mining Concepts and Techniques* Beijing: China Machine Press 146-76 (in Chinese)
- [3] Au W H and Chart K C C 2003 *IEEE T. Fuzzy Syst.* 11(2) 238-48
- [4] Yan P and Chen G Q 2004 *Fuzzy Syst. Math.* 18(1) 279-83
- [5] Ru J L 2013 *The research and implementation of algorithm in the employment of higher vocational colleges in the mining of association rules* Chengdu: Univ. Electr. Sci. Technol. China 2-6 (in Chinese)
- [6] Wen H 2013 Knowledge map mining of financial data *J. Tsinghua Univ. (Science and Technology)* 1(1) 68-76 (in Chinese)
- [7] Zhang X P 2013 Application of quantitative association rules in college employment information data *Comput. Technol. Dev.* 23(11) 199-203 (in Chinese)

Authors	
	<p>Laiquan Liu, born in August 14, 1979, Shaanxi Province of China</p> <p>Current position, grades: Hainan College of Software Technology, associate professor, senior engineer</p> <p>University studies: Shanxi Normal University (Bachelor of Science in Educational Technology), 2003; Tianjin University of Science & Technology (Master's degree in Control engineering), 2012</p> <p>Scientific interest: vocational education, multimedia application and data mining</p> <p>Publications: more than 15</p> <p>Experience: Hainan Academy on Computers of China (2008-)</p>
	<p>Li Lei, born in October 16, 1979, Shaanxi Province of China</p> <p>Current position, grades: Chongqing University of Arts and Sciences, lecturer</p> <p>University studies: Shanxi Normal University (Bachelor of Science in Computer science and technology), 2003; Chongqing University (Master's degree in Computer technology), 2013</p> <p>Scientific interest: web application development and data mining</p> <p>Publications: more than 5</p>
	<p>Yanrui Lei, born in December 12, 1980, Shaanxi Province of China</p> <p>Current position, grades: Hainan College of Software Technology, lecturer, engineer</p> <p>University studies: Shanxi Normal University (Bachelor of Science in Computer science and technology), 2003; Sun Yat-Sen University (Master's degree in Software engineering), 2013</p> <p>Scientific interest: web application development and data mining</p> <p>Publications: more than 15</p> <p>Experience: Hainan Academy on Computers of China (2008-)</p>

Information technology-based promotion of educational resource sharing

Jia Geng*

China Center for Industrial Security Research, Beijing Jiaotong University, Beijing - China

Received 1 March 2014, www.tsi.lv

Abstract

Information technology (IT) has revolutionary influence on the distribution of educational resources, and it plays a role in promoting the sharing of educational resources. Based on absorbing and inheriting the results of previous studies, this dissertation proposed to collect funds to construct large-scale digital education information resource database with centring on counties without building resource centre at basic education schools at all levels. In addition, the schools shall jointly expand and enrich the central educational resources to achieve sharing and co-construction of educational resources in the region. This dissertation also constructed a model of promoting sharing of educational resources by information technology in urban and rural areas.

Keywords: Information Technology, Educational Resource, Sharing

1 Introduction

In China, the unbalanced development of elementary education is the biggest problem preventing the education sector from an entirely balanced development, and the difference between urban and rural areas in terms of educational resources and quality teacher resources directly results in inequity in students' receipt of educational opportunities and the educational quality [1]. Information technology has advantages of quick knowledge dissemination, wide coverage and sharing of resources [2], [3]. Therefore, from our national conditions, the author wishes to make full use of information technology and networking to overcome the barriers of time and space so that more people, especially those in underdeveloped areas, can share the high-quality education resources and the elementary education can make great-leap-forward development in these areas.

Since 2008, the author has assumed two Beijing educational research projects Development and Construction of IT Subject Virtual Learning Platform and Research on Digital Campus Construction and Application Practices to conduct researches on promoting the sharing of educational resources by means of information technology. Based on the fact that the information-based hardware supporting environment for elementary education in Beijing has achieved great-leap-forward development, the practical research was conducted at a county (district) level for the unified construction of educational resource centre to achieve co-construction and sharing of educational resources in a region [4]. The research was to confirm that IT could really promote the sharing of educational resources, and analyse the effect of implementation. This research will provide new methods and approaches for narrowing the educational gap in the information age and it is of great

significance on exploring the urban and rural areas' sharing in elementary education in Beijing and even all over China.

2 Model of IT-based Promotion of Regional Sharing of Educational Resource

Combined with the actual conditions that the larger gap in education resource distribution in the implementation of the "School Access to ICTs" project in our country, taking into account the fact that in next few years the focus of China's development trend in basic education informationization is the construction of digitized educational resources, and based on the theory of educational equity and the characteristics of digital educational resources [5], [6], it is suggested to take each county as centre and collect funds for the construction of large digital resource centre of basic educational information to meet the use of education, teaching and management by basic education schools at all levels in the region. The basic education schools at all levels in this region will no longer build the school's educational resource centre, instead they can directly share all resources at the county's central educational resources centre (Figure 1), in order to achieve integrated development of basic education resources and ensure that students can have equal opportunity when using educational resources.

The county's central digitized educational resources centre is designed to solve the problem of resource sharing for basic education at all levels in this area and avoid the low-level and redundant construction of resources by every school. In this way, the resources can be constructed by the county's Department of Education in a unified way. As a result, the basic education schools at all levels can be free to focus on doing a good job in

* *Corresponding author* e-mail:jessie412@126.com

constructing the “School Access to ICTs” project with high-speed interconnection with the central resource centre and the teachers’ IT skills training. To ensure that schools can fully share the rich resources of the county’s central resource centre, the “School Access to ICTs” project must have the basic configuration, including the high-speed network interworking with county’s and city’s central resource centres; the network connection between school’s own information collection room and studios of teachers’ CAI courseware together with network curriculum; and the computer classrooms and multimedia classrooms for students with network connection. The qualified schools can consider the network connection between teachers’ living quarters and students’ dormitory, etc.

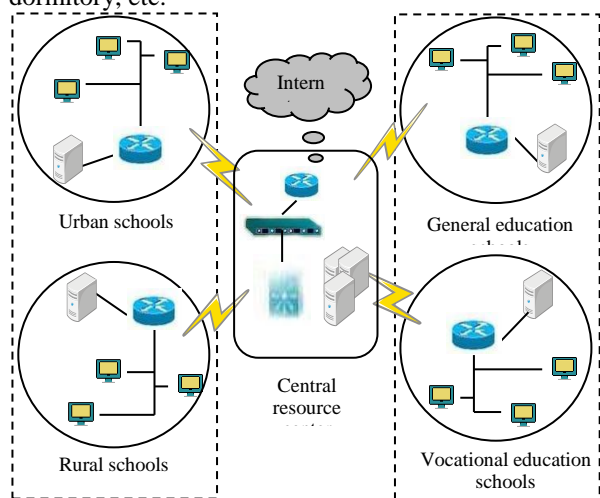


FIGURE 1 Model of IT-based Promotion of Regional Sharing of Educational Resource

3 Construction of Central Resource Centre

Linmei Liang and Xinmin Sang’s survey on the status quo and problems of regional education informationization found that the lack of overall and systemic planning and management of regional education informationization and the lack of effective mechanism for supervision and management of “investment-benefits” in regional education informationization resulted in serious waste of hardware in education informationization construction in some regions as well as the status quo of software resource lack and failure in sharing high-quality resources [7]. To advance the education informationization process equally and build regional education resource centres effectively, the focus shall be put on the overall planning and macro-management, changing from resource development orientation to the orientation of resource integration as well as joint-construction and sharing.

In order to build regional education resource centres more effectively to provide them with stronger guarantee in terms of quality, efficiency and sustainable development, etc. and give full play to the duties of the regional educational administrative departments in the

balanced development of education informationization to avoid redundant construction works, the author suggests the building of regional central resource centres for basic education. In one region, the construction of educational resources shall adopt the resource content construction mode with the organization works provided by grassroots educational administrative departments and the organization outline of local textbook knowledge or discipline catalogues. In the mode, the resources are provided by front-line teachers and supplemented by purchase, constructing in distributed way and using in shared manner.

3.1 THE CONNOTATION AND OBJECTIVES OF CENTRAL RESOURCE CENTER

The construction of educational resource centres at primary and secondary schools serve the education and teaching, and therefore they must comply with the characteristics of the primary and secondary education [8]. It is necessary to take into account not only the environment for teachers and students to take advantage of the resources but also the actual needs for resource types [9]. For regional central resource centres for primary and secondary education, its connotation can be specifically explained as follows:

First, the “School Access to ICTs” project is the premise based on which the regional central resource centres for primary and secondary schools are constructed. Only if all primary and secondary schools in the region implement the “School Access to ICTs” project, can they share the central resource centre through network.

Second, the central resource centre will permeate into the new curriculum reform of primary and secondary education. The contents of resource centre will be designed to serve the new personnel training, and the resource centre management will be aimed to ensure the normal operation and the user-friendly access.

Finally, the overall plan and the centralized financial resources for building have the profound significance to adapt to the concepts of autonomy, cooperation and inquiry learning in the network environment, improve the effectiveness of education informationization and create informationization conditions for sharing of quality educational resources in urban and rural areas. In addition, the balanced development of education informationization at primary and secondary schools can be promoted.

The core objective of constructing the regional central resource centre for basic education is to take full advantage of information technology, create changed ecological environment of teaching and learning mode in modern school education and improve the quality and efficiency of educational resources. It can help to narrow the gap between urban and rural schools’ education informationization. The construction of regional central resource centre for basic education shall be guided with

the new concept of teaching, reform concept and performance management concept, surpass the building mode of traditional resources and cultivate the creative talents who meet the needs of the information society.

3.2 OVERALL PLANNING FOR CENTRAL RESOURCE CENTER

3.2.1 Network Design Proposal

The central resource centre for primary and secondary education is mainly designed for all primary and secondary schools in the region, therefore the basic condition to ensure that schools can connect to the resource centre, and it is best to connect among schools. If there is no educational MAN built in this region, schools can only get access to the resource centre through the Internet. As a result, all individual school shall pay a lot to get access to Internet, and the single campus network is like an "isolated island of information". It is difficult to form a network within the region. They are not able to make better communication and sharing with each other. To establish the educational MAN has the advantage that the educational institutions in the region can be connected to the network and ultimately form a regional basic framework featured interconnection, interaction, information exchange, resource sharing and remote education. The educational MAN also plays a role in adjusting the regional education layout structure, optimizing educational resource distribution and achieving the overall development of regional education. The regional central resource centre for basic education should be built in the centre of the educational MAN. The schools in the region can directly share its resources and the interconnection between the schools can be achieved.

3.2.2 Construction of Resource Centre

In accordance with the *Technical Specifications for Educational Resources Construction* issued by the Ministry of Education, the commonly used information resources include nine categories, i.e. media materials (including text, graphic/image, audio, video and animation), test question, test paper, courseware and network courseware, case, document literature, answer to FAQ, resource index and network course. In addition, other types of resources, such as e-book, tool software and movie, etc., can also be added according to the actual needs.

To fully develop and make the best of resources in the construction of education resource centre, builders must store, organize and reveal information by scientific and rational classification method that meets people's way of thinking. There are three main classification attributes, namely discipline, applicable object and material types. These three categories are the fundamental basis for organizing resources (Figure2).

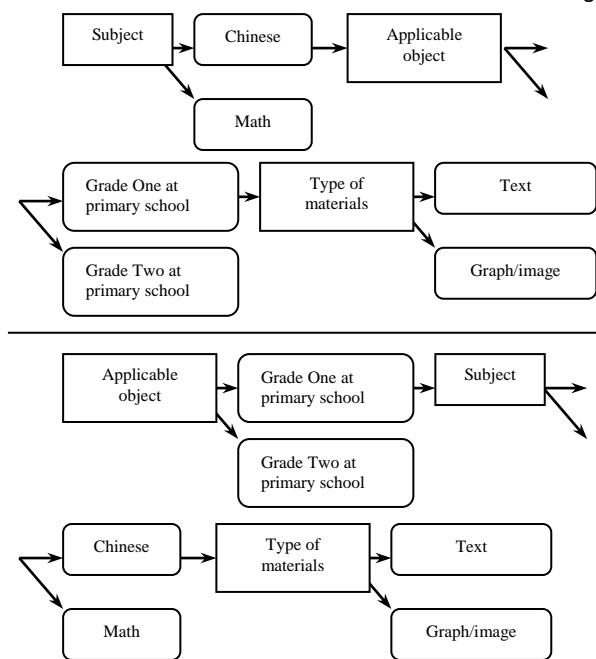


FIGURE 2 Education Resource Classification System

In fact, the construction of central resource centre is not completely from scratch, instead it is re-planned based on the integration of original high-quality resources and it is to continue to enrich resources and improve resource quality according to the new proposal standards. By combining a variety of educational resource classification theories and learning from some successful experience in building resource centre, the author proposed, in accordance with the principle of convenience for user, that the construction of regional resource centre should include four modules. They are the teaching and learning resource centre, information resources and library, the digital video library and the resource management system (Figure 3).

3.2.3 Resource centre management

The resource centre management includes resource management and system management. The resource management shall support user's operating features on resources, while the system management is to improve the operation security of the resource centre. The functions of resource management include resource acquisition, resource retrieval, resource upload and batch data entry, examination and release of resources, resource browsing, comment entry and display, resource filing, resource downloads, and usage record tracking, etc. The system management function includes data backup, user management, network fault management, network configuration management, network charging management, statistics and analysis, as well as data releasing and sharing.

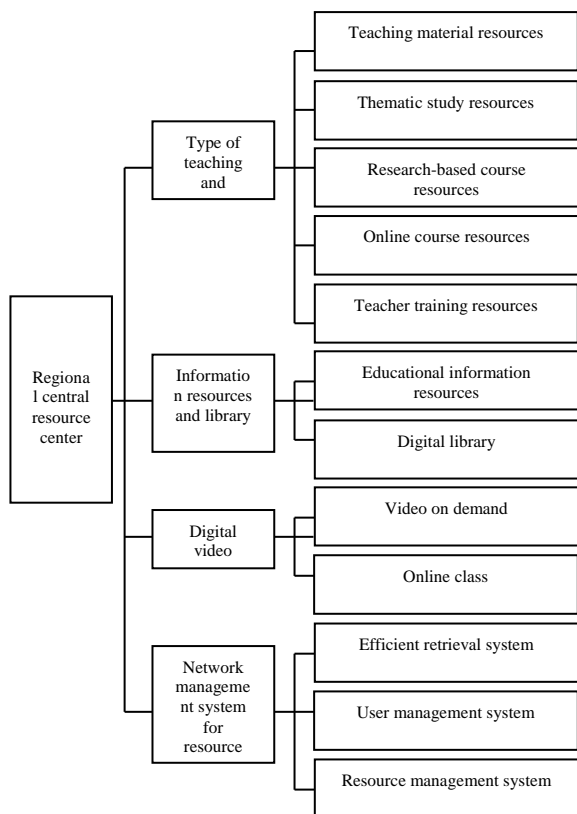


FIGURE 3 Function Module of Educational Resource Centre

For the logical structure of the management system of resource centre, please refer to Figure 4.

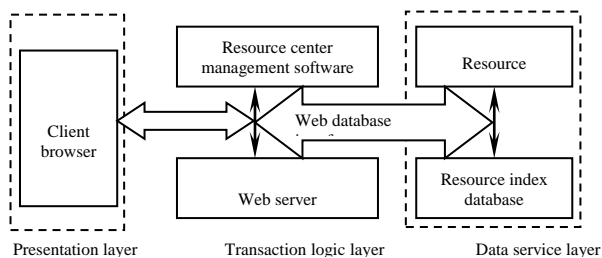


FIGURE 4 Logical Structure Chart of Resource Centre Management System

3.2.4 Application of resource centre

The ultimate goal of building resource centre is the application in teaching. Whether the resource centre can be applied flexibly and the utilization rate of resources is related to the effectiveness of resource centre construction [10]. In terms of resource centre application, the author provided two suggestions. One is to improve the user's application ability, and the other is to combine the resource centre application with the teaching reform.

To improve user's ability to use resources, training is a necessary way. The trainees should include teachers, students, technical personnel for education and teaching management personnel. Training of students should be integrated into the daily classroom teaching of information technology. The training for teaching

management personnel shall focus on getting familiar with the environment of the resource centre as well as some ideas and operations of information management. Teachers are the main user of resources, so the training for teachers is critical. Training for teachers shall include the concepts of using resource centre, the use and operation of resource centre, and, the most important one, the ability to integrate teaching by making use of resource centre. The training for teachers should adopt the mode combining theory and practice, and it shall achieve the effect of practicing while learning. During the specific training, the teachers shall be allowed to practice while learning; while when making use, teachers shall be allowed to learn while practicing. It means continuing to accumulate experience, mastering skills, enhancing concepts and thus continuing to improve the ability to use resources in the process of usage.

The role of the resource centre is not just to provide ancillary services for teaching. In fact, it is the core of teaching reform. The resource centre is not only teaching materials but also innovation works in teaching; the resource centre is not only a new form of teaching media but also new teaching communication system; the resource centre is closely associated with almost all aspects of the teaching reform. To allow the resource centre to better serve the teaching reform, we can take a series of measures, for example, the use of existing materials, the independent development of school-based teaching textbooks or localization of teaching materials; to use the resource centre to promote in-depth integration of information technology and curriculum; and to make use of the resource centre to realize distance communications between rural and urban students and teachers as well as distance education.

3.3 EVALUATION OF CENTRAL RESOURCE CENTER CONSTRUCTION

In the integration of educational resources, the central resource centre shall not only focus on quantity but also pay attention to the quality. In order to ensure the construction quality of the education and teaching resource centre, effective resource assessment mechanism should be established at macro level. Emphasis shall not only be laid on the completion of task and the achievement of verification results, but also on the assessment of all aspects of the overall process so as to standardize the construction management for resource centre and improve accuracy and reliability of verification results; at the micro level, the operating indicators for resource centre review shall be defined. The establishment of evaluation indicators shall highlight its educational characteristics and requirements for resource construction under the new curriculum standards with the fundamental starting point of quality-oriented education and innovation education.

Currently, the evaluation criterion for construction of central resource centre has not been unified yet. The

operability of various evaluation criteria is not so good. However, generally speaking, the common criteria are reflected in scientific feature, teaching, technical feature and standardization (Table 1). In addition, under the premise of conforming to the common criteria, there are

two-level evaluation indicators according to the characteristics of different resources. In general, there are a lot of evaluation criteria developed for different purposes by different departments and different people, and they have not been unified.

TABLE 1 Evaluation Principles for Central Resource Centre Construction

Principle	Requirements
Educational Principle	The fundamental principle for building resource centre. The collection and integration of any resources should work closely with education and teaching as well as subject knowledge points. It shall not only meet the needs of curriculum standards and instructional program but also have clear teaching objectives and obvious teaching effects
Scientific Principle	The technical requirements and specifications shall meet the standards published by the State. No wrongly written or mispronounced characters and ambiguous scientific errors can appear in the integrated resources. The resource classification and integration shall comply with the relevant national standards, and the resource attributes shall be marked in accordance with requirements of <i>Technical Specification for Educational Resource Construction</i>
Practical and Intuitive Principle	Construction of resource centre and integration of resources shall be easy to operate and use. Even teachers not so familiar with computer operation shall be able to easily find resources suitable for their teaching; in addition, for students who need help in learning shall be able to easily find the required learning resources
Interactive and Intelligent Principle	The content is not simple accumulation of materials or show of knowledge contents. It is more important to pay attention to its interaction and intelligence to show the cognitive involvement and the learning subjectivity in learning process
Interesting and Artistic Principle	The constructed and integrated resources shall highly interesting and can stimulate the students' motivation and interest in learning. In addition, the friendly interface featured visualization and creativity shall be provided to have strong visual appeal

4 Conclusion

Currently, China's social development has entered an unprecedented new stage. As one of the important factors that determine social development, elementary education is closely related to its era. In the information age, information technology is fully involved in the field of education, so the author advocates the information technology-based promotion of educational resource sharing to achieve the ultimate goal of closing the educational gap and finding the new ways, new approaches and new ideas to study the issue of educational equity in the information age.

The construction of central resource centre is a long-term dynamic development process, so we shall consider the dynamic and balanced development of high-quality education resources from the perspective of sustainable development. With the combination of resource construction and teaching reform and the implementation of research topics on practices of regional education informationization development accordingly, the effective use of educational resources shall be promote to accelerate in-depth reform of various disciplines; furthermore, resources shall be integrated and organized from the perspective of knowledge management while exploring the sustainable development way of resource construction.

References

- [1] Xin Li, Xiaoyong Hu, Rong Miao 2009 *China Educational Technology* 2 55-8 (in Chinese)
- [2] Song Surong 2007 Characteristics of Educational network information resources and retrieval approach *Teaching and on management* 18 48-9 (in Chinese)
- [3] Peng Hongguang, Ling Junfeng 2010 The mechanisms and strategies to promote balanced development of compulsory education with information technology *China Educational Technology* 10 33-39 (in Chinese)
- [4] Geng Jia, Zhang Gang 2012 Strategic thinking of promoting elementary education informationization in Xicheng District *Information technology education in primary and secondary schools* (Z1) 102-4 (in Chinese)
- [5] Han Xiaoguang 2009 Development of network information resources serving in teaching and scientific researches *Heilongjiang Science and Technology Information* 24 124 (in Chinese)
- [6] Liu Peng 2007 Educational resources have a distribution center *Shanghai Science and Technology News* 2007-02-28A01 (in Chinese)
- [7] Liang Linmei, Sang Xinmin 2004 *China Educational Technology* 8 21-4 (in Chinese)
- [8] Liu Yang, Gao Hongyuan 2005 *Foreign Educational Research* 7 69-72 (in Chinese)
- [9] Tang Zuoshou 2007 *China Educational Technology* 6 62-5 (in Chinese)
- [10] Qian Zhongping 2012 *Contemporary Education Science* 11 52-4 (in Chinese)

Author	
	<p>Jia Geng, born in April 12, 1983, Beijing, China</p> <p>Current position: Postdoctor of Information Management in China Center of Industrial Security Research, Beijing Jiaotong University</p> <p>University studies: Pedagogy Major, Tarlac State University, Philippines, Doctorate</p> <p>Scientific interest: Information Construction Subjects and Schools Management</p> <p>Publications: 10 Papers</p> <p>Experience: The author is the instructor and researcher of Modern Educational Information Technology Center of Xicheng District, Beijing, China, who takes charge of two "12th Five-Year Plan" information construction subjects and schools management; provides training for teachers on educational technology; acts as judge of competitions related to educational technology. As assistant commissioner in Beijing Municipal Commission of Education, the writer mainly took in charge of the project of quality report for 2013 Beijing primary and secondary school education informationization construction and application level and research on assessment indicator system. Academic essays, teaching designs and courseware have won First Prizes at different national educational competitions.</p>

Optimized preparation of γ -polyglutamic acid/chitosan nanocapsule

X Ma^{1,2}, M M Zhang¹, S W Chen^{3*}

¹ School of Perfume and Aroma Technology, Shanghai Institute of Technology, Shanghai- China

² Bright Dairy & Food Co. Ltd, Shanghai- China

³ Shanghai Jiaotong University affiliated Sixth people's hospital

Received 1 March 2014, www.tsi.lv

Abstract

This paper described the preparation of novel biodegradable nanocapsule based on self-assembly of γ -Polyglutamic acid (γ -PGA) and chitosan (CS). After the Plackett-Burman design (PBD), the impact of mass concentration and volume of γ -PGA and pH value of CS were characterized by size and PDI of the nanocapsule. A Box-Behnken design (BBD) was used to optimize the preparation of the nanocapsule. The optimized condition was: pH value of CS was 4.0; volume of γ -PGA was 18mL; mass concentration of γ -PGA was 0.1g/L. The Z-Ave and PDI of the nanocapsules prepared under the best conditions were 175 nm and 0.15 respectively. In this work, we have shown that nano-sized particles have been successfully assembled from the γ -PGA and chitosan without employing covalent linkages between these biopolymers. These results will provide a novel concept in the design of carrier systems composed of polyion complex (PIC).

Keywords: γ -Polyglutamic acid, chitosan, nanocapsule, optimization, response surface methodology

1 Introduction

γ -Polyglutamic acid (γ -PGA) is an anionic polymer which consists of D-or L-glutamic acid via α -amino and γ -carboxyl group to form γ -glutamyl bond (Figure 1). As a water-soluble biodegradable polymeric material, it is widely used because of the edible, non-toxic, cohesiveness and other characteristics. In the current study reports, γ -PGA is mainly used as drug carriers [1], bio-adhesive in high value-added field of medicine. What's more, γ -PGA can be used as flocculants, heavy metal chelating agent for water treatment [2], it can also serve as moisture holding agent, antifreeze, preservatives used in fruits, food, vegetables, health products, cosmetics and so on. There are a large number of free carboxyl groups in the γ -PGA molecular chain which can be modified by chitosan [3], polyethylene glycol [4], L-phenylalanine [5] and so on. The modified parts of γ -PGA could be as drug carriers such as amoxicillin [6], insulin [7]. What is more, γ -PGA can also be drug carriers direct binding with camptothecin [8], paclitaxel [9], penicillamine [10] and so on. As a new, safe and harmless biological adhesive γ -PGA can also be used to control continuous bleeding in tissue and repair aortic cutting.

Chitosan (CS) whose active group was amino was a deacetylated chitin product, containing β -[1 \rightarrow 4]-linked 2-acetamido-2-deoxy-D-glucopyranose and 2-amino-2-deoxy-D-glucopyranose units (Figure 2). Because of

these amino groups in the molecular structure, CS is a weak base and the only natural cationic polymers. CS has not only good biocompatibility, reliability safety and low immunogenicity, but also hypolipidemic, nor cholesterol, anti-bacterial and increasing immune physiological activities. As a good natural bio-medical material, CS can be used as wound dressing materials [11] and drug carries [12, 13].

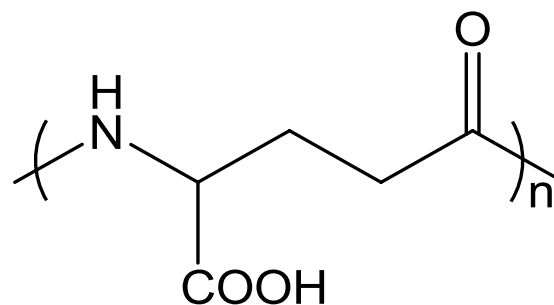


FIGURE 1 The molecular structure of γ -PGA

The response surface methodology (RSM) design was one of experimental design methods which could find improved or optimal process settings [14, 15]. The purpose of our research was:

- 1) The critical factors that affect the size of γ -PGA/CS nanocapsules were picked via a Plackett-Burman design (PBD);
- 2) The preparation conditions of γ -PGA/CS nanocapsules was optimized via a Box-Behnken design (BBD).

* Corresponding author e-mail: chenshiven@126.com

2 Materials and methods

2.1 MATERIALS

γ -Polyglutamic acid (γ -PGA) was prepared in our laboratory using the biosynthetic methods. γ -PGA was purified as follows: the crude products were dissolved in distilled water to give a mass concentration of 2% and the dialyzed against distilled water for 24h. The dialysate was lyophilized in form of white powder and used for further experiments. CS was purchased from Shanghai Plus Bio-Sci&Tech Co., Ltd. (Shanghai, China) (Grade: BR) with molecular weight of ~150,000 and 90% degree of deacetylation. The other reagents were purchased from Sinopharm Chemical Reagent Co., Ltd. (Shanghai, China). The γ -PGA aqueous solution was prepared with different concentration. The solution of CS was prepared with different concentration in aqueous solution with 1% acetic acid. The water used in the experiments is all distilled. 20% NaOH aqueous solution and 2mol/L HCl were used to adjust the pH of the γ -PGA aqueous solution and CS solution.

2.2 PREPARATION OF CS/ γ -PGA NANOCAPSULE

γ -PGA solution and CS solution were used for preparation of γ -PGA/CS nanocapsule. The preparation technique based on the self-assembly of polyelectrolytes, in which the anion polymer (PGA) interacts with the cation polymer (CS) at normal temperature and pressure. The preparation methods were as follow: the γ -PGA solution was dropwise added into the CS solution under continuous stirring; the mixture was further stirred for 60 min at ambient temperature after adding γ -PGA solution. The CS/ γ -PGA nanocapsules were obtained by lyophilization of the mixture.

The experimental design (Table 2) was contrived based on Design-Expert 7.0. Z-Ave was the response variable. After data processing by Design-Expert 7.0, A (pH value of CS), D (Mass concentration of γ -PGA) and J (Volume of γ -PGA) were selected for further analysis by Box-Behnken. The optimum preparation conditions selected by PBD were used for further analysis by BBD.

3.2 BOX-BEHNKEN DESIGN (BBD)

As one design method of RSM, BBD was used to optimize the preparation conditions of γ -PGA/CS nanocapsule. Three factors selected by PBD were tested in this design. Each factor was tested at 3 levels (-1, 0 and 1, Table 3). The experimental design (Table 4) was contrived based on Design-Expert 7.0. All experiments were performed in triplicate. Average size (Z- Ave) and polydispersity index (PDI) were the response variables.

2.3 SIZE AND ZETA-POTENTIAL MEASUREMENTS

The size and Zeta-potential of the nanocapsules were measured at 25 °C by a Malvern zetasizer Nano-ZS instrument (Malvern Instruments Ltd.). Each sample was measured three times and average serial data were calculated.

3 Optimization by RSM

3.1 PBD

PBD is an experimental design method for two-level and the most popular fractional design. It is suitable for optimizing multitudinous factors, since it makes it possible to pick up the relevant factors from a long list (16).

In the process of preparation conditions optimization, 9 factors (Table 1) were tested. Each factor was tested at both low (-1) and high (1) levels (Table 1).

TABLE 1 Levels and experimental factors of Plackett-Burman design

Code	Name	Unit	Levels	
			Low level (-1)	High level (1)
A	pH value of CS	/	3.0	6.0
B	pH value of γ -PGA	/	6.0	7.4
C	Mass concentration of CS	g/L	0.20	1.00
D	Mass concentration of γ -PGA	g/L	0.20	0.40
E	Stirring speed	r/min	50	150
F	Concentration of Mg ²⁺	mol/L	0.002	0.006
G	Reaction time after dropping	min	10	60
H	Dropping speed	mL/h	0.10	1.00
J	Volume of γ -PGA	mL	2	10

TABLE 2 Plackett-Burman experiment design and response values

St d No	R u n	A	B	C	D	E	F	G	H	J	K	L	Z- Ave/ nm
1	13	1	1	-1	1	1	1	-1	-1	-1	1	-1	793.0
2	5	-1	1	1	-1	1	1	1	-1	-1	-1	1	179.0
3	10	1	-1	1	1	-1	1	1	1	-1	-1	-1	779.0
4	1	-1	1	-1	1	1	-1	1	1	1	-1	-1	300.5
5	3	-1	-1	1	-1	1	1	-1	1	1	1	-1	219.5
6	15	-1	-1	-1	1	-1	1	1	-1	1	1	1	263.5
7	14	1	-1	-1	-1	1	-1	1	1	-1	1	1	170.0
8	4	1	1	-1	-1	-1	1	-1	1	1	-1	1	206.7
9	9	1	1	1	-1	-1	-1	-1	-1	1	1	-1	190.0
10	11	-1	1	1	1	-1	-1	-1	1	-1	1	1	738.5
11	7	1	-1	1	1	1	-1	-1	-1	1	-1	1	433.0
12	12	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	150.5
13	8	0	0	0	0	0	0	0	0	0	0	0	197.0
14	6	0	0	0	0	0	0	0	0	0	0	0	193.5
15	2	0	0	0	0	0	0	0	0	0	0	0	190.0

TABLE 3 Levels and experimental factors of BBD

Factors	Code	Level		
		-1	0	1
pH value of CS	X1	3.0	4.5	6.0
Volume of γ -PGA(mL)	X2	5	10	15
Mass concentration of γ -PGA (g/L)	X3	0.2	0.4	0.6

Table 4 Box-Behnken experiment design and response values

No.	Run	X1	X2	X3	Y1	Y2
1	1	-1	-1	0	340.5	0.2895
2	14	1	-1	0	213.5	0.1425
3	6	-1	1	0	326.0	0.2365
4	11	1	1	0	381.0	0.2805
5	2	-1	0	-1	293.5	0.4065
6	15	1	0	-1	233.5	0.1845
7	13	-1	0	1	367.5	0.2220
8	7	1	0	1	366.0	0.3925
9	9	0	-1	-1	238.5	0.3475
10	3	0	1	-1	224.5	0.2685
11	10	0	-1	1	295.5	0.2055
12	4	0	1	1	327.0	0.2045
13	12	0	0	0	271.0	0.2325
14	5	0	0	0	281.0	0.2020
15	8	0	0	0	305.5	0.2255

Using the Design-Expert 7.0, the analysis, the results and the second-order empirical model of each factor on the responses were obtained.

$$Y_1 = \alpha_0 + \sum_{i=1}^k \alpha_i X_i + \sum_{i=1}^{j-1} \sum_{j=1}^k \alpha_{ij} X_i X_j + \sum_{i=1}^k \alpha_{ii} X_i^2, i \neq j$$

$$Y_2 = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i=1}^{j-1} \sum_{j=1}^k \beta_{ij} X_i X_j + \sum_{i=1}^k \beta_{ii} X_i^2, i \neq j$$

where Y1 is the response value that is representative of Z-Ave; Y2 is the response value that was representative of PDI; α_0 , α_i , α_{ij} , α_{ii} , β_0 , β_i , β_{ij} , β_{ii} are the regression coefficients, X_i represent the variables of the system.

4 Results and discussion

4.1 FACTORS CHOSEN IN PBD

Many factors can affect the final results in preparation condition of γ -PGA/CS nanocapsule. In this research, 9 factors were chosen to estimate the relative variables. As polyelectrolytes γ -PGA and CS were different ionization in different pH value, so the pH value of solution might have an impact on the size of the nanocapsule. In the design factors A to D were the pH value and mass concentration of the two materials. In our previous work, we found that flocculation occurred between γ -PGA and CS if the concentration and the volume of two materials were not suitable. Therefore, except for the reaction conditions such as stirring speed, concentration of Mg^{2+} , reaction time after dropping and dropping speed the volume of γ -PGA was one of the factors. At the same time, the total volume of each sample was 20ml.

4.2 PBD RESULTS

To evaluate the quality of the model, an F-value test was conducted. The ANOVA of PBD is presented in Table 5. There is only a 1.43% chance that a "Model F-Value" this large could occur due to noise. Therefore, the model is significant and its R-Squared is 0.9645, that has to say the results are suitable for the experiment design and the further optimization. In this case D (Mass concentration

of γ -PGA) and J (Volume of γ -PGA) are significant model terms, and B (pH value of γ -PGA), E (Stirring speed), F (Concentration of Mg^{2+}), H (Dropping speed) are insignificant model terms. We chose pH value of CS, Mass concentration of γ -PGA and Volume of γ -PGA for further optimization. In the next experiments variables E and F took intermediate values, C, G and H took the high levels, while pH value of γ -PGA was natural value.

TABLE 5 The ANOVA of PBD

Source	Sum of Squares	df	Mean Square	F value	p-value Prob>F
Model	683387.5	9	75931.95	12.083580	0.0143
A	43224	1	43224.00	6.878539	0.0586
B	12818.4	1	12818.40	2.039882	0.2264
C	35730.25	1	35730.25	5.686006	0.0756
D	400332.3	1	400332.30	63.707680	0.0013
E	4531.853	1	4531.85	0.721186	0.4436
F	17495.6	1	17495.60	2.784198	0.1705
G	36212.05	1	36212.05	5.762678	0.0743
H	13682.25	1	13682.25	2.177353	0.2141
J	119360.9	1	119360.90	18.994730	0.0121
Curvature	73584.02	1	73584.02	11.709940	0.0267
Residual	25135.57	4	6283.89		
Lack of Fit	25111.07	2	12555.54	1024.942	0.0010
Pure Error	24.5	2	12.25		
Cor Total	782107.1	14			

4.3 ANOVA OF BBD

Based on the PBD, three factors (pH value of CS, Mass concentration of γ -PGA and volume of γ -PGA) significantly influenced the Z-Ave of the γ -PGA/CS nanocapsule. To define the optimum settings of these factor levels, a BBD with 15 experiments (Table 4) was used to estimate the model coefficients. The experimental points are located in the middle of a cube's edges (12 experiments, which used to factorial analysis) and at the centre of the cube. This ensured that independent estimates of the model's parameters were obtained. Regression analysis was used to estimate the regression coefficients of the model, each response can be described by a second-order empirical model, which is adequate for predicting the response in the experimental region.

A statistical test of the model fit was performed by comparing the variance due to the lack of fit with the pure error variance using the F-test. The Z-Ave analysis process of BBD is presented in Table 6. The Model F-value of 6.55 implies the model is significant. There is only a 2.61% chance that a "Model F-Value" this large could occur due to noise. In this case X3, X1* X2 and X1² are significant model terms. The "Lack of Fit F-value" of 2.73 implies the Lack of Fit is not significant relative to the pure error. There is a 27.91% chance that a "Lack of Fit F-value" this large could occur due to noise.

Three factors affect the Z-Ave of γ -PGA/CS nanocapsule in this order: $X_3 > X_2 > X_1$, which matches with the result of PBD.

TABLE 6 The Z-Ave ANOVA of BBD

Source	SS	df	MS	F value	p-value Prob>F
Model	37901.13	9	4211.24	6.55	0.0261
X1	2227.78	1	2227.78	3.46	0.1218
X2	3633.78	1	3633.78	5.65	0.0634
X3	16744.50	1	16744.50	26.04	0.0038
X1* X2	8281.00	1	8281.00	12.88	0.0157
X1* X2	855.56	1	855.56	1.33	0.3009
X2* X3	517.56	1	517.56	0.80	0.4107
X1^2	4941.56	1	4941.56	7.68	0.0393
X2^2	189.64	1	189.64	0.29	0.6104
X3^2	196.31	1	196.31	0.31	0.6044
Residual	3215.10	5	643.02		
Lack of Fit	2584.94	3	861.65	2.73	0.2791
Pure Error	630.17	2	315.08		
Cor Total	41116.23	14			

The quadratic model obtained by regression analysis showed as follow:

$$Y_1 = 285.83 - 16.69 * X_1 + 21.31 * X_2 + 45.75 * X_3 + 45.50 * X_1 * X_2 + 14.63 * X_1 * X_3 + 11.38 * X_2 * X_3 + 36.58 * X_1 * X_1 - 7.17 * X_2 * X_2 - 7.29 * X_3 * X_3$$

The R-squared and Adj R-squared of this model are 0.9218 and 0.7811, respectively. Therefore, there are insignificant terms in the model, this match with the ANOVA results. The C.V. % of 8.52 show that the model variation of the measured values is small; there is only 8.52% variability of the data. Moreover, Adeq Precision measures the signal-to-noise ratio, which value greater than 4 is desirable. The Adeq Precision of 7.230 in this model indicates an adequate signal (17).

The fit testing of regression model was shown in Figure 3. The Figure 3(a) showed that the normal plot of residuals was almost a straight line, indicating the potential of error distribution is normal. Predicted values and residuals showed irregular distribution as shown in the Figure 3(b), this matched with the residuals should be amorphous in the model fitting. Thus, the regression assumptions were reasonable, and fitting the regression model was appropriate. In sum, the model can be used to navigate the design space and has a certain degree of predictability.

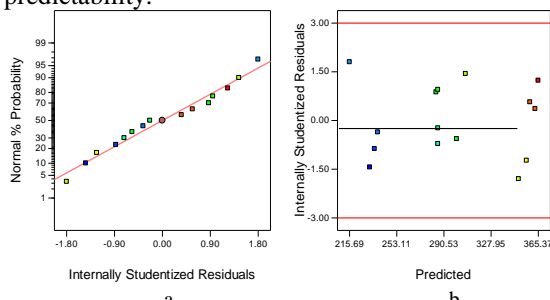


FIGURE 3 Regression model fit testing: a) Normal Plot of Residuals; b) Residuals vs. Predicted

In the PDI analysis process, the quadratic model F-value of 3.58 and p-value Prob>F of 0.0868 imply the model is insignificant. Therefore, we chose the modified quadratic model in which the insignificant item X2 was removed. The PDI analysis process of BBD is presented in Table 6. The model F-value of 4.84 and p-value Prob>F of 0.0353 imply the model is significant. In this case, there are many insignificant model terms and only X1* X3 is significant model terms. The modified quadratic model is as follow:

$$Y_2 = 0.22 - 0.019 * X_1 - 0.023 * X_3 + 0.048 * X_1 * X_2 + 0.098 * X_1 * X_3 + 0.019 * X_2 * X_3 + 0.031 * X_1 * X_1 - 0.014 * X_2 * X_2 + 0.050 * X_3 * X_3$$

The R-squared and Adj R-squared of this model are 0.8657 and 0.6867, respectively. At the same time, the C.V. % and Adeq Precision are 16.60 and 8.247. Based on the data process of PDI, we can conclude that second-order model is not most adequate for describing the preparation conditions of γ -PGA/CS nanocapsule.

TABLE 7 PDI ANOVA of BBD

Source	SS	df	MS	F value	p-value Prob>F
Model	0.069833	8	0.008729	4.835171	0.0353
X1	0.002984	1	0.002984	1.652750	0.2460
X3	0.004163	1	0.004163	2.306088	0.1797
X1* X2	0.00912	1	0.009120	5.051808	0.0657
X1* X3	0.038514	1	0.038514	21.333370	0.0036
X2* X3	0.001521	1	0.001521	0.842499	0.3941
X1^2	0.003563	1	0.003563	1.973380	0.2097
X2^2	0.000704	1	0.000704	0.390196	0.5552
X3^2	0.009347	1	0.009347	5.177138	0.0632
Residual	0.010832	6	0.001805		
Lack of Fit	0.010322	4	0.002580	10.109270	0.0920
Pure Error	0.000511	2	0.000255		
Cor Total	0.080665	14			

4.4 OPTIMIZING RESULTS BY RSM

The effect of three factors on the size of the γ -PGA/CS nanocapsules was further analysed using contour and 3D response surface plots, which were the graphical representations of the regression model. By simulating the experimental results using the empirical model, these plots (Figure 4) efficiently identified the optimum values of the variables.

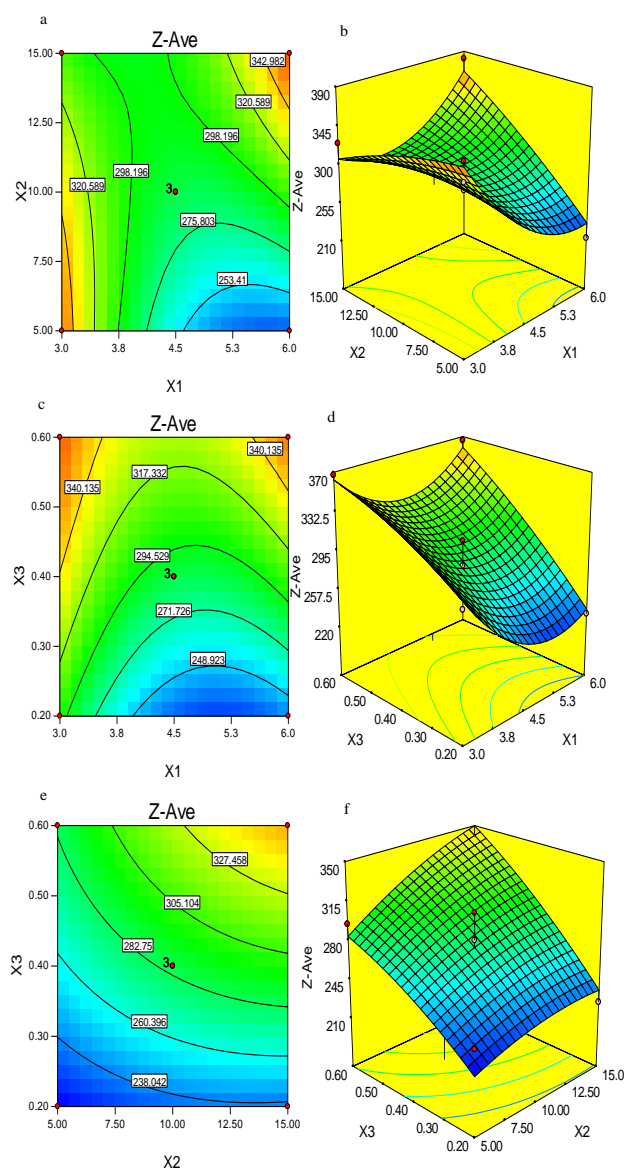


FIGURE 4 The contour and 3D response surface plots of two factors with the other factor at central levels; X1, X2, X3 represent pH value of CS, Volume of γ -PGA and Mass concentration of γ -PGA, respectively

Then it was convenient to understand the interactions between any two factors. From the Figure 4(a, b), we can know that pH value of CS and volume of γ -PGA had significant interaction. The Figure 4(c, d) showed pH value of CS and mass concentration of γ -PGA also had interaction. However, the interaction between volume of γ -PGA and mass concentration of γ -PGA was not significant (Figure 4(e, f)), the reason of that might be as follow: both of the two factors were affecting the total amount of γ -PGA.

References

- [1] Akagi T, Watanabe K, Kim H, Akashi M 2010 *Langmuir* **26** 2406–13
- [2] Bhatt R, Vries P, Klein JP, Tulinsky J, Lewis RA, Jack W, Singer J W 2007 *Polyglutamic acid-camptothecin conjugates and methods of preparation* USP: 7173041

4.5 VERIFICATION EXPERIMENT

After optimization, one set of experimental conditions with 4.0 of X1, 18ml of X2 and 0.1 of X3 was used as experimental conditions. The experimental results were listed in Table 7; the reproducibility and stability were very good (see Table 8). The size distribution of No.1 was showed in Figure 5.

TABLE 8 Verification experiment results

No.	1	2	3
Z-Ave/nm	177.5	174.4	173.1
PDI	0.169	0.130	0.150

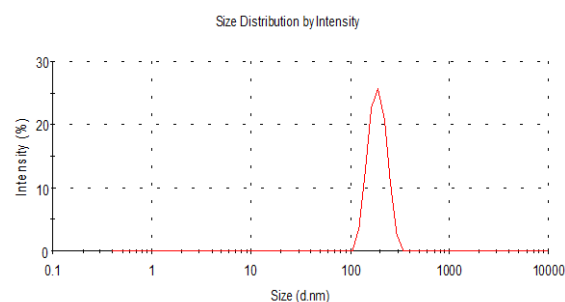


FIGURE 5 The size distribution of No.1

5 Conclusions

In this work, nanocapsules were successfully self-assembled from the γ -PGA and CS. PBD and BBD were used to optimize the preparation conditions. The optimized preparation conditions were that the pH value of CS was 4.0, the pH value of γ -PGA was natural pH value, the mass concentration of CS was 1g/L, the mass concentration of γ -PGA was 0.1g/L, the stirring speed was 100r/min, the concentration of Mg^{2+} was 0.004 mol/L, the reaction time after dropping was 60min, the dropping speed was 0.55 mL/h and the volume of γ -PGA was 18mL, in which the size of nanocapsule was 175 nm and its PDI was 0.15, it can be used as drug carrier or flavour carrier.

Acknowledgment

The authors gratefully acknowledge the financial support provided by Science and Technology Commission of Shanghai, China (No. 10540501100& No.LM201343) and by the Open Project Program of State Key Laboratory of Dairy Biotechnology, Bright Dairy & Food Co. Ltd., (No. SKLDB2013-09).

- [3] Chang C H, Lin Y H, Yeh C L, Chen Y C, Chiou SF, Hsu Y M, Chen YS, Wang C C 2010 *Biomacromolecules* **11** 133–42
- [4] Ljubimova J Y, Black K L, Ljubimov A V, Eggehard H 2008 *Multifunctional Pharmaceutical Nanocarriers* **4** 233–62
- [5] Keresztessy Z, Bodnár M, Ber E, Hajdu I, Zhang M, Hartmann JF, Minko F, Borbély, J 2009 *Colloid & Polymer Science* **287** 759–65

- [6] Kurtaran H, Erzurumlu T 2006 *The International Journal of Advanced Manufacturing Technology* **27** 468–72
- [7] Mi F L, Wu Y Y, Lin Y H, Sonaje K, Ho Y C, Chen C T, Juang J H, Sung H W 2008 *Bioconjugate Chemistry* **19** 1248–55
- [8] Montgomery D C 2009 *Design and analysis of experiments*. Beijing, China: Posts Telecom Press
- [9] Prencipe G, Tabakman S M, Welsher K, Liu Z, Goodwin A P, Zhang L, Henry J, Dai H J 2009 *Journal of the American Chemical Society* **131** 4783–7
- [10] Sonaje K, Chen Y J, Chen H L, Wey S P, Juang J H, Nguyen H N, Hsu C W, Lin K J, Sung H W 2010 *Biomaterials* **31** 3384–94
- [11] Tong R, Christian D A, Tang L, Cabral H, Baker J R Jr, Kataoka K, Discher D, Cheng J 2009 *Nanopolymeric therapeutics MRS Bulletin* **34** 422–31
- [12] Tsao C T, Chang C H, Lin Y Y, Wu M F, Young T H, Han J L, Hsieh K H 2011 *Carbohydrate Polymers* **84** 812–9
- [13] Wadhwa S, Mumper R J 2010 *Molecular pharmaceuticals* **7** 854–62
- [14] Xu X H, He M Z 2010 *Experimental design and application of design-expert SPSS*. Beijing, China: Science Press 36-59
- [15] Yan S F, Zhu J, Wang Z C, Yin J 2011 *European Journal of Pharmaceutics and Biopharmaceutics* **78** 336–45
- [16] Yao J, Xu H, Wang J 2007 *J Biomater Sci Polymer Edn* **18** 193–204
- [17] Zhang H, Jiang Z, Guo C 2009 *The International Journal of Advanced Manufacturing Technology* **41** 110–21

Authors	
	<p>Ma Xia, born in May 29, 1971, Shandong, China</p> <p>Current position: Doctor of Fermentation Engineering, professor and master's supervisor in Shanghai institute of technology</p> <p>University studies: Fermentation engineering in Tianjin University of Science & technology</p> <p>Scientific interest: Biomaterials biosynthesis & applications fermentation of function food, drug & fragrance controlled release materials</p> <p>Publications: 10 Patents, 40 Papers</p> <p>Experience: Ma Xia, is professor, advisor of graduates. Her Ph.D was obtained from Tianjin University of Science & Technology in 2003. Post-doctor was finished in College of Chemistry and Bioengineering, Donghua University in 2008. Her research area focuses on new biologic material, fermentation technology, functional food, and flavour. She has taken charge in several scientific projects. About 30 papers were published. 4 natural invention patents were applied. 2 third progress prize of Shandong Science and Technology and 2 second progress prize of Shandong Science and Technology were gained.</p>
	<p>Zhang MianMian, born in August 5, 1988, Henan, China</p> <p>Current position: A graduate student in Shanghai institute of technology</p>
	<p>Chen ShiWen, born in January 28, 1971, Shandong, China</p> <p>Current position: Doctor of Neurosurgery, associate professor and master's supervisor in Shanghai JiaoTong University Affiliated Sixth People's Hospital</p> <p>University studies: Clinical medicine in The Second Military Medical University</p> <p>Scientific interest: Biologic material being used on the medical aspect, Brain injury and Cerebrovascular disease</p> <p>Publications: 40 Papers</p> <p>Experience: Chen ShiWen, is an associate professor. His M.D was obtained from The Second Military Medical University in 2005. Post-doctor was finished in Medical College of Shanghai JiaoTong University in 2011. His research area focuses on biologic material being used on the medical aspect, brain injury and cerebrovascular disease. He has taken part in several scientific projects. More than 40 papers were published.</p>



Authors' index

Ai Yue Wei	271	Kozachenko Yurii	7	Wang Xianming	249
Amirgaliyev Beibut	193	Lei Li	283	Wang Xinyuan	267
Bai Bing	63	Lei Lianghai	237	Wang Yang	271
Bi Jianxin	237	Lei Yanrui	283	Wang Zheng	204
Chao Chih-Feng	77	Li Chenghua	33	Wang Zhi Jian	271
Chen Guangxue	133	Li Dong-ping	254	Wei Jian Ping	44
Chen Longmiao	93	Li Hongkai	183	Wen Zhi Hui	44
Chen Shi Wen	293	Li Ming Ming	88	Wu Li	267
Chen Suying	52	Li Shilong	209	Wu Shaohua	72
Chu LiLi	144	Li Yinglin	52	Wu Y P	215
Dai Shao Hua	44	Lin Gui	93	Xing Bao	231
Dai Zhendong	183	Lin Rui-Lin	179	Xing Tingyan	170
Deng Xiuquan	63	Liu Laiquan	283	Xu Bing	159
Fan X G	215	Liu Rufeif	188	Xu Lei	52
Fang Bo	88	Liu Wanli	224	Yang Lianhe	52
Fei Song	198	Liu Zhaohua	174	Yang Shiqing	84
Fu Qiang	93	Lu Xiushan	188	Yang Yuxia	174
Gao Dehua	63	Ma Xia	293	Yedilkhan Didar	193
Gao Guohui	49	Ma Xing-bin	68	Yin Zhigang	49
Gao Qinhe	124	Miao Lanfang	84	Yu Chuanqiang	124
Gao Quan Chen	58	Miao Qian	124	Yu Yang	276
Geng Jia	288	Miao Zhenjiang	204	Yue Binbin	164
Gu Qiong	249	Ou G Zh	215	Zhang De Feng	151
Guo Xuhong	72	Pan Xin	98	Zhang Fu	164
Guo Zhiqiong	63	Pashko Anatolii	7	Zhang Guoying	164
Hagan Paul	13	Qin LiLi	144	Zhang Hong Tu	44
He Wan	258	Shi Yongfang	104	Zhang Hua	33
He Yan	271	Su C Q	26	Zhang Junxia	39
Hong Ming-Huwi	77	Su Junqing	13	Zhang Mian Mian	293
Hou Hongqing	124	Sun Hongbin	98	Zhang Q X	215
Huang Q M	26	Sun Wei	276	Zhang Shang	170
Huang Wen Hu	88	Sun Wenlei	104	Zhang Xin	111
Huang Xiao-dan	115	Tan Li Jun	88	Zhang Yakun	164
Huang Xing	72	Tang Ye	88	Zhao Hongxing	39
Huang Yanhua	104	Tao Feng	21	Zhao Lei	133
Jia Zhenzhen	21	Tian Hongyan	209	Zhe Cui	198
Jiang Cui-pin	68	Tian Maoyi	188	Zheng Wu	249
Jiang Hong	104	Tu Junxiang	263	Zhou J H	215
Jiang Xiaoping	33	Wang G	26	Zhou Ling	151
Kairanbay Magzhan	193	Wang H Y	26	Zhu Chang Zhi	58
Kenshimov Chingiz	193	Wang Shuren	13		
Kong Xiangsheng	254	Wang Wei	115		

Cumulative Index

Mathematical and Computer Modelling

Yurii Kozachenko, Anatolii Pashko Accuracy of Simulations of the Gaussian random processes with continuous spectrum

Computer Modelling & New Technologies 2014 18(3) 7-12

This paper investigates algorithms for the construction of sub-Gaussian models for the Gaussian stationary random processes with continuous spectrum. Estimates for random processes with analytical correlation functions retrieved and improved existing ones. Algorithms for simulation of random processes with given accuracy and reliability in various function spaces were constructed.

Keywords: Gaussian process, simulation, sub-Gaussian model, model accuracy, model reliability

Shuren Wang, Junqing Su, Paul Hagan Energy dissipation characteristics of sandstone cutting under mechanical impact load

Computer Modelling & New Technologies 2014 18(3) 13-20

Based on the sandstone experiment by using the linear rock cutting machine at the school of mining engineering, University of New South Wales (UNSW), the computational model was built by using particle flow code (PFC), and the energy dissipation characteristics was studied considering these parameters such as cutting speed, cutting depth, rock strength, rock temperature and rock confining pressure. The results showed that the specific energy and acoustic emission of sandstone were proportional to the cutting speed and rock strength, there was an optimal depth of rock cutting with the cutting depth increasing under the confining pressure condition, and the specific energy and acoustic emission of sandstone changed obviously due to the influence of the rock temperature and rock confining pressure.

Keywords: Rock Cutting, Energy Dissipation, Specific Energy, Acoustic Emission, Particle Flow Code

Zhenzhen Jia, Feng Tao Gas explosion characteristics and its control technologies in closed fire zone

Computer Modelling & New Technologies 2014 18(3) 21-25

The closure measures of fire zones are taken after gas explosion in the working face, which can bring two problems: whether closure measures will lead to a secondary gas explosion in closed fire zones or not, and what will be the rough time interval between taking measures and gas explosion occurrence. To solve these problems, gas accumulation characteristics, oxygen concentration characteristics and fire sources in the fire closure process were analysed, and then the characteristics and rules of gas explosion were acquired, in addition, the pressure change and gas accumulation model in closure zones under three conditions (the temporary closure wall only in air inlet laneway, only in air return laneway, or both in air inlet laneway and air return laneway) were obtained. Finally, the measures and technologies to prevent and control gas explosion were introduced in the fire closure process of working face.

Keywords: Closed Fire Zone, Gas Explosion Characteristics, Gas Accumulation Model, Fire Sources, Control Technologies

G Wang, H Y Wang, Q M Huang, C Q Su Numerical modelling of rock cross-cut coal uncovering based on ANSYS

Computer Modelling & New Technologies 2014 18(3) 26-32

Outbursts of coal and gas could be induced by rock cross-cut coal uncovering. ANSYS is used to numerically simulate the stress, strain and energy of surrounding rock of roadway during the process of rock cross-cut coal uncovering. Modelling results show that there is a banding tension stress zone in roof and floor of roadway after excavation. Principal and shear stress concentration are formed in the upper and under area of the anterior heading face, which is symmetrically distributed between medial axis of roadway, while stress, strain and strain energy of overlying strata above the coal seam approximately keep invariant. The occurrence of the stress concentration in upper area of the anterior heading face could contribute to the instability and failure of coal and rock mass. The area with weak destruction-resisting ability is the most easily to be the releasing port of outburst and should be regarded as the key region in outburst prevention.

Keywords: Rock Cross-cut Coal Uncovering, Numerical Modelling, ANSYS, Stress Distribution

Hua Zhang, Xiaoping Jiang, Chenghua Li UKF-based underground intrusion localization algorithm for optical-fibre sensing perimeter protection

Computer Modelling & New Technologies 2014 18(3) 33-38

To improve the precision of the underground intrusion localization in the optical-fibre sensing perimeter protection application, an intrusion localization algorithm based on the Unscented Kalman Filter (UKF) is presented. The geometrical relationships of the sensors and the intruder are analysed and the state equation and the measurement model are deduced. Then the UKF algorithm is used to estimate and track the location of the intruder. The simulations demonstrate that the algorithm improves the intrusion localization precision and the intruder can be tracked even if not enough sensors detect the intrusion signal.

Keywords: Optical-Fibre Sensor, Underground Intrusion Detection, State Estimation, Unscented Kalman Filter

Junxia Zhang, Hongxing Zhao A novel compressed air solar energy photo-thermal generating electricity system

Computer Modelling & New Technologies 2014 18(3) 39-43

On the basis of comparing the solar photovoltaic and photo-thermal generating electricity advantages and disadvantages, to overcome the phase-change losses caused by water evaporated into vapour, a compressed air solar energy photo-thermal generating electricity system was proposed in the present work. Air was compressed with compressor and heated with solar heater so as to get high temperature and pressure. High-temperature and high pressure air passes into turbine to generate electricity. The entire design is simple and compact, safe and reliability, energy saving and environmental protection. Thermodynamic cycle analysis was carried out. It comes to conclusion that practical efficiency depends on pressure ratio, the compressor and turbine efficiencies and solar photo-thermal conversion efficiency.

Keywords: Compressed air, solar energy, photo-thermal, generating electricity

Zhi Hui Wen, Jian Ping Wei, Hong Tu Zhang, Shao Hua Dai An experimental study on coal gas desorption laws with different particle size

Computer Modelling & New Technologies 2014 18(3) 44-48

Based on the self-designed coal containing methane gas desorption law experimental system and coal gas desorption kinetics law, the gas desorption law for different particle size coal under isothermal-isobaric condition are measured by combining combined simulation measurement with the theoretical analysis. The effect of particle size on coal gas desorption laws are obtained through fitting analysis on experimental data: 1. Under the same sorption and desorption conditions, coal particle size was inversely related to total desorption gas amount in the same period. 2. For coal with any particle size, the gas desorption amount firstly increased as time, and finally reached the maximum value, which was equal to the gas sorption amount unit per mass. 3. For coal with higher gas and coal outburst risk trend, the effect of particle size on initial desorption velocity and total desorption amount with time was less. 4. Within the limit particle size, the gas initial desorption velocity had a negative relationship with particle size. Finally we theoretically analysed the effect of particle size on coal gas desorption laws.

Keywords: Particle size, Gas desorption, Gas adsorption, Pore

Zhigang Yin, Guohui Gao Application of fractal theory to dam deformation forecast

Computer Modelling & New Technologies 2014 18(3) 49-51

Based on the safety observation data of dam, the establishment of the prediction model of dam deformation is very important for safe operation of the reservoir. The early deformation curve of the dam has self-similar fractal feature. The fractal interpolation function can be applied to not only processing the dam prototype observation data but also forecasting the rule of early dam deformation. In this paper, the reservoir dam deformation can be analysed and predicted by the fractal interpolation function. Analysis shows that, the method for dam deformation prediction of maximum error is 8.0%. Therefore, regarding the half-year short-term forecast, the forecast result obtained from fractal interpolation function method could be reliable.

Keywords: dam, the fractal interpolation function, forecast

Yinglin Li, Lianhe Yang, Suying Chen, Lei Xu Three dimensional simulation of weft knitted fabric based on surface model

Computer Modeling & New Technologies 2014 18(3) 52-57

In order to enhance the visual simulation effect of weft knitted fabric in weft knitting CAD system, a three dimensional surface model of weft knitted structures is developed and realized on the computer screen. It is assumed that the yarn centreline is continuous when it is stressed to bend, and the cross-section of yarn is circular and uniform. Three dimensional geometric models are created by using second order continuous NURBS curves. According to the trend of the yarn centreline, a surface model of the yarn is created by sweeping approach. With developing tools of Visual C++ and OpenGL graphics library, the surface model of loop is achieved by the method of mesh chips on computer screen, and the three dimensional simulation of weft knitted fabric is realized. The simulation effect based on surface model can be seen from the simulation results of basic weft knitted fabric.

Keywords: Weft knitting, Loop structure, NURBS; Simulation, Surface model

Chang Zhi Zhu, Quan Chen Gao Three-dimensional computer simulation of soil nailing support in deep foundation pit

Computer Modelling & New Technologies 2014 18(3) 58-62

The computer application program that is applied based on the finite difference method. By taking the soil nailing support structure in Shijiazhuang city as an example, the three-dimensional computer numerical model of deep foundation pit is set up; the horizontal displacement and the ground settlement of the deep foundation pit are simulated in the process of excavation and support. The simulation result is consistent with the test result. The results show that the deformation behaviour of the deep foundation pit can be analysed by using three-dimensional computer simulation technology in actual project. The method overcomes the deficiency of theoretical analysis method and offers effective guidance for design and construction of foundation pit excavation and support.

Keywords: Three-dimensional computer simulation, Soil nailing support, Deep foundation pit

Bing Bai, Xiuquan Deng, Zhiqiong Guo, Dehua Gao Organizations learning mechanism in the cyber society ecology system: an agent-based simulation

Computer Modelling & New Technologies 2014 18(3) 63-67

In this paper, we try to focus on the cyber society ecology system, which is a naturally occurring complex system to a certain stage of networks economic development. Based on the multi-agent simulation methodology, this paper analyses adaptive learning among organizations of cyber society ecology system, and then designs interaction rules of agents and simulation parameters, and finally the simulation results are analysed.

Keywords: Cyber Society Ecology System, Adaptive Learning, Agent-Based Simulation

Xing-bin Ma, Cui-pin Jiang The approach of fixed asset management based on the shortest path

Computer Modelling & New Technologies 2014 18(3) 68-71

We often meet with shortest path problem in National Undergraduate Mathematical Contest and practical life. The definition of shortest path problem is introduced, Dijkstra algorithm and 0-1 Programming Method to solve the shortest path problem are given. A practical problem is given and is calculated by these two methods.

Keywords: mathematical modelling, shortest path problem, Dijkstra algorithm, 0-1 Programming

Xuhong Guo, Xing Huang, Shaohua Wu Correlation between austempering parameters and hardness of austempered ductile iron based on artificial neural network

Computer Modelling & New Technologies 2014 18(3) 72-76

Mechanical properties of ADI mainly depend on the austempering parameters, which include austenitizing temperature and time, austempering time and temperature, apart from chemical composition, alloyed elements and casting parameters. In this paper, an investigation has been conducted on the prediction model of mechanical properties of ADI between austenitizing temperature and time, austempering temperature and time as inputs and Vickers hardness of samples after austempering as the outputs based on artificial neural network. There are two types of data of the model: training data and testing data. The former data come from the published literature and 12 experimental data used for network testing. The research results of the model shows that the predicted values approach to the measured data in most of the testing samples and the maximum margin of error between experimental and predicted data is 4.682%.

Keywords: Austempered Ductile Iron, Artificial Neural Network, Austempering Parameters, Mechanical Properties, Prediction

Model

Chih-Feng Chao, Ming-Huwi Horng Firefly algorithm for training the radial basis function network in ultrasonic supraspinatus image classification*Computer Modelling & New Technologies 2014 18(3) 77-83*

The physicians observed the echo-texture and the shape changes of supraspinatus to decide the severity of rotator cuff disease in the clinical standard ultrasound examination. It is not reliable because the accuracy of visual observation depends on the experience of physicians. This article proposes a new algorithm called **Firefly RBF** network to training the radial basis function neural network by applying the firefly algorithm for classifying the different supraspinatus disease groups that are normal, tendon inflammation, calcific tendonitis and tendon tears of the ultrasound supraspinatus images based on the texture analysis technology. The texture features are generated from four methods those are the grey-level co-occurrence matrix, the texture spectrum, the fractal dimension and the texture feature coding method to analyse the tissue characteristic of supraspinatus. The F-score measurement are used to select powerful features those are generated from the four texture analysis methods for comparison in the training stage, meanwhile, the proposed Firefly RBF network is used to discriminate test images into one of the four disease groups in the classification stage. Experimental results showed that the percentage of correct classification was more than 93.7% that is superior to other methods in the classification of ultrasonic supraspinatus images.

Keywords: Radial basis function network, Firefly algorithm, Ultrasonic supraspinatus image, Texture analysis

Shiqing Yang, Lanfang Miao Material simulation based on Phong illumination model*Computer Modelling & New Technologies 2014 18(3) 84-87*

Realistic material simulation is one of the major works in rendering realistic graphics. In this paper, we have studied and discussed the effects of rendering realistic graphics by simulating some materials under different illumination conditions and material attributes and emission based on the OpenGL graphics technology and Phong illumination model. Experimental results show effects of emission on specular and non-specular materials and demonstrate the emission effects to some degree.

Keywords: Phong illumination model, Realistic graphics, Material simulation, Emission

Li Jun Tan, Bo Fang, Ming Ming Li, Ye Tang, Wen Hu Huang Parametric identification for GHM and application of viscoelastic damper*Computer Modelling & New Technologies 2014 18(3) 88-92*

The GHM (Golla-Hughes-McTavish) model is extensively utilized by structural designers for studying complex structures with viscoelastic damping treatments in engineering. A kind of shear-type viscoelastic damper is investigated, and the damper is modelled with GHM model. The parameters of GHM model are identified by curve fitting and a detailed experiment in complex frequency domain. The comparison results show that the method proposed in the present paper to determine the parameters of GHM model is correct. A whole-spacecraft vibration isolation experiment is practically performed, and the results show that using the method to design the WSVI (Whole-spacecraft Vibration Isolator) is effective for isolating structure vibrations.

Keywords: viscoelastic damper, GHM model, vibration experiment, WSVI

Longmiao Chen, Qiang Fu, Gui Lin Study on the sealing properties of the sealing structure for the rotating chamber of a certain cased telescoped ammunition gun*Computer Modelling & New Technologies 2014 18(3) 93-97*

To solve the spherical transient high pressure gas seal problem for the rotating chamber of a certain medium calibre cased telescoped ammunition gun, a self-impacted combined sealing system was newly designed which can be placed at semi combustible cartridge of the cased telescoped ammunition. The sealing mechanism of the structure was analysed and simulation studies on the comprehensive properties of the sealing structure were carried out via the FEM dynamic response. In addition, the simulation and verification tests were conducted to test the sealing performance of the sealing structure. The results of the simulation analysis and the experiments demonstrate that the designed sealing structure has a good sealing performance and can solve the spherical transient high pressure gas seal problem for the rotating chamber of the medium calibre cased telescoped ammunition gun, and it is expected to offer a reference value to solve related problems in engineering.

Keywords: Sealing for the Rotating Chamber, FEM, Test Verification, Cased Telescoped Ammunition

Information and Computer Technologies

Xin Pan, Hongbin Sun A self-adaptive selective method of remote sensing image classification algorithms
Computer Modelling & New Technologies 2014 18(3) 98-103

Remote sensing image classification algorithms, which can obtain information of land use\cover quickly and inexpensively have been widely used in the field of GIS. The quality of classification results is not only affected by the quality of remote sensing data, but also affected by the character of classification algorithm. At present, despite a lot of algorithms have been proposed, but users usually meet difficulties in algorithm selection due to single classification algorithm cat not applicable to all classification cases. This study proposes a self-adaptive selective method for remote sensing image classification algorithms based on data complexity evaluation, through data complexity evaluation, our method can distinguish remote sensing data's character even from same satellite sensor and give user recommendation of algorithm selection. Experiments indicate that the algorithms selected by this method can achieve higher classification accuracy, which provides the recommendation for the selection of appropriate classification models to users.

Keywords: Remote Sensing image, Classification, Algorithms Evaluate, Data Complexity

Hong Jiang, Wenlei Sun, Yongfang Shi, Yanhua Huang A study and implementation on the data reduction based on the curvature of point clouds
Computer Modelling & New Technologies 2014 18(3) 104-110

In the process of Reverse Engineering (RE), higher density of measured data from all kinds of parts with complex curved surface will not only lead to lower efficiency in computing, storing and data processing, but also affect the fairness of reconstructed surface. According to the advantages and disadvantages of common algorithms, an algorithm for data reduction is proposed in this paper, in which the neighbourhood search method based on the point cloud's curvature is used. With the utilization of proposed algorithm, high precision and the desired effect can be ensured. Finally, a roller bit's data cloud, as an example, is reduced efficiently and validly by the algorithm in this paper.

Keywords: Reverse Engineering, Data Reduction, Point Cloud, Curvature

Xin Zhang Multi-objective improved algorithm for flow allocations in hazardous chemicals logistics preference paths
Computer Modelling & New Technologies 2014 18(3) 111-114

The flow allocation of paths was a key stage of the transportation network's efficiency, particularly in the hazardous chemicals logistics network where many weights were stochastic. Over the years, a variety of methods (or heuristics) have been proposed to solve this complex optimization problem, with good results in some cases just with limitations in the special fields. In this work, we develop an algorithm for model multi-objective that combines ideas from stochastic weight. Our method performs well even when the order of magnitude and/or the range of the parameters were unknown. The method refines iteratively a sequence of parameter distributions through preference combined with partial exemplifying from a historical prior defined over the support of all previous iterations. We exemplify our method with multi-objective improved models using both simulated and real experimental data and estimate the weight efficiently even in the absence of a priori knowledge about the weight.

Keywords: Hazardous chemicals transportation, Flow allocation, Multi-objective optimization, Path preference, Control

Wei Wang, Xiao-dan Huang Spatially aware in implicit human robot interaction
Computer Modelling & New Technologies 2014 18(3) 115-123

Implicit interaction pattern between the human user and the robot is important for reducing cognitive burden and enhancing cooperation effect. Given that the spatially aware is a foundation for human-robot cooperation, for the existing researches of robots, in this paper, a reachable space for a serial robot arm with a fixed monocular vision system and five degrees of freedom (5-DoF) was built. Based on the link frame with D-H notation, analysis and simulating experiments were carried on to show the reachable space in three dimensions. In addition, multi ultrasonic sensors are used to detect the space realizing proximity controlled.

Keywords: Human Robot Cooperation, Spatially Aware, Reachable Space, Proximity Controlled, Implicit Pattern

Hongqing Hou, Qian Miao, Chuanqiang Yu, Qinhe Gao Study on T-S fuzzy sliding mode control based on a new reaching law

Computer Modelling & New Technologies 2014 18(3) 124-132

Fuzzy Sliding Mode controller based on fuzzy T-S Model is designed for the nonlinear, uncertainties and fast variable time characterizes question of bank-to-turn aircraft control and guidance system model. BTT control model was obtained by using T-S modelling method, asymptotically stable sliding surface was designed, and a new sliding mode reaching law is proposed. Based on the new reaching law, sliding mode stable tracking controller is designed. At last, the rationality and the effectiveness of the designed T-S fuzzy sliding mode stable tracking controller with the new reaching law are verified by the theoretical proof and the simulation experiments.

Keywords: Bank-to-turn Aircraft, Fuzzy Control, Sliding Mode Control, Reaching Law

Lei Zhao, Guangxue Chen A printer reverse characterization model based on BP neural network

Computer Modelling & New Technologies 2014 18(3) 133-143

For colour printer, there are very complicated nonlinear relation between its printed colour chromatic values and input digital image pixel values. In the research, data sets of printed colour chromatic values and their digital image pixel values are classified by hue angle range, the data in each hue angle range is taken as learning samples to create BP neural network. With improved combined method of additional momentum factor and variable learning rate, BP neural network of each hue angle range is trained and created. The experiment result shows that, with appropriate structure and classified learning samples, the reverse characterization model based on ten BP neural networks can be trained in relative short time; the colour errors between the experimental printed colour chromatic values and computed printed chromatic values are far less than the threshold of human eyes, i.e. the reverse characterization model achieves rather high accuracy.

Keywords: BP neural network, Hue angle range, Data classification, Colour management

LiLi Chu, LiLi Qin Communication technology in the application of the smart grid

Computer Modelling & New Technologies 2014 18(3) 144-150

By studying the significance of the smart grid, combined with a regional substation point location, line conditions, existing and future business development, existing communications equipment status, etc., the author initially sets the smart grid communications infrastructure deployment and network planning, in order to use the most reasonable communication technologies to support rapid development of smart grid. In a certain city with electric power communication network to the actual construction goal, we should complete the city power system communication network covering the whole deployment. At the same time combined with network energy efficiency project, we should analyse the already formed network, provide effective optimization model of energy efficiency and practical algorithm, and analyse its rationality through the simulation analysis, further improve the overall network in order to make it efficient to run.

Keywords: Smart grid, Data network, Network energy efficiency, Energy efficient routing strategy

Ling Zhou, De Feng Zhang A distributed multicast routing algorithm based on bone node set for mobile IP

Computer Modelling & New Technologies 2014 18(3) 151-158

Multicast routing is an important issue in network communication. In order to optimize the multicast routing cost and lessen the transmission delay for mobile IP communication, an idea of bone node set is introduced and the distributed multicast routing algorithm is designed based on the idea firstly. At the same time, the algorithm is implemented according to centre version and distributed version in detail, respectively. Then its necessary data structures, time complexity and message complexity are analysed in theories according to order of sequence for distributed operation. At last, simulation experiments are done in a 7×7 mesh topology and the results show that the designed algorithm can optimize the routing cost for multicast routing and reduce the transmission delay greatly compared to some same type algorithms. The distributed routing algorithm with the simple complexity can be efficiently used in large-scale mobile IP network.

Keywords: Distributed routing, mobile IP, bone node set, performance analysis, simulation

Bing Xu The study of campus network traffic monitoring platform

Computer Modelling & New Technologies 2014 18(3) 159-163

From the view of practical campus network traffic monitoring platform, one kind of solution based on the model of SNMP and NETFLOW network management frame was put forward to elaborate the designed overall structure of campus network traffic monitoring platform, data acquisition, traffic plotting and so on. Using Visual C++6 to design this platform, not only the key technology and methods for realizing the campus network traffic monitoring platform could be achieved, but also the network traffic monitoring and management should be completed. The implementation of this platform can efficiently monitor the network traffic.

Keywords: Campus traffic, Network traffic, Traffic monitoring, management platform, VC++6

Fu Zhang, Yakun Zhang, Binbin Yue, Guoying Zhang The gait analysis on the sloping walking of goat

Computer Modelling & New Technologies 2014 18(3) 164-169

The 18° sloping walking state and movement rules of the goat was researched by the high-speed video camera system. The movement process and the imaging results of goat in 18° sloping fields were recorded in the computer. The experiment imaging results of goat movement process were analysed by SigmaScan software and Matlab software, the results showed that gait parameters and angle change curve of each leg on 18° slope was obtained. The research will provide the basis of the experimental data for bionic design of agricultural machinery of goats sloping walking mechanism.

Keywords: Goat, High-speed camera, Sloping fields, Gait

Shang Zhang, Tingyan Xing RSSI Enhanced indoor LBS platform design

Computer Modelling & New Technologies 2014 18(3) 170-173

The fast development of WSN (Wireless Sensor Network) provide the solution to indoor localization application. To which, the position accuracy become the essential problem need to solve. This paper introduces in detail the composition of the whole system and the design of localization algorithm based on RSSI. Through computing the relevance of adjacent nodes of target tag, the enhanced localization method is introduced and reasonable system design prove the possibility to build a high accuracy indoor localization system based on WSN.

Keywords: WSN, RSSI, indoor localization, LBS

Zhaohua Liu, Yuxia Yang Study on semi-global matching algorithm extended for multi baseline matching and parallel processing method based on GPU

Computer Modelling & New Technologies 2014 18(3) 174-178

This paper extended semi-global matching algorithm into multi baseline matching to improve matching reliability, especially studies kernel function optimization strategies and GPU threads' executing scheme of matching cost cube computing and aggregating, and realized its fine granularity parallel processing based on GPU. The experiment results using three UCD aerial images based on Tesla C2050 GPU showed that MVLL's semi-global optimize algorithm can improve matching effectiveness and productiveness.

Keywords: semi-global matching, multi baseline matching, dynamic programming algorithm, GPU, parallel processing

Rui-Lin Lin Industrial product innovative design of toilet sensor timer

Computer Modelling & New Technologies 2014 18(3) 179-182

This study came up with an innovative product design for toilet safety timer, which can be installed in public lavatories or toilets at home. When an elderly person is in a toilet for a period of time longer than the time set in the device, a warning will be sent to the outside. The unique feature is the safety device of the sensor timer with the innovative structure containing an infrared sensor, a CPU, a timer, and a warning device. Users can set up the time on the timer. In case of an emergency such as users falling by accident or being in a coma, a warning will be sent out in time to get help to prevent tragedy from happening.

Keywords: Toilet Safety, Timer, Industrial Product

Hongkai Li, Zhendong Dai Mechanism design and flow estimation method of a hydraulic actuated robot

Computer Modelling & New Technologies 2014 18(3) 183-187

With the extension of robot applications, robot with high adaptability and high load capacity become a new focus in the recent years. Wheeled robots have the advantages of high load and speed, but this is limited in specified substrate. Legged robots inspired by the legged animals could move on rough terrain, so it was selected as a robot prototype for

the high adaptability and high load capacity robot. In this paper, the structure of a hydraulic actuated quadruped robot was proposed. And then the kinematics of single leg was analysed. To estimate the required flow, a trot gait with 50% duty cycle is schemed and the trajectories of feet were planned. Then the total flow of the system required was calculated with the planned motion. The result can be taken as a reference to optimize the robot mechanism and select the hydraulic system.

Keywords: hydraulic actuated, quadruped robot, mechanism design, flow estimation

Maoyi Tian, Rufei Liu, Xiushan Lu Ground point filtering method of vehicle-borne laser point cloud in urban street

Computer Modelling & New Technologies 2014 18(3) 188-192

Through the analysis of the spatial characteristics of vehicle-borne laser point cloud data in urban street, a method to extract ground points accurately from point cloud data is proposed. Firstly, three-dimensional virtual grid is used to organize point cloud. Secondly, the initial low ground point in a grid is extracted by level plane constraint (LPC) method, and then a multi-scale neighbourhood analysis (MSNA) method is taken to optimize the low ground points further. Finally, the ground points from original point cloud data are filtered based on the local slope. The experiment shows that this method can effectively extract the ground points.

Keywords: Vehicle-borne Laser, Urban Street, Ground Point Filtering, MSNA

Beibut Amirgaliyev, Magzhan Kairanbay, Chingiz Kenshimov, Didar Yedilkhan Development of automatic number plate recognition system

Computer Modelling & New Technologies 2014 18(3) 193-197

Today, the automatic number plate recognition (ANPR) system is a key aspect in traffic congestion. This will help minimizing the different kind of violations in the road. Advanced systems for tracking and fixing stolen, unauthorized vehicles are based on automated number plate recognition. This paper's main objectives is to review other methods and develop, at the same token evaluate our proposed approach. A very short review is performed on the various methods of number plate recognition systems. Further explanations of the proposed algorithm is illustrated in graphical forms to show how algorithm works. The paper is going to be concluded with test and evaluation results.

Keywords: ANPR, Plate area, Segmentation, OCR

Song Fei, Cui Zhe Study on HDFS improvement scheme based on the GE code and dynamic replication strategy

Computer Modelling & New Technologies 2014 18(3) 198-203

There is a lot of valuable information in the massive amounts of data. Any loss of data may result in a great loss. Data security cannot be ignored. There are varieties of data disaster recovery technologies. However, most of these techniques depend on the hardware devices or data redundancy greatly. This paper presents a distributed data disaster recovery technology that minimum dependence on data redundancy and hardware system redundancy. In addition, this technology has nothing to do with the user equipment and application data structures. The test proved that this new data disaster recovery method can not only enhance disaster recovery capabilities and reduce the redundancy of the system greatly, but also suitable for large-scale distributed data disaster recovery.

Keywords: data disaster tolerance, HDFS, GE code, dynamic replication

Zheng Wang, Zhenjiang Miao Design a media art installation based on fuzzy controlling system

Computer Modelling & New Technologies 2014 18(3) 204-208

As an art installation showed at Houtan station of Shanghai Metro Line7 and 2010 Shanghai World Expo Museum, "smart suspension ball" system displayed the rational sense of form and order of the controlling. The article focused on how to use appropriate fuzzy strategy to make the movements of the art installation more accurate under detailed experimental data. Another point of the article is to consider how to make this art installation to be a product with network, being combined and modular after upgrading the hardware and software of the installation. The performance of this upgraded product will bring more beautiful visual effects of controlling and technique. It will also be a successful case of integrating of science and technology into product designing for development of creative industries in China.

Keywords: Digital Media Art, Fuzzy Controlling, Installation Design, Arduino

Operation research and decision making

Shilong Li, Hongyan Tian Decision-making model of the urban regeneration construction project based on environment improvement

Computer Modelling & New Technologies 2014 18(3) 209-214

Urban regeneration is regarded as a more deliberate and harmonious progress of development, which takes a series of more sensible and multiple regeneration ways to achieve the goals, such as protection, repair, reuse and redevelopment. It is suggested that urban regeneration has positive effects on urban development and social economic system. Improvements in urban systems mainly originate from the environmental improvement, including ecological environment, social environment and the neighbourhood environment. Urban regeneration construction project and its decision-making model become the important research contents of urban regeneration. This paper discusses the urban regeneration from angle of project, and suggests a mathematical decision-making model of the urban regeneration construction project, which considers the risk, cost and environment constraints. Meanwhile, it is suggested that the joint exploitation is an important selection criterion. A case simulation is suggested in this paper in order to test the strategy model.

Keywords: Urban Regeneration, Decision-making Model, Project

Q X Zhang, Y P Wu, G Zh Ou, X G Fan, J H Zhou Displacement prediction of liangshuijing landslide based on time series additive model

Computer Modelling & New Technologies 2014 18(3) 215-223

The evolution of landslide displacement is affected by many factors. This paper studied the displacement monitoring data of Liangshuijing Landslide with Factor Analysis Method and found that the dominant factors influencing landslide displacement were in decreasing sequence: cumulative rainfall of anterior two months > rainfall of current month > the average reservoir level of current month > reservoir level fluctuation of current month. The paper selected three typical GPS monitoring points (ZJC09, ZJC11, ZJC13) of Liangshuijing Landslide to forecast their displacements by adopting the time series additive model on basis of the conclusion of previous factor analysis. The accumulative displacement of Liangshuijing Landslide can be divided into trend term and random term. The polynomial fitting was used for trend term displacement prediction. BP neural network model was used for the random displacement prediction. The final calculation results indicated that combination of factor analysis method and time series additive model could generate a reasonable and accurate prediction of landslide.

Keywords: Displacement prediction, Time series, Liangshuijing Landslide, Factor analysis, BP neural network

Liu Wanli Study on the decision value of analysts' recommendations

Computer Modelling & New Technologies 2014 18(3) 224-230

This paper documents a relationship between analysts' recommendations and the stock price reaction in China. Using a new methodology that combines the event of stock dividends and transfer of reserves to common shares, the author provides evidence of the decision value of analysts' recommendations that is different from the mature market. The results show that analysts' cumulative rating values positively relate to the cumulative abnormal returns. Favourable ratings result in the lower cumulative abnormal returns. The cumulative number of analyst rating agencies negatively relates to the cumulative abnormal returns. In general, analysts' information does not bring abnormal returns for investors.

Keywords: Analysts' Recommendations, Rating, Stock Dividends and Transfer of Reserves to Common Shares

Bao Xing A dual capacity sourcing model of disruption management for an injured power system

Computer Modelling & New Technologies 2014 18(3) 231-236

Great loss would be caused when power system lost its critical capacity by the impact of extreme events. Disruption management of State Grid Zhejiang Electric Power Company of China (shorted for SGZEPC) suffered in 2008 was firstly investigated, and then a dual sourcing model of regular and expedite capacity during recovery periods is correspondingly presented in this paper. A mathematical model of capacity procurement in a multi recover periods is constructed at the aim of minimizing the disruption cost of injured power system. Three meaningful managerial insights are obtained through sensitivity analysis on key parameters, which is helpful for manager to make decision during the disruption period.

Keywords: disruption management, injured power system, dual capacity sourcing, multi-period

Jianxin Bi, Lianghai Lei Analysis on the dynamic effects of the aggregate supply, aggregate demand and macroeconomic policies of China based on SVAR model

Computer Modelling & New Technologies 2014 18(3) 237-248

After studying the financial crisis using the AS-AD model and the SVAR model, the paper analyses the dynamic effects of the aggregate supply, aggregate demand and macroeconomic policies of China. Then, combining the real macroeconomic environment of China at present, the paper discusses the Keynesian AD-AS model, gives the constraint conditions of SVAR model according to the economic meaning of China, makes an empirical study based on five selected variables including supply, demand, fiscal expenditure, monetary expenditure and interest rate and their monthly data correspondingly, and lastly analyses the results of empirical study to make recommendations on current macroeconomic policy adjustment of China.

Keywords: AD-AS model, SVAR model, impulse response, variance analysis

Qiong Gu, Wu Zheng, Xianming Wang Study on Xiangyang's population and aging trend prediction based on discrete population development equation model

Computer Modelling & New Technologies 2014 18(3) 249-253

Population problem is an important factor that influences economy and social development of China. This paper takes the statistic data of 6th census in 2010 in Xiangyang as the accordance to establish a discrete model of population development equation, to analyse the population aging trend in the future in Xiangyang from a short period, and further to predict the long-term population development trend and aging population change condition in Xiangyang in the case of different total fertility rate to provide reference accordance for the government to make relevant social and economic decisions.

Keywords: population aging, population development equation, discrete model, total fertility rate

Dong-ping Li, Kong Xiangsheng A Study on Fast Assessment of Medium and Small Earthquake

Computer Modelling & New Technologies 2014 18(3) 254-257

The key to rapid assessment of earthquake losses is to identify the seismic intensify area. The information about the scale of earthquake may help the government and the relative department to make countermeasures, dispose disaster rescue action and strive for foreign aid. In this paper, the data of history earthquakes of Zhejiang Province and its surrounding areas, after being processed by GIS system, are used to access the length of the earthquake axis parameters. Then the data are compared with the tectonic structure of the area to determine the classification. After that, the rapid assessment of earthquake model is applying to the axis parameters of earthquakes, which have impact in Zhejiang Province. The model can provide references of earthquake rapid assessment.

Keywords: Fast Earthquake Loss Assessment, GIS, Earthquake disaster emergency

Wan He Deep neural network based load forecast

Computer Modelling & New Technologies 2014 18(3) 258-262

Accurate electrical load forecast has great economic and social value. In this paper, we study deep neural networks based load forecast approaches. We first analyse the critical features related to load forecast. Then we present details of deep neural networks and pre-training technologies, including RBM pre-training and discriminative pre-training. We compare the performances of different neural network models and show the advantages of the proposed methods using a rather large data set of loads.

Keywords: Load Forecast, Deep Neural Networks, Pre-training, RBM

Junxiang Tu Retrieving product information of collaborative enterprises based on Bayesian network

Computer Modelling & New Technologies 2014 18(3) 263-266

There exist many differences in nomenclature and descriptions of products and parts in collaborative enterprises, which greatly hinder the retrieval and sharing of web-based product information. In this paper, we present an extended Bayesian network for retrieving and integrating the product information of collaborative enterprises based on product ontology. This approach not only reduces the complexity of existing ontology mapping methods, but also increases the efficiency of product information integration.

Keywords: Product Information Retrieval, Bayesian Network, Ontology, Collaborative Enterprises

Li Wu, Xinyuan Wang Geoinformatics-based study on the regionalization of ecological function in the Chaohu Lake Basin, East China

Computer Modelling & New Technologies 2014 18(3) 267-270

Ecological function regionalization is a kind of geographic spatial division, which is based on the spatial differentiation of ecosystem functions. Based on an analysis of the primary features of the ecological environment of Chaohu Lake Basin in Anhui Province, the principles, bases, methodology and nomenclature of ecological function regionalization were determined. As the sub-valley is an independent geographical unit within the lake basin, its ecosystem sustains ecological integrity from the upstream through to the downstream. Therefore, ensuring the monitoring and management of the regional ecological environment in the sub-valley unit is of great importance to the conservation and ecological restoration of the regional ecosystem. Through extraction of land use information from remote sensing data, and sub-valley division from DEM analysis, this paper discusses the methodology of sub-valley ecological function regionalization in the research area based on the application of geoinformatics technology (e.g. RS and GIS technology). The ecological function regionalization of the Chaohu Lake Basin is then calculated, and the five ecological function regions and twelve sub-regions are subdivided. This study has an important practical relevance for the integrated management of the ecological environment of the Chaohu Lake Basin, and provides scientific grounds for the improved industrial distribution, ecological hazard prevention and reduction, environmental protection and construction planning in this area.

Keywords: RS and GIS, ecological function regionalization, ecological environment, Chaohu Lake Basin

Yue Wei Ai, Yan He, Zhi Jian Wang, Yang Wang A new method of digital manufacturing of orthoses

Computer Modelling & New Technologies 2014 18(3) 271-275

The proportion of disabled people is rising and now represents 1 billion people—15% of the global population, which leads to increasingly demand for orthotic device. However, moulds for orthoses manufacturing through traditionally manual technique are often dedicated, and this causes problems such as long lead time, lack of flexibility, low-efficiency and material waste, further leading to serious financial burns and environmental pollution as well. In this paper, an innovative method is proposed to replace traditionally dedicated moulds with reconfigurable moulds utilizing screw-pins that are directly transferred to the vacuum forming of thermoplastic material at low cost for the fabrication of orthoses. In the developed system, the fast reconstruction of human body anatomy based on the 3D digital scanning, is introduced firstly, the reconfigurable mould utilizing screw-pins is then generated and machined based on the reconstructed human body anatomy. After this, vacuum forming is performed on the reconfigurable mould, which could be reused for different anatomical shape variations by adjusting screw-pins. Additionally, an intelligent database is developed and a lot of reconstructed anatomies, the best practices of experienced orthotists, optimal parameters for 3D digital scanning, reconfigurable mould generation and machining and vacuum forming are stored, which will allow rapid recall of the stored information to reduce too much man-machine interaction and expertise dramatically.

Keywords: Orthoses, Digital manufacturing, Reconfigurable moulds, Fast reconstruction

Wei Sun, Yang Yu Evolutionary game and simulation of organizational information security investment

Computer Modelling & New Technologies 2014 18(1) 276-282

To investigate the evolution law of organizational information security investment, this paper analyses evolutionary stable strategies of organizational information security investments using evolutionary game theory and verifies the evolutionary stable strategies through the simulation based on Repast, a multi-agent simulation platform. First, according to the bounded rationality of actual organizations, this paper sets up the evolutionary game model of organizational information security investment. And then, we investigate the evolutionary stable strategies by replicator dynamics. Finally, we simulate the evolutionary game by Repast based on Java programming language, and the experimental results verify the evolutionary stable strategies obtained from the theoretical analysis. The research results can be used to predict the long-term stable trend of organizational information security investment, state that investment cost is the key for organizations to choose the strategy, and provide decision support for organizational information security investment.

Keywords: evolutionary game, information security investment, multi-agent simulation, evolutionary stable strategy

Innovative Education

Laiquan Liu, Li Lei, Yanrui Lei The application of fuzzy association rules in the employment data mining of a higher vocational college

Computer Modelling & New Technologies 2014 18(3) 283-287

Data mining is able to extract potentially useful information from plentiful seemingly unrelated data. A high efficiency is therefore obtained using these useful data in work or study. Association rules mining is a significant branch in data mining. It mirrors the implicit relations among transactions in mass data. In addition, association rules can intuitively reflect the associations among item sets in data, and the relations are established according to the frequencies of the item sets appearing in data. This method, which explains its rules clearly and is easily to understand, therefore is different from the traditional statistical method. This research introduced and applied the mining algorithms of fuzzy association rules to the employment data analysis of a higher vocational college, in order to find significant association rules from numerous data and provide guidance for the education and employment in the future, therefore improving the employment rate further.

Keywords: Association rules, Data mining, Research, Application

Jia Geng Information technology-based promotion of educational resource sharing

Computer Modelling & New Technologies 2014 18(3) 288-292

Information technology (IT) has revolutionary influence on the distribution of educational resources, and it plays a role in promoting the sharing of educational resources. Based on absorbing and inheriting the results of previous studies, this dissertation proposed to collect funds to construct large-scale digital education information resource database with centring on counties without building resource centre at basic education schools at all levels. In addition, the schools shall jointly expand and enrich the central educational resources to achieve sharing and co-construction of educational resources in the region. This dissertation also constructed a model of promoting sharing of educational resources by information technology in urban and rural areas.

Keywords: Information Technology, Educational Resource, Sharing

Nanoscience and Nanotechnology

Xia Ma, Mian Mian Zhang, Shi Wen Chen Optimized preparation of γ -polyglutamic acid/chitosan nanocapsule

Computer Modelling & New Technologies 2014 18(3) 293-298

This paper described the preparation of novel biodegradable nanocapsule based on self-assembly of γ -Polyglutamic acid (γ -PGA) and chitosan (CS). After the Plackett-Burman design (PBD), the impact of mass concentration and volume of γ -PGA and pH value of CS were characterized by size and PDI of the nanocapsule. A Box-Behnken design (BBD) was used to optimize the preparation of the nanocapsule. The optimized condition was: pH value of CS was 4.0; volume of γ -PGA was 18mL; mass concentration of γ -PGA was 0.1g/L. The Z-Ave and PDI of the nanocapsules prepared under the best conditions were 175 nm and 0.15 respectively. In this work, we have shown that nano-sized particles have been successfully assembled from the γ -PGA and chitosan without employing covalent linkages between these biopolymers. These results will provide a novel concept in the design of carrier systems composed of polyion complex (PIC).

Keywords: γ -Polyglutamic acid, chitosan, nanocapsule, optimization, response surface methodology