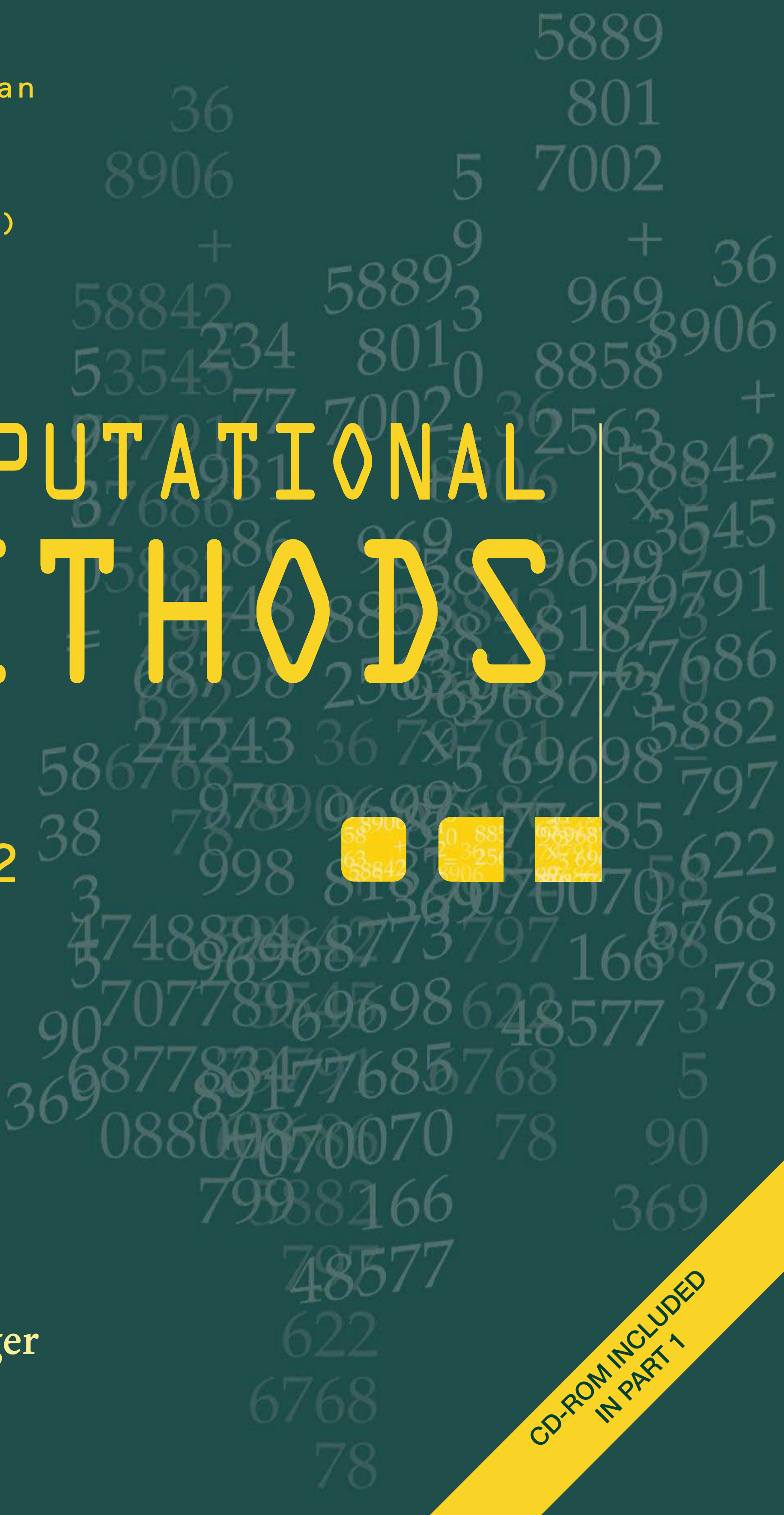
G.R. Liu
V.B.C. Tan
X. Han
(Editors)



COMPUTATIONAL METHODS

Part 2

CD-ROM INCLUDED
IN PART 1

Springer

COMPUTATIONAL METHODS

Computational Methods

Edited by

G.R. LIU
National University of Singapore, Singapore

V.B.C. TAN
National University of Singapore, Singapore

and

X. HAN
Hunan University, P.R. China

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-3952-2 (HB)
ISBN-13 978-1-4020-3953-2 (HB)
ISBN-10 1-4020-3953-0 (e-book)
ISBN-13 978-1-4020-3953-9 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

Printed in the Netherlands.

CONTENTS

PLENARY TALKS

BEM
(Organized by Ch. Zhang)

BIO-SYSTEM MODELING

CFD/COUPLE

COMPUTATIONAL METHODS IN GEOMECHANICS (Organized by A. Murakami and A. Yashima)

COMPUTATIONAL METHODS FOR MATERIALS ANALYSIS

COMPUTATIONAL TECHNIQUES
(Organized by R.S.R Gorla)

FRACTURE/DAMAGE/FATIGUE MECHANICS

GENERALIZED FEM AND EXTENDED/FICTITIOUS/EULERIAN FEM

HEALTH MONITORING

HEAT TRANSFER/FLUID FLOW

INVERSE & OPTIMIZATION/SEARCH METHODS

MACHINE LEARNING
(Organized by Y. C. Liang)

MESHFREE/PARTICLE METHODS

MULTI-PHYSICS/MULTISCALE ANALYSIS

RIGID BODIES/SOLIDS/STRUCTURES

ROCK/CONCRETE/SOIL

SHOCK & IMPACT MECHANICS

SPORTS DYNAMICS & ENVIRONMENTAL ANALYSIS

VIBRATION/EIGENVALUE ANALYSIS

PREFACE

The *First International Conference on Computational Methods* (ICCM04), organized by the Department of Mechanical Engineering, National University of Singapore, was held in Singapore, December 15–17, 2004 with great success. The conference brought together over 200 engineers and scientists to promote and exchange ideas, solutions and experiences on computational methods.

This conference proceeding contains some 280 papers and features 4 plenary papers, 17 keynote papers and 8 invited papers from more than 30 countries/regions. The papers cover a broad range of topics such as meshfree particle methods, Generalized FE and Extended FE methods, inverse analysis and optimization methods. Computational methods for geomechanics, machine learning, vibration, shock, impact, health monitoring, material modeling, fracture and damage mechanics, multi-physics and multi-scales simulation, sports and environments, etc., are also included. We hope that the proceedings will provide an informative, timely and invaluable resource for engineers and scientists working in the important areas of computational methods.

We would like to thank all authors and participants for sharing their ideas and developments in the area of computational methods. We are also appreciative of all the session and minisymposium organizers for their hardwork in putting up many interesting sessions for the conference. We wish to acknowledge the help we received in reviewing papers from members of the organizing committee, the international advisory panel and the international scientific committee. Special thanks are due to Hui Leng and Vivienne for their assistance in conference management and preparation of the proceedings. Finally, the organizers would like to express their thanks to Prof. Shih Choon Fong, President, National University of Singapore, for his strong support and guidance during the organization of this conference.

G.R. Liu
V.B.C. Tan
X. Han

TOPOLOGICAL OPTIMIZATION OF CONTINUUM STRUCTURE WITH GLOBAL STRESS CONSTRAINTS BASED ON ICM METHOD

Y.K. Sui, H.L. Ye and X.R. Peng
Numerical Simulation Center for Engineering, Beijing University of Technology, Beijing 100022, China

Abstract Continuum topology optimization with stresses constraints is not only important but also complicated. Because stress is a local quantity, a large number of constraints must be considered, and the complication of optimization algorithm and sensitivity analysis is also increased. A global stress constraints method based on ICM (Independent Continuous Mapping) method, which takes minimum weight and strain energy of structure with multi-load-case as design objective and constraint respectively, is suggested in this paper. Dual quadratic programming is applied to solve the optimal model for continuum structure established in this paper. As a result, the number of constraints is reduced and the local optimal solution of the weight is also avoided. Two numerical examples are discussed, and their results show that the present method is effective and efficient.

Keywords: structural topological optimization, global stress constraints, ICM method, continuum structure.

1. INTRODUCTION

Topology optimization has been extremely active since pioneering work of Bendsoe and Kikuchi [1]; a review on this regard is given in reference [2]. ICM (Independent Continuous Mapping) method is used to analyse an optimization model for continuum structure where the topological variables are independent and continuous [3]. Topological optimization for continuum structures has been extensively studied with displacement, volume and frequency constraints which are global. However there has been relatively little research

G. R. Liu et al. (eds), Computational Methods, 1003–1014.

done on stress constraints. The reason is that stress is a local quantity and nonlinear with respect to the design variables. Major difficulties that topological optimization problems with stress constraints have to deal with are a large number of constraints, thus the complication of both optimization algorithm and sensitivity analysis is significantly increased. The fully stressed design was proposed to tackle such a problem without needing sensitivity analysis [4–6]. A global stress constraints method based on ICM method and the von Mises' yield criterion, which takes minimum weight and strain energy of structure with multi-load-case as design objective and constraint respectively, is presented in this paper. Then a new optimal model of continuum structure is established, and the dual quadratic programming is applied to analyse this optimal model. As a result, the number of constraints is reduced and the local optimal solution of the weight is avoided. Numerical examples show that the present model is effective and efficient.

2. TOPOLOGY OPTIMIZATION OF CONTINUUM STRUCTURE WITH GLOBAL STRESS CONSTRAINTS

2.1 The strategy of global stress constraints

In order to reduce the complication of the sensitivity analysis and the optimization algorithm, the stress constraints are replaced by the strain energy constraints. It can be found from the von Mises' yield criterion that distortion energy is a key variable of the bending body. Therefore it can be written as

$$e_{il}^{f}/V_i \leq (1+\nu)\overline{\sigma_i}^2/(3E), \tag{1}$$

where e_{il}^{f}, $V_i E$, ν and $\overline{\sigma_i}$ represent distortion energy of element, volume of element, elastic modulus of material, Poisson's ratio, and allowable stress, respectively. It is difficult to derive the form of distortion energy from the strain energy of element. However, this drawback can be overcome in terms of the stain energy. From the following inequality expression

$$e_{il}^{f} < e_{il} \tag{2}$$

we can readily obtain

$$e_{il} \leq (1+\nu)\overline{\sigma_i}^2 V_i \big/ (3E) \tag{3}$$

where e_{il} represents the strain energy of i-th element subjected to the l-th load case.

Whatsmore, some filter functions according to the ICM method are listed in the following, that is

$$w_i = f_w(t_i)w_i^0;\ \bar{\sigma}_i = f_\sigma(t_i)\bar{\sigma}_i^0;\ k_i = f_k(t_i)k_i^0 \tag{4}$$

where w_i^0, $\bar{\sigma}_i^0$ and k_i^0 represent the natural weight, allowable stress and stiffness of element, respectively. The stiffness filter function $f_k(t_i)$ on the left-hand side of the Equation (3) is further introduced and that is

$$e_{il}^v f_k^v(t_i)/f_k(t_i) \leq (1+v)\bar{\sigma}_i V_i/(3E) \tag{5}$$

where superscript v represents the number of iteration. It can be obtained by simplifying the above equations

$$e_{il}^v f_k^v(t_i) \leq (1+v)(\bar{\sigma}_i)^2 V_i/(3E)f_k(t_i) \tag{6}$$

Finally it can also be gotten by summing up these equations over the whole continuum structure

$$\sum_{i=1}^{N}[e_{il}^v f_k^v(t_i) - (1+v)(\bar{\sigma}_i)^2 V_i/(3E)f_k(t_i)] \leq 0 \tag{7}$$

It is noted that Equation (7) can readily be derived from Equation (6). However, the converse may not be true. Although this kind of global stress constraint cannot guarantee the stress to be satisfied at every point, it is satisfied in a global sense. The exact satisfaction of stress constraints can be performed by the simple optimization such as cross-sectional or shape optimization. From Equation (7), the local stress constraints can be transformed into global constraints, which is critical to establish a new optimal model.

2.2 Optimal model with global stress constraints

The topology optimization model of continuum structure with multi-load-cases, based on ICM method, can be established as follows. The weight is taken as the design objective, and the strain energy with multi-load-cases is considered to be the constraint:

$$\begin{cases} \textit{find}\ t = (t_1, \ldots, t_N)^T \\ W = \sum\limits_{i=1}^{N} f_w(t_i)w_i^0 \to \min \\ \text{s.t.}\ \sum\limits_{i=1}^{N}[e_{il}^v f_k^v(t_i) - d_i f_k(t_i)] \leq 0 \\ \qquad 0 \leq \underline{t_i} \leq t_i \leq 1 \\ \quad (i = 1, \ldots, N; l = 1, \ldots, L) \end{cases} \tag{8}$$

where $t = (t_1, \ldots, t_N)^T$ denotes the design variables, and $d_i = (1 + v)(\bar{\sigma}_i)^2 V_i / (3E)$. Define $f_w(t_i) = t_i^{\alpha}$, $f_k(t_i) = t_i^{\beta}$,where α, β are constant parameters, assuming $\alpha = 1$, $\beta = 3$ in our analysis. Because every element involves only one design variable, we can conveniently use dual programming to solve this model. Then Equation (8) can be converted into

$$\begin{cases} \text{find } \lambda \in E^L \\ \phi(\lambda) \to \max \\ \quad \text{s.t. } \lambda \geq 0 \end{cases} \tag{9}$$

where

$$\phi(\lambda)| = \min_{\underline{t_i} \leq t_i \leq 1} \left\{ S(t, \lambda) = \sum_{i=1}^{N} t_i^{\alpha} w_i^0 + \sum_{l=1}^{L} \lambda_l \sum_{i=1}^{N} \left(e_{il}^v f_k^v(t_i) - d_i t_i^{\beta} \right) \right\} \tag{10}$$

From the Kuhn–Tucker condition, one obtains

$$\partial S / \partial t_i = \alpha w_i^0 t_i^{\alpha - 1} - \sum_{l=1}^{L} \lambda_l \beta d_i t_i^{\beta - 1} \begin{cases} \quad > 0 \ (t_i^* = \underline{t_i}) \\ = 0 \ (\underline{t_i} < t_i^* < 1) \\ \quad < 0 \ (t_i^* = 1) \end{cases} \tag{11}$$

Solving Equation (10) we get

$$t_i^* = \begin{cases} \underline{t_i} \ (t_i^* \leq \underline{t_i}) \\ t_i \ (\underline{t_i} < t_i^* < 1) \\ 1 \ (t_i^* \geq 1) \end{cases} \tag{12}$$

where

$$t_i = \left(\beta d_i \sum_{l=1}^{J} \lambda_l / \left(\alpha w_i^0 \right) \right)^{1/(\alpha - \beta)} \tag{13}$$

An assumption $n_a = \{ i \,|\, \underline{t_i} < t_i^* < 1 \}$ is applied. If $i \notin n_a$ is known, we would have $t_i^* = \underline{t_i}$ or $t_i^* = 1$ and

$$\partial t_i^* / \partial \lambda_k = 0 \tag{14}$$

If $i \in n_a$ is prescribed, we would have

$$\partial t_i^* / \partial \lambda_k = \beta d_i (t_i^*)^{\beta - \alpha + 1} / \left(\alpha \left(\alpha - \beta \right) w_i^0 \right) \tag{15}$$

Differentiating Equation (10) with respect to λ by using dual quadratic theory, we get

$$\partial \phi(\lambda) / \partial \lambda_l = h_l(t_i^*(\lambda)) = \sum_{i=1}^{N} \left(e_{il}^v - d_i f_k(t_i^*) \right) = \sum_{i=1}^{N} \left(e_{il}^v - d_i (t_i^*)^{\beta} \right) \tag{16}$$

$$
\begin{aligned}
\partial^2\phi(\lambda)/(\partial\lambda_l\partial\lambda_k) &= \sum_{i=1}^{N}(\partial h_l(t_i^*)/\partial t_i^*)(\partial t_i^*/\partial\lambda_k) \\
&= -\sum_{i=1}^{N} d_i\beta(t_i^*)^{\beta-1}\partial t_i^*/\partial\lambda_k
\end{aligned} \tag{17}
$$

Substitute Equation (19) into Equation (21), we obtain

$$
\partial^2\phi(\lambda)/(\partial\lambda_l\partial\lambda_k) = -\sum_{i\in n_a}\left(d_i^2\beta^2(t_i^*)^{2\beta-\alpha}/(\alpha(\alpha-\beta)w_i^0\right) \tag{18}
$$

In the neighbourhood of λ^0, we expand $\phi(\lambda)$ into Taylor series and neglect all the terms higher than second order, the following model can be derived

$$
\begin{cases}
\text{find } \lambda \ \in \boldsymbol{E}^L \\
-\phi(\lambda) = -\sum\limits_{l=1}^{L}\sum\limits_{k=1}^{L}\boldsymbol{C}_i\lambda_l\lambda_k/2 - \sum\limits_{l=1}^{L}\boldsymbol{H}_{il}\lambda_l \to \min \\
\text{s.t.} \qquad \lambda \geq 0 \qquad (\boldsymbol{j} = 1, \ldots, \boldsymbol{J})
\end{cases} \tag{19}
$$

where $C_i = -\sum\limits_{i\in n_a}\beta d_i(t_i^*)^\beta/(\alpha-\beta)$

$$
H_{il} = \sum_{i=1}^{N}\left(e_{il}^v f_k^v(t_i) - d_i(t_i^*)^\beta\right) + \sum_{i\in n_a} d_i\beta(t_i^*)^\beta/(\alpha-\beta)
$$

The solution of λ can be found from Equation (19). Substituting λ into Equation (18), the corresponding t^* can also be found, and the next structural analysis will consequently be carried out. The numerical program will be stopped until the following condition is satisfied.

$$
\Delta\boldsymbol{W} = \left|(\boldsymbol{W}^{(v+1)} - \boldsymbol{W}^{(v)}/\boldsymbol{W}^{(v+1)}\right| \leq \varepsilon \tag{20}
$$

2.3 Continuous topology variable is mapped inversely into discrete topology variable

After topology optimization, a continuous solution should be obtained, which can be transformed into discrete solution by a defined value δ. If the value of topology variable is in the region of $(0, \delta)$, this element denoted by topology variable should be deleted from the entire discrete structure. Such element is referred to as 0 element. On the other hand, if the value of topology variable is in the region of $[\delta, 1]$, this element which is named as 1 element will be preserved. As usual, the value δ is ranged from 0.1 to 0.5.

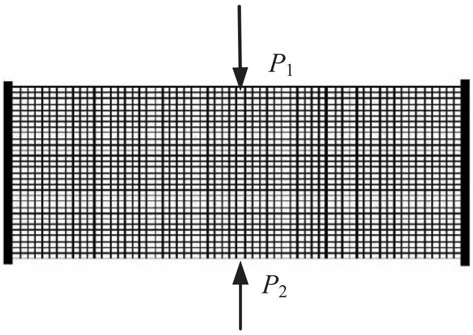


Figure 1. The finite element model.

3. NUMERICAL EXAMPLES

It is noted that the sizes and parameters of the following examples are dimensionless.

Example 1 A two-dimensional rectangular flat plate 120 × 60, which is fixed at two sides and loaded, is shown in Figure 1. In the model, $P_1 = P_2 = 2000$, the thickness is adopted to be 6.0, the density of material is used to be 1.0, the elastic modulus of material is 2.1×10^7, and the Poisson's ratio is 0.25. The design domain is discretized into 60 × 30 meshes by using 4-node membrane element. Each load case is corresponding to only one concentrated force. The allowable stress is 155, and the weight of structure is 43200. The convergence precision is limited to be 0.001. δ is 0.1. Then the optimal topology structure with simple load case is given in Figures 2 and 3, and the topology structure with multi-load-case is shown in Figure 4. The iterative number with respect to weight is depicted in Figure 5. Furthermore some stress fringes are given in Figures 6–8. It can be seen that the distribution of stress after optimization tends to be uniformity. There is no element which von Mises' stress exceeds allowable stress. The number of elements in the range of topology variables is given in

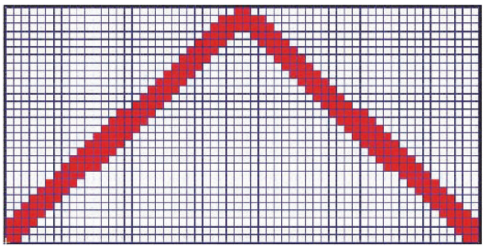


Figure 2. The optimal topology structure with load case 1.

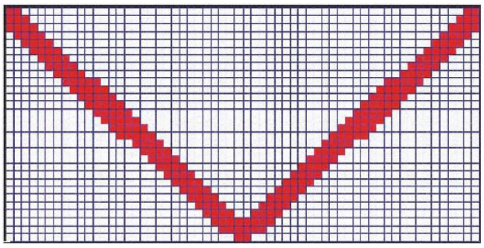


Figure 3. The optimal topology structure with load case 2.

Table 1. It can be seen that the value δ should be adopted in the region (0.1, 0.5). Such δ has little effect on the optimal topology structure.

Example 2 Second example is identified with the preceding one mesh of calculation parameters, except that the central load is located on the endpoint the right-hand side of the structure and the only left-hand side is fully fixed (shown in Figure 9). Calculation cases are also same as the example mentioned previously. The allowable stress is 155. The convergence precision is 0.001. δ is 0.1. The results after topological optimization are shown in Figures 10–13. Compared with reference [7], it can be seen the number of iteration is decreased. And some stress fringes are given in Figures 14–16. It can be seen that the distribution of stress after optimal topology tends to uniformly. In addition, there are only 8 elements which maximum von Mises' stresses are higher than allowable stresses. It can be performed by sectional optimization to satisfy the stress constraint. The number of elements in the range of topology variables is given in Table 2.

4. CONCLUSION

In this paper, the global stress constraints topological optimization problem has been investigated. Based on ICM method and the von Mises' yield criterion,

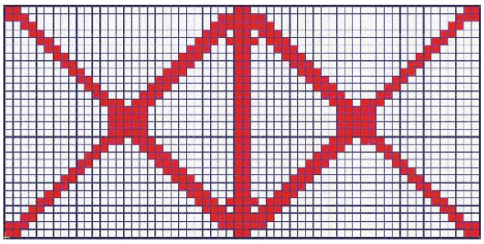


Figure 4. The optimal topology structure with two load cases.

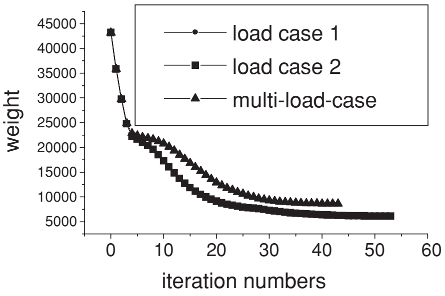


Figure 5. The iterative history of weight.

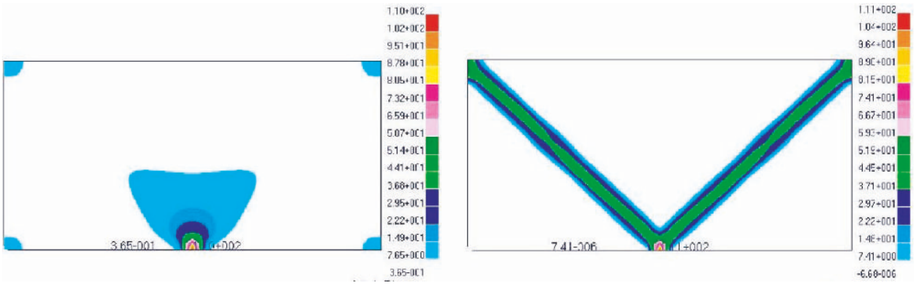


Figure 6. The stress fringe of original structure and topological structure with load case 1.

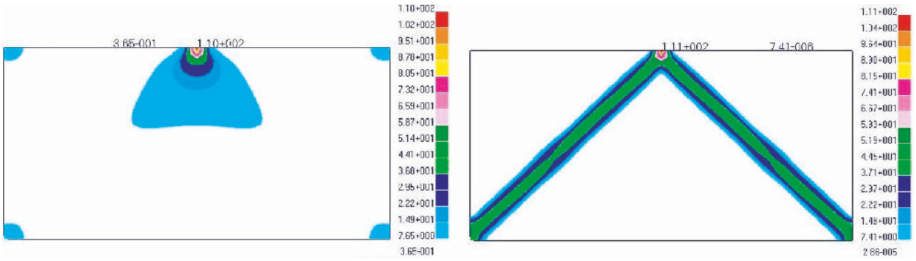


Figure 7. The stress fringe of original structure and topological structure with load case 2.

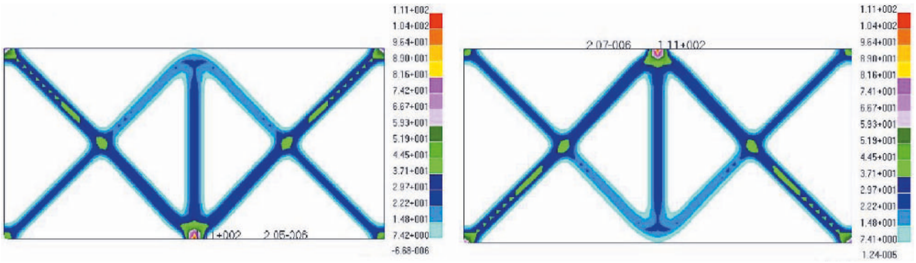


Figure 8. The stress fringe of topological structure of multi-load-cases with load case 1 and load case 2.

Table 1. The number of elements in the range of topology variables.

Range of variable		0.9–1	0.7–0.9	0.5–0.7	0.3–0.5	0.1–0.3	<0.1
Number of elements	Load case1	222	16	2	2	0	1558
	Load case2	222	16	2	2	0	1558
	Two load cases	320	12	19	5	0	1444

Table 2. The number of elements in the range of topology variables.

Range of variable		0.9–1	0.7–0.9	0.5–0.7	0.3–0.5	0.1–0.3	<0.1
Number of elements	Load case1	526	20	8	7	7	1232
	Load case2	526	20	8	7	7	1232
	Two load cases	534	89	17	8	10	1142

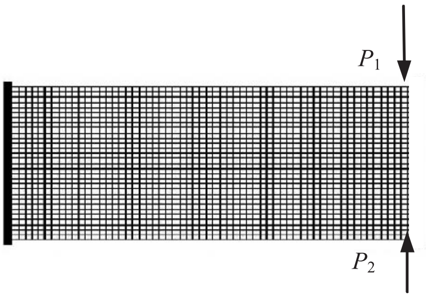


Figure 9. The finite element model.

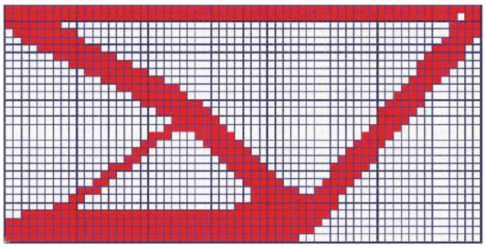


Figure 10. The optimal topology structure with load case 1.

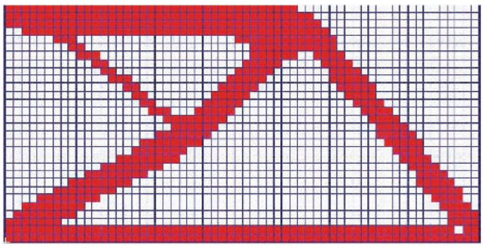


Figure 11. The optimal topology structure with load case 2.

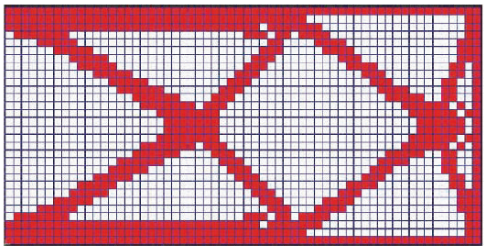


Figure 12. The optimal topology structure with two load cases.

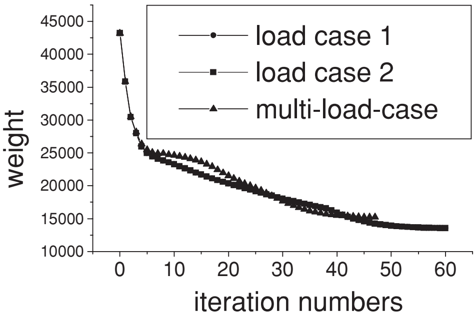


Figure 13. The iterative history of weight.

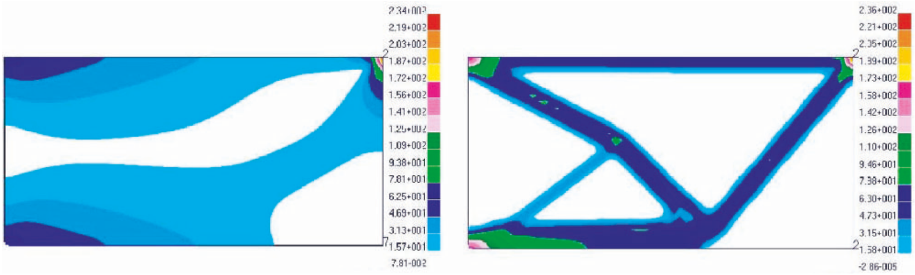


Figure 14. The stress fringe of original structure and topological structure with load case 1.

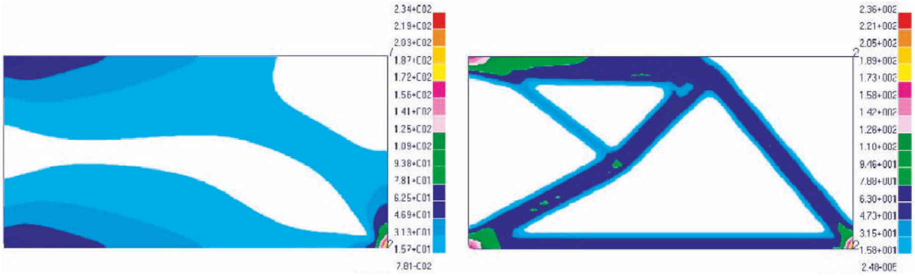


Figure 15. The stress fringe of original structure and topological structure with load case 2.

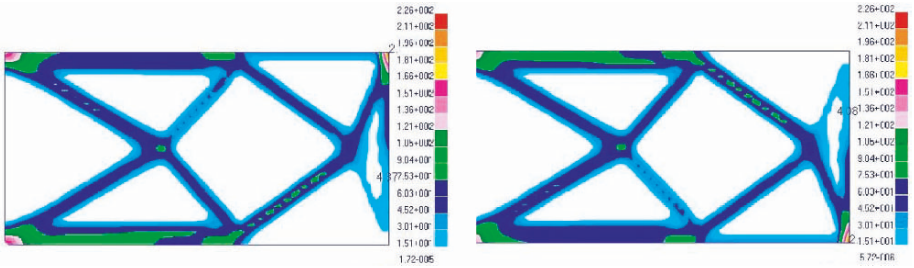


Figure 16. The stress fringe of topological structure of multi-load-cases with load case 1 and load case 2.

a new optimal model for continuum structure with stain energy constraints is established. Furthermore, this model is analysed by using dual quadratic programming. It can be seen from the result that the number of constraints is reduced, the sensitivity analysis of structure is avoided, and the efficiency of calculation is increased. Numerical example shows that the present model is effective. In addition, the weight of topological structure may change when the continuous topological variables are discretized as 0 or 1, which lead to some gray elements are transformed into 0 or 1 element.

ACKNOWLEDGEMENTS

Supported by National Natural Science Foundation of China (10472003), Beijing Natural Science (3042002) and Beijing Educational Committee (KM200410005019) Foundations.

REFERENCES

1. M.P. Bendsoe and N. Kikuchi (1988), Generating optimal topologies in structural design using a homogenization method. *Computer Methods in Applied Mechanics and Engineering*, 71, pp. 197–224.
2. H.A. Eschenauer and N. Olhoff (2001), Topology optimization of continuum structures: a review. *Applied Mechanics Review*, 54, pp. 331–389.
3. Y.K. Sui (1996), *Modeling, Transformation and Optimization: New Developments of Structural Synthesis Method*. Dalian University of Technology Press (in Chinese).
4. G. Rozvany (2001), Stress ratio and compliance based methods in topology optimization—a critical review. *Structural and Multidisciplinary Optimization*, 21, pp. 109–119.
5. J. Wang and G.D. Cheng (1997), Optimal topology design of thin plate with stress constraints. *Acta Mechnica Solida Sinics,* 18, 4, pp. 317–322 (in Chinese).
6. J.H Rong and J.S. Jiang (2003), A structural topology evolutionary optimization method based on stresses and their sensitivity. *Acta Mechnica Sinica*, 3535, 5, pp. 584–591 (in Chinese).

7. Y. Xin (1999), The Exist-null Combination Method for the Optimum Design of Structural topology and Its Computer Implementation, Ph.D. Dissertation, Dalian University of Technology (in Chinese).
8. R.J. Yang and C.J. Chen (1996), Stress-based topology optimization. *Structrual Optimization*, 12, pp. 98–105.

TOPOLOGICAL OPTIMIZATION OF FRAME STRUCTURES UNDER MULTIPLE LOADING CASES*

Yun Kang Sui[1], Jia Zheng Du[1] and Ying Qiao Guo[2]
[1]*Numerical Simulation Center for Engineering, Beijing University of Technology, Beijing 100022, Chin*
[2]*Lab. Mechanics, Materials & Structures, University of Reims Champagne-Ardenne, Reims 51687, France*

Abstract Structural topological optimization is to seek the best path of transmitting forces for structures. The topological optimization under multiple loading cases involves the balance of many paths of transmitting forces. Based on the Independent Continuous Mapping (ICM) method, the optimization problem under multiple loading cases is solved under three conditions, namely local constraints, global constraints, and their combination. In this paper, local constraints are firstly analyzed by envelope method and average method. Secondly, global constraints are uniformly calculated with mathematical programming (MP). Thirdly, local and global constraints are processed by synthesizing above two methods. Finally, the results are compared. From the present numerical examples, it is shown that the envelope method and the MP or their combination can be used to efficiently model and accurately simulate the topological optimization problem under multiple loading cases.

Keywords: frame structures, topological optimization, envelope method, average method, mathematical programming.

1. INTRODUCTION

The concept of structural topological optimization was proposed by Michell [1], but the research works were mainly concentrated on the truss and continuum

---

*Supported by National Natural Science Foundation of China (10472003), Beijing Natural Science (3042002) and Beijing Educational Committee (KM200410005019) Foundations.

G. R. Liu et al. (eds), Computational Methods, 1015–1022.

structures [2–6]. The topological optimization of frame structures under single loading case was studied in previous works [7, 8]. However, the topological optimization problem under multiple loading cases is more difficult to be solved because it involves the balance of many paths of transmitting forces [9–13]. The optimization problem for frame structure under multiple loading cases is presented in this paper. The main idea is that our approach can deal with three conditions, namely local constraints, global constraints and their combination. Firstly, local constraints are analyzed by envelope method and average method. Secondly, global constraints are uniformly calculated with mathematical programming (MP). Thirdly, local and global constraints are processed by synthesizing above two methods. Finally, the results are compared.

2. TOPOLOGICAL OPTIMIZATION MODEL BASED ON ICM METHOD [14, 15]

The topological optimization model of frame structures with stress and displacement constraints under single loading case is expressed as follows:

$$\begin{cases} \text{Find} \quad t_i \quad (i = 1, \ldots, n) \\ \text{Make} \quad w = \sum_{i=1}^{n} w_i \to \min \\ \text{Subject to} \quad \sigma_i \le \bar{\sigma}_i \quad (i = 1, \ldots, n) \\ u_j \le \bar{u}_j \quad (j = 1, \ldots, J) \\ 0 \le t_i \le 1 \quad (i = 1, \ldots, n) \end{cases} \tag{1}$$

where t_i is the topological variable, n is the number of variables, J is the number of displacement constraints.

Based on the Independent Continuous Mapping (ICM) method, three filter functions for element weight, element allowable stress and element stiffness are introduced, noted $f_1(t)$, $f_2(t)$ and $f_3(t)$. The relations of element weight, element allowable stress and element stiffness are given by

$$w_i = f_1\,(t_i) w_i^0 \tag{2}$$

$$\bar{\sigma}_i = f_2\,(t_i) \bar{\sigma}_i^0 \tag{3}$$

$$k_i = f_3\,(t_i) k_i^0 \tag{4}$$

where w_i, $\bar{\sigma}_i$ and k_i are, respectively, the element weight, element allowable stress and element stiffness related to t_i; w_i^0, $\bar{\sigma}_i^0$ and k_i^0 are the same physical quantities related to the initial topological variables.
Assuming

$$f_1\,(t) = t^{\alpha},\ f_2\,(t) = t^{\beta},\ f_3\,(t) = t^{\gamma} \tag{5}$$

where α, β and γ are constant related to the structural types, which can be obtained according to the relations of w_i, $\bar{\sigma}_i$ and k_i or by numerical experiments. For the frame structures, if the axial force is ignored, we can obtain $\alpha = 1$, $\beta = 1.5$, $\gamma = 2$.

3. OPTIMIZATION PROBLEM UNDER MULTIPLE LOADING CASES WITH STRESS CONSTRAINTS

The stress constraint is a local constraint, which can be approached in zero-order approximation and transformed into movable lower limits of topological variables with the full stress criterion.

The full stress criterion gives the following relation:

$$\sigma_i^*/\bar{\sigma}_i = 1 \tag{6}$$

Substituting Equation (6) to Equation (3), we can obtain

$$\sigma_i^*/(f_2(t_i^*)\bar{\sigma}_i^0) = 1 \tag{7}$$

So

$$t_i^* = (\sigma_i^*/\bar{\sigma}_i^0)^{2/3} \tag{8}$$

where t_i^* is the continues topological variable satisfying the stress constraints.

For the problem with multiple loading cases, the envelope method and the average method are used as following.

The envelope method takes the maximum values of the topological variables under all loading cases, which can be expressed as

$$t_i = \max_{l=1,\cdots,L}(t_{il}^*) \tag{9}$$

The average method takes the average values of the topological variables under all loading cases, which can be written as

$$t_i = \sum_{l=1}^{L} t_{il}^*/L \tag{10}$$

4. OPTIMIZATION PROBLEM UNDER MULTIPLE LOADING CASES WITH DISPLACEMENT CONSTRAINTS

Unlike the stress constraint discussed previously, the displacement constraint is a global constraint. According to unit virtual load method, the displacement with respect to the j^{th} displacement constraint can be denoted as follows:

$$u_j = \sum_{i=1}^{n}\sum_{e\in i}\int_e \left(\frac{M_y\bar{M}_y}{EI_y} + \frac{M_z\bar{M}_z}{EI_z} + \frac{M_x\bar{M}_x}{GI_p} + \frac{\mu_y Q_y\bar{Q}_y}{GA} + \frac{\mu_z Q_z\bar{Q}_z}{GA} + \frac{N\bar{N}}{EA}\right)dx \tag{11}$$

Since the element stiffness is proportional to the Young's modulus, according to the Equation (4), Equation (11) can be simplified to

$$u_j = \sum_{i=1}^{n}\frac{1}{f(t_i)}u_{ij}^0 = \sum_{i=1}^{n}\frac{1}{t_i^2}u_{ij}^0 \tag{12}$$

where $u_{ij}^0 = \sum\limits_{e^0\in i}\int_{e^0}\left(\frac{M_y\bar{M}_y}{EI_y} + \frac{M_z\bar{M}_z}{EI_z} + \frac{M_x\bar{M}_x}{GI_p} + \frac{\mu_y Q_y\bar{Q}_y}{GA} + \frac{\mu_z Q_z\bar{Q}_z}{GA} + \frac{N\bar{N}}{EA}\right)dx$ is a constant and can be derived from the results of structural analysis, e^0 is the element belonging to the i^{th} topological variables.

Therefore we can obtain the explicit formulation:

$$\sum_{i=1}^{n}\frac{1}{t_i^2}u_{ij}^0 \le \bar{u}_j \tag{13}$$

For the problem with multiple loading cases, the explicit expressions for every loading case can be written as

$$\sum_{i=1}^{n}\frac{1}{t_i^2}u_{ijl}^0 \le \bar{u}_j \tag{14}$$

where l is the number of loading cases, u_{ijl}^0 is the constant corresponding to the j^{th} constraint under the l^{th} loading case.

Taking into account all the constraints in one model, we can obtain the following optimization model:

$$\begin{cases} \text{Find} & t_i \quad (i = 1, \ldots, n) \\ \text{Make} & w = \sum\limits_{i=1}^{n} w_i = \sum\limits_{i=1}^{n} t_i w_i^0 \to \min \\ \text{s.t.} & \sum\limits_{i=1}^{n}\frac{1}{t_i^2}u_{ijl}^0 \le \bar{u}_j \quad (j = 1, \ldots, J;\ l = 1, \ldots, L) \\ & 0 \le t_i \le 1 \quad (i = 1, \ldots, n) \end{cases} \tag{15}$$

5. OPTIMIZATION PROBLEM UNDER MULTIPLE LOADING CASES WITH STRESS AND DISPLACEMENT CONSTRAINTS

Generally, the stress and displacement constraints are required in optimization problem. For such a problem, the methods dealing with local and global constraints should be combined together.

The model with stress and displacement constraints is difficult to be solved for the stress and displacement constraints are two different types of constraints. Thus, the stress constraints should be processed with the envelope or average method, and transformed into the movable lower limits of topological variables as follows:

$$t_i \geq \underset{\sim}{t_i} = t_i^* \tag{16}$$

Therefore, we can represent the model with stress and displacement constraints:

$$\begin{cases} \text{Find} & t_i \quad (i = 1, \ldots, n) \\ \text{Make} & w = \sum_{i=1}^{n} w_i = \sum_{i=1}^{n} t_i w_i^0 \rightarrow \min \\ \text{s.t.} & \sum_{i=1}^{n} \frac{1}{t_i^2} u_{ijl}^0 \leq \bar{u}_j \quad (j = 1, \ldots, J;\ l = 1, \ldots, L) \\ & \underset{\sim}{t_i} \leq t_i \leq 1 \quad (i = 1, \ldots, n) \end{cases} \tag{17}$$

where $\underset{\sim}{t}_i = \sum_{l=1}^{L} t_{il}^*/L$ (the average method) or $\underset{\sim}{t}_i = \max\limits_{l=1,\ldots,L} (t_{il}^*)$ (the envelope method).

6. NUMERICAL EXAMPLES

A spatial frame structure with 390 beams is presented in Figure 1. The bottom of the structure is fixed. There are two forces ($F_1 = F_2 = 100$ kN) on the top, which are depicted as two loading cases: Case I and Case II. The material properties are: Young's modulus $E = 200$ GPa, Poisson ratio $v = 0.3$, density $\rho = 7800$ kg/m^3, allowable stress $\sigma_a = 160$ MPa. The initial sectional moment of inertia of all elements is 8.333E-6 m^4.

Taking structure's weight as objective function, every beam element as independent topological variable, the structure is optimized with different constraints under the two loading cases (Case I and Case II).

Just with the stress constraint (the maximal stresses do not exceed the allowable stress, 160 MPa), the problem is, respectively, solved by the envelope method

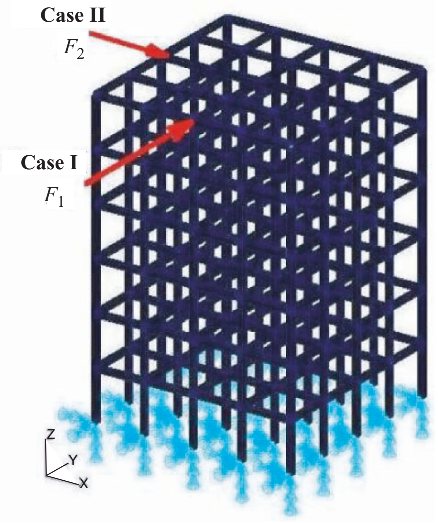


Figure 1. Spatial frame structure.

and the average method. The results for comparison are shown in Table 1 and the optimal topology is given in Figure 2 (the left figure is obtained by the envelope method and the right figure by the average method).

With the stress and displacement constraints (the maximal stresses do not exceed the allowable stress, 160 MPa, and the maximal displacements do not exceed 20 mm), the problem is analyzed by combining mathematical programming with the envelope method or the average method. The results for comparison are shown in Table 2 and the optimal topology is given in Figure 3 (the left

Table 1. Result comparison just with stress constraints.

Methods	Iteration number	Maximal stress (MPa)	Structural weight (kg)
Envelope method	52	104.926	7800
Average method	37	45.9477	19,188

Table 2. Result comparison with stress and displacement constraints.

Methods	Iteration number	Maximal stress (MPa)	Maximal displacement (mm)	Structural weight (kg)
Envelope method and MP*	35	107.388	14.2865	7488
Average method and MP	16	53.6709	8.61115	14,976

* MP is mathematical programming.

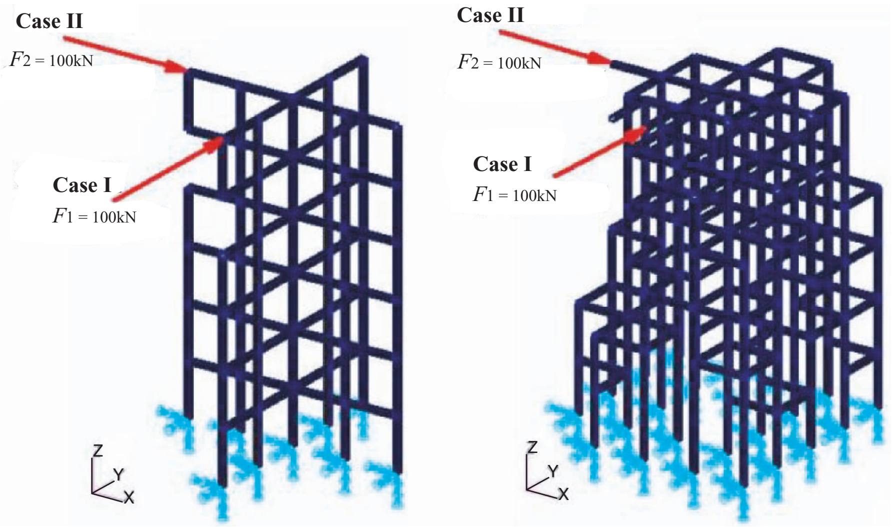


Figure 2. Optimal topology just with stress constraints.

figure is obtained by combining mathematical programming with the envelope method and the right figure by combining mathematical programming with the average method).

From the Tables 1 and 2 and Figures 2 and 3, we can conclude that the structural weight is much lighter and the optimal topology is more reasonable by the envelope method and the MP or their combination.

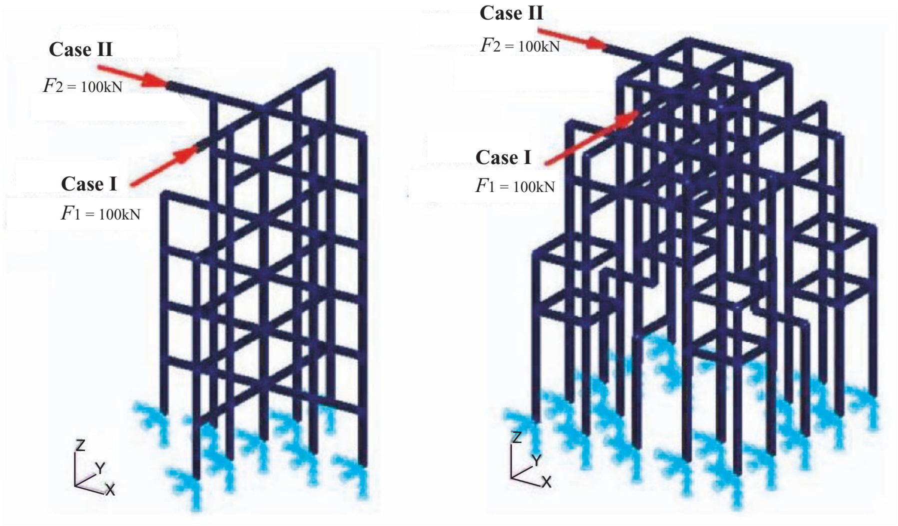


Figure 3. Optimal topology with stress and displacement constraints.

7. CONCLUSIONS

The methods to solve the optimization problem of frame structures under multiple loading cases are presented. Numerical examples show the envelope method and the MP or their combination can be used to efficiently model and accurately simulate the topological optimization problem.

REFERENCES

1. G.M. Michell (1904), The limits of economy of materials in frame structures. *Philosophical Magazine*, 8, pp. 589–597.
2. W. Dorn, R. Gomory and H. Greenberg (1964), Automatic design of optimal structures. *Journal de Mechanique*, 3, pp. 25–52.
3. U. Ringertz (1985), On topology optimization of trusses. *Engineering Optimization*, 9, pp. 209–218.
4. U. Kirsch (1989), Optimal topologies of truss structures. *Computer Methods in Applied Mechanics and Engineering*, 72, pp. 15–28.
5. G.D. Cheng and Z. Jiang (1992), Study on topology optimization with stress constraints. *Engineering Optimization*, 20, pp. 29–148.
6. M. Bechers and C. Fleury (1997), A primal–dual approach in truss topology optimization. *Computers & Structures*, 64, pp. 77–88.
7. Y.K. Sui, X.C. Ren and L.C. Long (2001), Topological optimization of frame structures. *Acta Mechanica Solida Sinica*, 22, pp. 95–100.
8. Y.K. Sui, X.C. Ren, L.C. Long and B.R. Ye (2003), Topological optimization of frame based on ICM method. *Chinese Journal of Computational Mechanics*, 20, pp. 286–289.
9. Y.F. Wang and H.C. Sun (1995), Optimal topology designs of trusses with discrete size variables subjected to multiple constraints and loading cases. *ACTA Mechanica Sinica*, 27, pp. 365–369.
10. Y.K. Sui, D.Q. Yang and B. Wang (2000), Topological optimization of continuum structure with stress and displacement constraints under multiple loading cases. *ACTA Mechanica Sinica*, 32, pp. 171–179.
11. Y.J. Wang, C. Liu and Y.R. Wang (2002), On topology-type optimization of trusses under multiple load cases. *Mechanics in Engineering*, 24, pp. 27–30.
12. J. Wang and G.D. Cheng (2003), Topology optimization design of the continuum structure for multiple loading conditions with stress constraints. *Journal of Mechanical strength*, 25, pp. 55–57.
13. F.H. Yu and J.H. Rong (2004), Topology optimization of structures under multiple loads. *Mechanics in Engineering*, 26, pp. 50–54.
14. Y.K. Sui (1996), *Modelling, Transformation and Optimization—New Developments of Structural Synthesis Method.* Dalian University of Technology Press, Dalian, China.
15. Y.K. Sui and D.Q. Yang (1998), A new method for structural topological optimization based on the concept of independent continuous variables and smooth model. *ACTA Mechanica Sinica*, 18, pp. 179–185.

OPTIMAL DISPLACEMENT CONTROL SIMULATION OF ELECTRIC-MECHANICAL COUPLED TRUSSES

L.C. Long and Y.K. Sui
Numerical Simulation Center for Engineering, Beijing University of Technology, Beijing 100022, China

Abstract Finite element equations of electric-mechanical coupled piezoelectric truss structures are deduced using linear piezoelectricity and Hamilton's principle. An optimal control model of piezoelectric smart trusses is established. It makes the maximum value of the controlled nodal displacements to be minimized and subjected to the strength requirements of the structure and the characteristic of the actuators. Control variables of the model are the voltages of actuators. The model is transformed to a sequential linear programming ultimately. For static indeterminate trusses in some case, the maximum stress in the structure can be decreased to allowable limit by adjusting the length of the member if it exceeds the allowable limit. When the number of actuators is limited, if they are distributed reasonable and have sufficient active elongation, effective control can be realized also. The efficiency of the control model is simulated by numerical method.

Keywords: smart truss, displacement control, electric-mechanical coupling, optimization.

1. INTRODUCTION

Smart trusses have advantages such as less weight, easy for design and manufacture, safe and expandable on the track. It can be used as supporting structures of spatial apparatus. A trend penetrates to fields of mechanical, automobile, robot and so forth is appeared. Piezoelectric materials have piezoelectric effect or reverse piezoelectric effect. Using these two properties, respectively, sensor or actuator can be made. Piezoelectric materials possess advantages such as good controllability, stable performance and larger induced distance and otherwise, so they are applied in broad fields. Much research work on piezoelectric structures has been done.

G. R. Liu et al. (eds), Computational Methods, 1023–1035.

Finite element method is an effective numerical analysis method for structures. Piezoelectric smart trusses include electric-mechanical coupled element, no other than synthesized the mechanical quantity and the electricity quantity to describe it perfectly. The relation between the input excitation and the output response should be set in the finite element method of the piezoelectric smart trusses.

Ha and Chang [1] analyzed composite structures containing distributed piezoelectric sensors and actuators using electric-mechanical coupled finite element method. Tzou and Tseng [2] used piezoelectric hexahedral element in analysis of smart continua, and Tzou and Ye [3]analyzed piezoelectric structures with laminated piezoelectric triangle shell element. Chen and Varadan [4] deduced hybrid finite element formulation for periodic piezoelectric arrays subjected to field loading. Lammering et al.[5] researched optimal placement of piezoelectric actuators in adaptive truss structures based on the electric-mechanical coupled properties of adaptive trusses. Tian et al. [6] used finite element method analyzing buckling and post-buckling of piezoelectric plates. Ding et al. [7] analyzed free vibration of rectangular pyroelectric plates of transverse isotropy based on the three-dimensional coupled piezothermoelasticity theory. Liu et al. [8] formulated the general electric-mechanical coupled dynamic equations of finite element models based on piezoelectricity, elastic mechanics and Hamilton's principle. Li et al. [9] did research on electric-mechanical coupled finite element analysis and experimental study on smart truss structures. Sui and Shao [10] used piezoelectric actuators to strengthen truss structures. Long et al. [11] used smart structure to improve the precision of the structures.

Precision improvement of the spatial structures becomes important since they may encounter stochastic external forces and have errors caused by manufacturing on earth or assembling on the orbits. Considering electric-mechanical coupled properties, optimal shape control of trusses with piezoelectric ceramic actuators is simulated in this paper. The piezoelectric ceramic actuators are regarded as multi-layer stacks. Equations of finite element models for smart trusses are deduced based on Hamilton's principle. The optimal control model of the structures takes voltages as objective.

2. ELECTRIC-MECHANICAL COUPLED FINITE ELEMENT EQUATION

2.1 Micropiezoelectric laminate

Piezoelectric actuators are manufactured of piezoelectric laminates stacked and glued. The sketch is shown as Figure 1. It is parallel connected in electrics

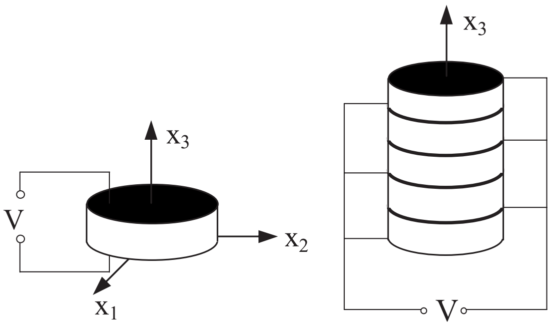


Figure 1. (a) Piezoelectric laminate and (b) piezoelectric stack.

and series connected in mechanics. The polarized direction of the microlaminate is along the thickness and is coincided with axes x_3 and its material is transversely isotropic.

Assuming, $\sigma_3 \neq 0$ and $\sigma_1 = \sigma_2 = \sigma_4 = \sigma_5 = \sigma_6 = 0$, then only strain component ε_3 need to be considered since only deformation in the direction of thickness is concerned.

For the effect of electric field, only E_3 is taken into account. The other components $E_1 = E_2 = 0$ and $\partial E_3/\partial x_1 = \partial E_3/\partial x_2 = 0$. If σ_3, E_3 are taken as independent variables and ε_3, D_3 as attributive variables, piezoelectric equations of microelement can be obtained from linear piezoelectric principle

$$\begin{cases} \varepsilon_3 = c_{33}^E \sigma_3 + d_{33} E_3 \\ D_3 = d_{33}\sigma_3 + \varepsilon_{33}^{\sigma} E_3 \end{cases} \tag{1}$$

The formula above is the piezoelectric equation of the microelement. Intension of electric field $E = V/l_t$, and from the first equation of (1), the output force of the piezoelectric element when its displacement is zero can be expressed as

$$F = \sigma_3 A = -\frac{d_{33} E_3 A}{c_{33}^E} = -\frac{e_{33} A V}{l_t} \tag{2}$$

When axial stress σ_3= 0, the output displacement of the piezoelectric microelement is

$$\delta_3 = l_t \varepsilon_3 = d_{33} E_3 l_t = d_{33} V \tag{3}$$

The output displacement of the piezoelectric microelement is direct proportion to voltage. The proportional constant is d_{33}. Microelements are parallel connected in electrics and series connected in mechanics, which piece number is n. When voltage of every piezoelectric microelement is V, the output displacement of each one, δ_i is

$$\delta_i = \frac{l_t c_{33}^E F_i}{A} + d_{33} V \tag{4}$$

On the ideal condition, each piezoelectric microelement is the same in shape and property, and no energy consumption in it. Axial force: $F_i = F$. The output displacement of each piezoelectric microelement, δ_i, is the same in phase and value, then the total displacement is linear plus of which of each microelement, i.e.,

$$\delta = \sum_{i=1}^{n} \delta_i = n\delta_i = \frac{nl_t c_{33}^E F}{A} + nd_{33}V \tag{5}$$

The output displacement of the actuator is increased to n times of ones in this way (where n is the number of microelements). In electrics, the linear plus relation exists also as the same for electric quantity Q.

$$Q = \sum_{i=1}^{n} Q_i = nQ_i = nAD_3 = nd_{33}A\sigma_3 + n\varepsilon_{33}^{\sigma}AE_3 = nd_{33}F + \frac{n\varepsilon_{33}^{\sigma}AV}{l_t} \tag{6}$$

That is

$$Q = n(d_{33}F + c_s V) \tag{7}$$

where $c_s = \varepsilon_{33}^{\sigma}A/l_t$.

2.2 Field function of piezoelectric stack

Assuming the length, equivalent density and Young's modulus of piezoelectric stack in active members are l_a, ρ_a and Y_a, respectively. The displacements of two ends are u_i and u_j, respectively. The electric potentials of upper and lower plate electrodes of every piezoelectric microelement are ϕ_b and ϕ_a, respectively, as shown in Figure 2. A piezoelectric stack is formed of n pieces of piezoelectric microelements as shown in Figure 1b. Neglecting the thickness of the gluing layer, the plate electrode and the edge effect, the electric field in a piezoelectric stack can be regarded as a uniform electric field. The electric potential is linear distribution in the direction of x_3. The axes x_3 is coincide with the axes of the stack.

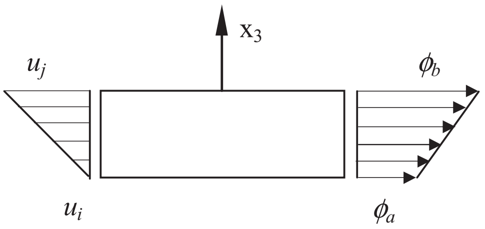


Figure 2. Displacement and electric potential distribution of piezoelectric microelement.

Assuming the displacement function $u(x_3)$ of piezoelectric stack is linear, then

$$u\left(x_3\right)=\frac{u_j-u_i}{l_a}x_3+u_i=\left(1-\frac{x_3}{l_a}\right)u_i+\frac{x_3}{l_a}u_j=\left[N_u\right]\left\{u^{(a)}\right\} \tag{8}$$

where, $[N_u]=\{\,1-x_3/l_a,\ \ x_3/l_a\}$ is the displacement shape function matrix of piezoelectric stack; $\{u^{(a)}\}=\{u_i,u_j\}^T$ is the column matrix of displacement of the stack; And $0\leq x_3\leq l_a$. Corresponding strain is

$$\varepsilon\left(x_3\right)=\frac{\partial u\left(x_3\right)}{\partial x_3}=\left[B_u\right]\left\{u^{(a)}\right\} \tag{9}$$

where $[B_u]=\left\{-1/l_a,1/l_a\right\}$ is strain-shape function matrix.

When $(k-1)l_t\leq x_3\leq kl_t$, then

$$\varphi\left(x_3\right)=\frac{\phi_b-\phi_a}{l_t}x_3-(k-1)\phi_b+k\phi_a=\left[N_\phi\right]\left\{\phi^{(i)}\right\} \tag{10}$$

where $[N_\phi]=\{k-x_3/l_t,\ \ 1-k+x_3/l_t\}$ is electric potential shape function; $\left\{\phi^{(i)}\right\}=\{\phi_a,\ \phi_b\}^T$ are electric potentials of plate electrodes a and b. The electric potential is always continuous, therefore the field intensity distribution is

$$E_3\left(x_3\right)=-\frac{\partial\varphi\left(x_3\right)}{\partial x_3}=-\left[B_\phi\right]\left\{\phi^{(i)}\right\} \tag{11}$$

where, $[B_\phi]=\{-1/l_t,\ 1/l_t\}$ is the shape function of electric field intension. The value of E_3 is unchanged in the range [0, l_a] except ($n-$ 1) discontinuity points (at gluing layers). Therefore the E_3 is regarded as continuous in the range [0, l_a]. This hypothesis is convenient for modelling and almost causes any error.

2.3 Electric-mechanical coupled finite element equations for piezoelectric stack element

Taking the piezoelectric part of the active bar to analysis first, and order: $g_{33}^E=1/c_{33}^E$. Equation (1) transformed as

$$\begin{cases}\sigma_3=g_{33}^E\varepsilon_3+c_d\dot{\varepsilon}_3-g_{33}^Ed_{33}E_3\\ D_3=g_{33}^Ed_{33}\varepsilon_3+c_dd_{33}\dot{\varepsilon}_3+(\varepsilon_{33}^\sigma-g_{33}^Ed_{33}^2)E_3\end{cases} \tag{12}$$

where c_d is viscous damping coefficient of piezoelectric material. Substituting Equations (9), (10) and (11) in Equation (12), we obtained

$$\begin{cases}\sigma_3=g_{33}^E[B_u]\{u^{(a)}\}+c_d[B_u]\{\dot{u}^{(a)}\}-g_{33}^Ed_{33}[B_\phi]\{\Phi^{(a)}\}\\ D_3=g_{33}^Ed_{33}[B_u]\{u^{(a)}\}+c_dd_{33}[B_u]\{\dot{u}^{(a)}\}+(\varepsilon_{33}^\sigma-g_{33}^Ed_{33}^2)[B_\phi]\{\Phi^{(a)}\}\end{cases} \tag{13}$$

From Hamilton' principle

$$\int_{t_0}^{t_1} \delta L \, dt + \int_{t_0}^{t_1} \delta' W \, dt = 0 \tag{14}$$

where $L = T - U$ is Lagrangian function; T is kinetic energy of the system; U is potential energy of the system; $\delta' W$ is virtual work done by external forces. For piezoelectric stack element, kinetic energy is

$$T = \int_{l_a} \frac{1}{2} \rho_a A \left\{ \dot{u}^{(a)} \right\}^T [N_u]^T [N_u] \left\{ \dot{u}^{(a)} \right\} dx \tag{15}$$

Density of potential energy is

$$\begin{aligned} \bar{u} = {} & \frac{1}{2} \{u^{(a)}\}^T [B_u]^T g_{33}^E [B_u] \{u^{(a)}\} + \frac{1}{2} \{u^{(a)}\}^T [B_u]^T c_d [B_u] \{\dot{u}^{(a)}\} \\ & + \frac{1}{2} \{u^{(a)}\}^T [B_u]^T g_{33}^E d_{33} [B_\phi] \{\Phi^{(a)}\} + \frac{1}{2} \{u^{(a)}\}^T [B_u]^T g_{33}^E d_{33} [B_\phi] \{\Phi^{(a)}\} \\ & + \frac{1}{2} \{\dot{u}^{(a)}\}^T [B_u]^T c_d d_{33} [B_\phi] \{\Phi^{(a)}\} + \frac{1}{2} \{\Phi^{(a)}\}^T [B_\phi]^T (\varepsilon_{33}^\sigma - g_{33}^E d_{33}^2) [B_\phi] \{\Phi^{(a)}\} \end{aligned} \tag{16}$$

Total potential energy of the system is

$$U = \int_{l_a} \bar{u} A \, dx \tag{17}$$

Non-potential force work includes mechanical force work and electric force work. Virtual work of the external forces is

$$W^F = \delta \{u^{(a)}\}^T \{F^{(a)}\} \tag{18}$$

where $\{F^{(a)}\} = \{F_i, F_j\}^T$ are axial forces of the two ends. Virtual work of electric forces is

$$W^E = -\delta \{\Phi^{(a)}\}^T \{Q^{(a)}\} \tag{19}$$

where $\{Q^{(a)}\} = \{nQ_i, nQ_j\}^T$. Substitute Equations (15), (16), (17), (18) and (19) into Equation (14), we obtained the electric-mechanical coupled finite element equations for piezoelectric stack element

$$\begin{cases} \left[M^{(a)}\right] \left\{\ddot{u}^{(a)}\right\} + \left[C^{(a)}\right] \left\{\dot{u}^{(a)}\right\} + \left[K^{(a)}\right] \left\{u^{(a)}\right\} + \left[K_{uv}^{(a)}\right] \left\{\Phi^{(a)}\right\} = \left\{F^{(a)}\right\} \\ \left[K_{vu}^{(a)}\right] \left\{u^{(a)}\right\} + \left[C_{uv}^{(a)}\right] \left\{\dot{u}^{(a)}\right\} + \left[K_{vv}^{(a)}\right] \left\{\Phi^{(a)}\right\} = -\left\{Q^{(a)}\right\} \end{cases} \tag{20}$$

where $[K_{vu}^{(a)}] = [K_{uv}^{(a)}]^T$, $[M^{(a)}]$ is the mass matrix of element; $[K^{(a)}]$ is the stiffness matrix of element; $[C^{(a)}]$ is damp matrix of element; $[K_{uv}^{(a)}]$ is the generalized stiffness matrix; $[C_{uv}^{(a)}]$ is the generalized damp matrix; $[K_{vv}^{(a)}]$ is dielectric coefficient stiffness matrix.

2.4 Electric-mechanical coupled finite element equation for structures

2.4.1 Electric-mechanical coupled finite element equation for piezoelectric active member

Assuming that an active member is combined of three parts which an actuator in the middle and host parts at two ends, as shown in Figure 3. It can be divided into three elements: host part elements ①, ③ and piezoelectric element ②. Through Equation (21), the finite element equations of element ② can be obtained:

$$\begin{cases} [M_2]\left\{\ddot{u}^{(a)}\right\} + [C_2]\left\{\dot{u}^{(a)}\right\} + [K_2]\left\{u^{(a)}\right\} + [K_{uv}]\left\{\Phi^{(a)}\right\} = \left\{F^{(a)}\right\} \\ [K_{vu}]\left\{u^{(a)}\right\} + [C_{uv}]\left\{\dot{u}^{(a)}\right\} + [K_{vv}]\left\{\Phi^{(a)}\right\} = -\left\{Q^{(a)}\right\} \end{cases} \tag{21}$$

Assembling the mass matrix, damp matrix and stiffness matrix of elements ①, ② and ③ to member matrixes, respectively, then we obtained:

$$\begin{cases} [M^c]\left\{\ddot{u}^{c)}\right\} + [C^c]\{\dot{u}^c\} + [K^c]\{u^c\} + \left[K_{uv}^c\right]\{\Phi^c\} = \{F^c\} \\ \left[K_{uv}^c\right]\{u^c\} + \left[C_{uv}^c\right]\{\dot{u}^c\} - \left[K_{vv}^c\right]\{\Phi^c\} = -\{Q^c\} \end{cases} \tag{22}$$

For static problem, the derivative item of Equation (22) can be removed and then simplified as

$$\begin{cases} [K^c]\{u^c\} + \left[K_{uv}^c\right]\{\Phi^c\} = \{F^c\} \\ \left[K_{uv}^c\right]\{u^c\} - \left[K_{vv}^c\right]\{\Phi^c\} = -\{Q^c\} \end{cases} \tag{23}$$

As $Q_i = -Q_j$, and assuming the electric potential of plate electrode i is 0, i.e. $\phi_I = 0$, then, $\phi_j = V$. Condense the internal freedoms of Equation (23) according to the equality of internal forces of node i and j, remove the middle

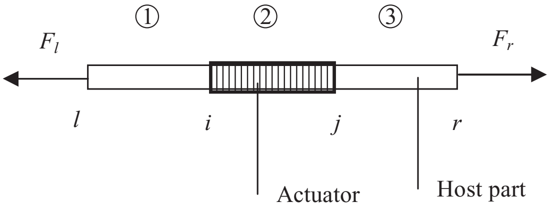


Figure 3. Sketch of an active member.

freedoms, and substitute Equation (22) into Equation (20), obtain the following equations.

$$\begin{cases} [K^e]\{u^e\} + \left[K^e_{uv}\right] V = \{F^e\} \\ \left[K^e_{uv}\right]\{u^e\} - \left[K^e_{vv}\right] V = -\{Q^e\} \end{cases} \tag{24}$$

2.4.2 Electric-mechanical coupled finite element equation for structures

The electric-mechanical coupled finite element equation of piezoelectric active member was obtained in the last section as Equation (24). For general members without active part, the finite element equation for static problem is

$$[K^e]\left\{u^e\right\} = \left\{F^e\right\} \tag{25}$$

Assuming a structure is combined with piezoelectric active members and general members without active part, their finite element equation of members are Equations (24) and (25), respectively. We can obtain the electric-mechanical coupled finite element equation of the whole structure by transforming the finite element equations of members to global coordinates and assembling them into relative coefficient matrixes.

$$\begin{cases} [K]\{X\} + [K_{uv}]\{V\} = \{F\} \\ [K_{uv}]^T\{X\} - [K_{vv}]\{V\} = -[B]^T\left[K_q\right]\{Q\} \end{cases} \tag{26}$$

where, $[K]_{n\times m}$ is the stiffness matrix of the structure; $[K_{uv}]_{n\times m}$ is the general stiffness matrix of the structure; $\{V\}_{m\times 1}$ is the driving voltage column matrix of active members; $[K_{vv}]_{n\times m}$ is the equivalent dielectric stiffness matrix of the structure; $[B]_{n\times m}$ is the position matrix of active members, it is consisted of the direction cosine of active members; $[K_q]_{n\times m}$ is the equivalent coefficient matrix of electric loads; $\{F\}_n$ is the external force column matrix and $\{Q\}_m$ is the electric charge column matrix.

3. OPTIMAL DISPLACEMENT CONTROL OF ELECTRIC-MECHANICAL COUPLED TRUSS

A programming optimal nodal displacement control model is presented utilizing the electric-mechanical coupled property of piezoelectric actuator imbedded, which takes control voltages as control variables and takes precision as objective, and strength are ensured simultaneously.

3.1 Establish the control model

3.1.1 Bring out of the control model

In a general case, a structure has n bars, p bars has imbedded actuator, and q nodal displacement components need to be controlled. Allowable voltages of kth actuator is $-\overline{v}_k \leq v_k \leq \overline{v}_k$; Allowable axial force of it is $\underline{N}_k \leq N_k \leq \overline{N}_k$. A control model is established for problem as:

$$\begin{cases} \text{Find} \quad v_k & (k = 1, 2, \ldots, p) \\ \text{make} \quad \max \left| u_j \right| \quad \rightarrow \quad \min & (j = 1, 2, \ldots, q) \\ \text{s.t.} \quad \sigma_i^c \leq \sigma_i \leq \sigma_i^t; \ \ \underline{N}_k \leq N_k \leq \overline{N}_k; \ \ -\overline{v}_k \leq v_k \leq \overline{v}_k & (i = 1, 2, \ldots, n) \end{cases} \tag{27}$$

where, u_j is the jth nodal displacement component to be controlled; σ_i is the actual stress of ith bar; σ_i^t and σ_i^c are allowable tensile and compressive stresses of ith bar, respectively. This is an undifferentiable programming. The processing of treatment for this control model is discussed as following.

3.1.2 Treatment of the control model

Model (27) cannot be solved directly because firstly, the objective function and stress constraint functions are implicit functions to control variables, and secondly, the model is an un-differentiable programming. Therefore a series of treatment is needed.

Explicit expression of nodal displacement and internal force of bars. For a linear structure, the relation between nodal displacement and internal force of the bar to the control voltages can be expressed as

$$u_j = u_j^0 + \sum_{k=1}^{p} u_{jk} \cdot v_k \tag{28}$$

in which, $u_{jk} = \partial u_j / \partial v_k$ is sensitivities of displacement; u_j^0 denotes nodal displacement component while all voltages are 0. Internal force of the bar is

$$N_i = N_i^0 + \sum_{k=1}^{p} n_{ik} \cdot v_k \tag{29}$$

where, $n_{ik} = \partial N_i^e / \partial v_k$ is Sensitivities of internal force; N_i^0 denotes internal force of element while all the voltages are 0.

Treatment of objective function. In order to transform the undifferentiable programming to an easy solved differentiable programming, treat the objective function of model (27). Introduce a new variable ε, make

$$-\varepsilon_j \le u_j \le \varepsilon_j \quad (j = 1{,}2, \text{ and } q) \tag{30}$$

For whole structure,

$$\varepsilon = \max_{(j\in q)} (\varepsilon_j) \tag{31}$$

Transformation of constraints. Assume the allowable tensile and compressive stresses of host part material of ith bar are σ_i^t and σ_i^c, respectively, on account of the axial force of the host part of an active bar can be expressed as the product of stress and sectional area, $\sigma_i \cdot A_i$, the allowable axial force of an active bar is

$$\begin{cases} N_i^t = \min\left\{\sigma_i^t \cdot A_i,\ \overline{N}_i\right\} & (N_i \ge 0) \\ N_i^c = \max\left\{\sigma_i^c \cdot A_i,\ \underline{N}_i\right\} & (N_i < 0) \end{cases} \tag{32}$$

The stress constraints and axial force constraints of actuators are transformed as

$$N_i^c \le N_i^0 + \sum_{k=1}^{p} n_{ik} \cdot v_k \le N_i^t \tag{33}$$

3.1.3 Formation of the control model

Through the deducing and transforming above, substitute explicit expression of constraints into control model (27), yields

$$\begin{cases} \text{Find} & v_k,\ \varepsilon \quad (k = 1, 2, \ldots, p) \\ \text{make} & \varepsilon \ \rightarrow \ \min \\ \text{s.t.} & \sum\limits_{k=1}^{p} u_{jk} \cdot v_k - \varepsilon \le -u_j^0;\ -\sum\limits_{k=1}^{p} u_{jk} \cdot v_k - \varepsilon \le u_j^0 \ (j = 1, 2, \ldots, q) \\ & \sum\limits_{k=1}^{p} n_{ik} \cdot v_k \le N_i^t - N_i^0; - \sum\limits_{k=1}^{p} n_{ik} \cdot v_k \le -N_i^c + N_i^0 \ (i = 1, 2, \ldots, n) \\ & v_k \le \overline{v}_k, \quad -v_k \le \overline{v}_k, \quad \varepsilon \ge 0 \end{cases} \tag{34}$$

Model (34) is a standard linear programming model, through which ε and v_k can be obtained.

4. NUMERICAL EXAMPLE AND DISCUSS

A 10-bar static indeterminate truss is shown in Figure 4. For the material in the host part of the bars, the Young's modulus is $E = 70$ Gpa. The allowable

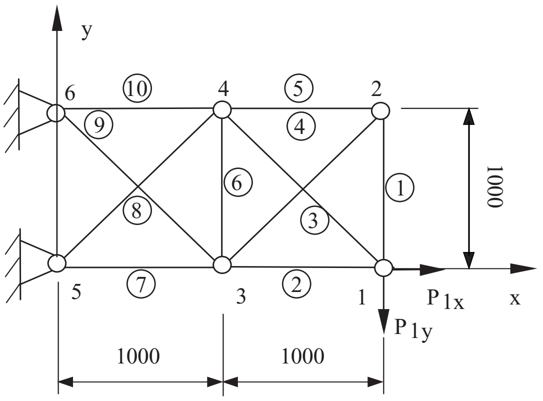


Figure 4. 10-bar static indeterminate truss.

tensile and compressive stresses are $\sigma^t = 30$ MPa and $\sigma^c = -30$ MPa, respectively. The cross-sectional area of each bar is 1cm^2. The Young's modulus of the piezoelectric material is $E = 63$ GPa. The piezoelectric actuator is formed of 400 piezoelectric pieces. The thickness of each one is 0.5 mm. The allowable tensile and compressive axial forces of it are 3 kN and −3 kN, respectively. The nodal displacement component of node 1 along y-direction requires to be minimized. Maximum allowable voltage is 500 V. Two cases are considered here: (1) Built-in actuators are installed in all bars. (2) Only one bar has a built-in actuator. The main results are listed in Table 1 and Table 2.

In case (1), when there is no actuator in the structure, the nodal displacement component of node 1 along y-direction is 0.3101 mm, maximum stress in the structure is 14.0 MPa. Main results are listed in Tables 1 and 2.

Table 1 shows that when built-in actuators are installed in all bars, the control model can make the controlled nodal displacement to be zero. The maximum stress in the structure is decreased. When only one built-in actuator in the structure and its maximum allowable voltage is limited in 500 V, the effect of the control varies with the position of the actuator. Table 2 shows that, it has the best control effect when the actuator is installed in bar ⑦. Next, when the actuator at certain position, some improvement of precision may cause

Table 1. Main results of case 1.

	Max control voltage of actuators/(V)	Max stress in the structure/(MPa)	Max value of controlled nodal displacement component/(mm)
Before control	0	14.0	0.3101
After control	357	13.5	0

Table 2. Controlled results when only one actuator in the structure.

Actuator at the bar number	Controlled disp. when defm. of the actuator is no limited/(mm)	Max stress when defm. of the actuator is no limited/(MPa)	Controlled disp. when max voltage of the actuator is 500 V/(mm)	Max stress when max voltage of the actuator is 500 V/(MPa)
1	0	24.9	0.2765	13.8
2	0	25.2	0.2677	14.2
3	0	25.2	0.2541	14.2
4	0	24.9	0.2636	13.8
5	0	24.9	0.2765	13.8
6	0.2756	30.0	0.3055	12.5
7	0	15.1	0.2020	12.4
8	0	27.9	0.2580	16.3
9	0	28.6	0.2598	11.8
10	0	18.7	0.2049	15.6

the maximum stress increasing. If the control voltage of the actuator is not limited or the piezoelectric constant can changes as needed to make the active deformation of the actuator large enough, the controlled nodal displacement component can be reduced to zero by one actuator except when the actuator is at the position of bar ⑥. But the maximum stress in the structure increases to different extent.

5. CONCLUSION

An optimum control model was constructed, which takes the maximum controlled displacement component of piezoelectric truss as objective, and takes the controlled voltage of actuators as variables, subjects to strength requirement of the structure and characteristic of the actuator. When the number of actuator is limited but distribution of it is rational and the actuator has enough active deformation, the optimal control can be realized still. The models can realize displacement control of single point or multiple points effectively. For static determinate structure, control will not cause varying of stresses.

ACKNOWLEDGEMENTS

Supported by National Natural Science Foundation of China (10472003), Beijing Natural Science (3042002) and Beijing Educational Committee (KM200410005019) Foundations

REFERENCES

1. S.K. Ha and F.K. Chang (1992), Finite element analysis of composite structures containing distributed piezoelectric sensors and actuators. *AIAA Journal*, 30, 3, pp. 772–780.
2. H.S. Tzou and C.I. Tseng (1994), Thin piezoelectric hexahedron finite element applied to design smart continua. *Finite Element in Analysis and Design*, 16, 1, pp. 27–42.
3. H.S. Tzou and R. Ye. (1996), Analysis of piezoelectric structures with laminated piezoelectric triangle shell element. *AIAA Journal*, 34, 1, pp. 110–115.
4. L.C. Chen and V.K. Varadan (1994), Hybrid finite element formulation for periodic piezoelectric arrays subjected to field loading. *International Journal of Numerical Methods in Engineering*, 37, pp. 2987–3003.
5. R. Lammering, J.H. Jia and C.A. Rogers (1994), Optimal placement of piezoelectric actuators in adaptive truss structures. *Journal of Sound Vibration,* 171, 1, pp. 67–85.
6. X.G. Tian, Y.P. Shen and J.X. Gao (2000), Buckling and post buckling analysis of piezoelectric plates using finite element method. *ACTA Mechanica Solida Sinica*, 21, 2, pp. 123–130 (in Chinese).
7. H.J. Ding, F.L. Guo and P.F. Hou (2000), Free vibration of simply supported transversely isotropic pyroelectric rectangular plates. *ACTA Mechanica Sinica*, 32, 4, pp. 402–411.
8. Z.X. Liu, Y.W. Yang and W. Cai, and Q.S. Li (1997), Electromechanical coupled finite element and dynamic equations. *Journal of Shanghai Jiaotong University*, 31, 7, pp. 54–59 (in Chinese).
9. J.B. Li, J.M. Ge, and Q.F. Yang (1998), Electromechanical coupled finite element analysis and experimental study on intelligent truss structure. *Journal of Taiyuan University of Technology*, 29, 6, pp. 577–581 (in Chinese).
10. Y.K. Sui and J.Y. Shao (2002), Characteristics of strength control of adaptive structure with electromechanical coupling. *ACTA Mechanica Solida Sinica,* 1, 15, pp. 49–56 (in Chinese).
11. L.C. Long, Y.K. Sui, and B.R. Ye (2003), Optimal control method of intelligent space structures. *Journal of Beijing University of Technology*, 29, 2, pp. 133–137 (in Chinese).

PROTEIN SECONDARY STRUCTURE PREDICTION METHODS BASED ON RBF NEURAL NETWORKS

N. Jing[1,2], B. Xia[1], C.G. Zhou[3] and Y. Wang[3]
[1]*Key Laboratory of Marginal Sea Geology, Guangzhou Institute of Geochemistry & South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China*
[2]*Graduate School of CAS, Beijing, China*
[3]*College of Computer Science & Technology, Jilin University, Changchun, China*

Abstract To solve the complicated non-linear mode-sorting problem of protein secondary structure prediction, the chapter proposed a new method based on radial basis function neural networks and learning from evolution. It also discussed the influence of data selection and structure design on the performance of the networks. The results indicate that this method is feasible and effective.

Keywords: protein secondary structure, radial basis function neural networks, evolutionary information, amino acids sequence.

1. INTRODUCTION

Protein secondary structure prediction is one of the most important tasks in bioinformatics. Sequence determines structure determines function. Polypeptide chains can fold into complicated 3D structure in correspondence to its function. The key step in predicting the folding of a protein is to predict its secondary structure. Unfortunately, experiments aimed at extracting such information cannot keep pace with the rate at which raw sequence data are being produced. Computational techniques allow for biological discovery based on the protein sequence itself or on their comparison to protein families. Given a protein sequence, the secondary structure prediction problem is to predict whether each amino acid is in a helix, strand or neither. H, E and C represent helix, strand and non-routine structure, respectively. Some stretches of sequence show a particular preference to be in one of these three states.

G. R. Liu et al. (eds), Computational Methods, 1037–1043.

One of the most popular secondary structure prediction algorithms is PHD. Using BP networks, the algorithm takes multiple alignments of protein sequences as input. The gaining upon ability, sorting ability and learning velocity of radial basis function neural networks are all better than BP networks. This study discussed how to use radial basis function neural networks to predict protein secondary structure. Further, it analysed how to use evolutionary information to enhance the prediction accuracy.

2. EVOLUTIONARY INFORMATION

There is long-range information in multiple sequence alignments. Some residues can be replaced by others without changing structure. However, not every amino acid can be replaced by any other. On the contrary, one evolutionary step (exchange of one residue) can destabilize a structure. Residues substitution patterns observed in protein families are highly specific for a particular structure, and thus, contain more information about structure than single sequence.

Multiple alignments can produce position specific profiles. Profiles are, quite simply, a numerical representation of a multiple sequence alignment. Imbedded within a multiple sequence alignment is intrinsic sequence information that represents the common characteristics of that particular collection of sequences, frequently a protein family. By using a profile, one is able to use these imbedded, common characteristics to find similarities between sequences with little or no absolute sequence identity, allowing for the identification and analysis of distantly related proteins. In the predictive methods using evolutionary information such as PHD, the prediction accuracy is improved effectively.

3. METHODS

3.1 Constructing appropriate database

Constructing a database for this study is very important because of the following two points. First, the records in PDB are in detail, so the corresponding part must be extracted. Second, since the training and testing sets must be big enough, a number of proteins must be chosen from PDB. Choosing effective training set is the basis of this research. Some sequences mainly consist of one type of the structures, so we must be careful when choosing the training set. To achieve higher prediction accuracy, the training set must be big enough and include all kinds of structures in proportion. In this study 106 proteins were chosen to construct the training and testing set. Prediction of secondary structure makes sense for proteins, which are not homologous to any other proteins with known 3D structure; otherwise, the secondary structure can be predicted by homology with higher accuracy than by any existing secondary structure prediction algorithm.

3.2 A prediction method using second level RBFNN

Radial basis function neural networks are feed-forward networks. RBF networks stand for a class of neural network models, in which the hidden units are activated according to the distance between the input units. RBF networks combine two different types of learning: supervised and unsupervised. In this study we use supervised learning. Radial basis functions are characterized by the fact that their response decreases (or increases) monotonically with distance from a central point. The hidden neurons represent computing units that perform a non-linear transformation on the input vector by means of radial basis functions that constitute a basis for this transformation.

Basically there are twenty kinds of amino acids. Before feeding the amino acids sequence into neural networks, they need to be changed into binary code of 21 units (each unit representing one of the amino acids or spacer). The training of the cluster layer is carried out first. Once completed, the output layer is subjected to supervised learning as used in the feed-forward network with backpropagation. In order to train the network, several parameters need to be specified. The maximum learning rate parameter ($\mathbf{L}_{\text{max}}$) ranging between 0 and 1, the minimum learning rate parameter ($\mathbf{L}_{\text{min}}$) ranging between 0 and 1 and the maximum number of epochs used for training (**N**). They are used to establish the learning gradient. This gradient is needed to update the weights of a cluster node whose input weight vector has the minimum squared Euclidean distance to the input vector pattern. The maximum learning rate also serves as the initial learning rate. In updating the weight of each dimension, the following formula is applied:

$$\mathbf{W}_{\text{new}} = \mathbf{W}_{\text{old}} + [\mathbf{L} \times (\mathbf{I}_{\text{value}} - \mathbf{W}_{\text{old}})] \tag{1}$$

indent $\mathbf{W}_{\text{old}}$ represents the weight before updating. $\mathbf{W}_{\text{new}}$ represents the updated weight. **L** is learning rate. $\mathbf{I}_{\text{value}}$ is input pattern value. After all input patterns have been run through the network (this constitutes an epoch), the learning rate itself is updated via:

$$\mathbf{L}_{\text{new}} = \mathbf{G} \times \mathbf{L}_{\text{old}} \tag{2}$$

G is learning gradient. The gradient is determined by the following fomula:

$$\mathbf{G} = \mathbf{L}_{\text{max}} - \mathbf{E}_{\text{complete}} \times (\mathbf{L}_{\text{max}} - \mathbf{L}_{\text{min}}) \tag{3}$$

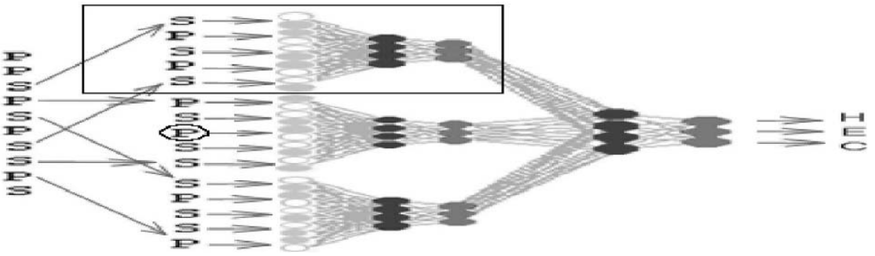


Figure 1. Simplified demonstration of the method based on second level RBFNN.

$\mathbf{E}_{\text{complete}}$ means epochs completed. The weight correction $\mathbf{w}_{\text{ij}}$ of neuron i and j is calculated by backpropagation:

$$\mathbf{w}_{\text{ij}} = \mathbf{L}_{\text{constant}} \times \mathbf{G}_{\text{local}} \times \mathbf{I}_{\text{j}} \tag{4}$$

The learning rate constant $\mathbf{L}_{\text{constant}}$ and the non-normalized minimum average squared error of BP algorithm must be specified. $\mathbf{G}_{\text{local}}$ represents local Gradient. $\mathbf{I}_{\text{j}}$ means input Signal of neuron j.

Window size is an important parameter, which means the number of residues entering the networks every time. In order to take more consideration of the correlation between adjacent residues, second level RBF neural networks is introduced. The first level changes sequences into structure. The second level uses the output of the first level as input, changing structures into final results.

Figure 1 is simplified demonstration, in the actual predicting process each amino acid in the sequence should be changed into 21 units, not just P and S as shown in the figure. That is to say, the 20 kinds of amino acids and a spacer are simplified into two kinds. Let's see the rectangle area first. In demonstration the window size is 5, but actually the window size is 15. Every time adjacent 15 residues enter the sub-network as a group. The output is the corresponding secondary structure of the central residue, in Figure 1 S. Now see the whole figure. In the first level window size of the 3 sub-networks (in Figure 1 it is 3, actually it is 11) is all 15 and their output is the central residue (from top to bottom S, P, S). In the second level there are 11 structures of adjacent 11 residues as input, again the output is the secondary structure of the central residue (shown in Figure 1 the residue P with a black circle around it). In the actual predictive process the sequences slide through the window, every adjacent 15 × 11 residues are as an whole input group. For every group of 15 × 11 residues the corresponding output represents H, E or C – the secondary structure of the central residue.

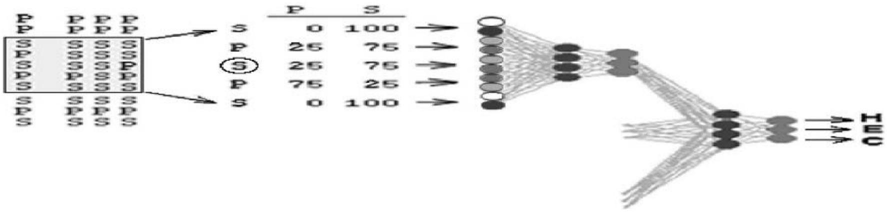


Figure 2. Feeding evolutionary information into a neural network system.

3.3 Using evolution information

Multiple alignments can produce position specific profiles, which describe crucial information about structure and reflect evolutionary constraints. The multiple alignment is converted into a profile: for each position, the vector of amino acid frequencies is calculated based on the alignment. The neural net is applied sequentially to all protein sequence positions to predict secondary structure in every position as a state from three possible alternatives. Prediction at a given position depends on amino acid frequencies (in the profile) at that position and neighbouring positions within a range defined by the window for which inputs are collected.

Use profile as input instead of binary code this time. When similar sequences are found, a profile-base multiple sequence alignment is generated.

Figure 2 is also simplified demonstration. A sequence family is aligned (shown are the sequence of unknown structure and three aligned relatives). For each sequence position of profile is compiled that gives the percentage of S or P in the alignment (shown in centre for window of five adjacent residues). Instead of using binary input units (0 or 1), now the profile is fed into the first neural network. Then the output is again fed into a second level network. The following steps are the same as I have explained before (same as Figure 1).

4. RESULTS

The first step of this chapter is to use second level RBF neural networks to predict protein secondary structure. The window size of the second layer is decided by experimenting. The results show that the performance does not always become better with the increasing window size. When the window size of the sub-networks in the first layer is 15 and the window size of the second layer is 11, the performance is the best. When the window size is too small, the performance will be influenced. Table 1 shows the results of the method based on second level RBF neural networks.

Table 1. The prediction results of second level RBFNN.

Second level RBF neural networks			
Prediction accuracy	Cα	Cβ	Ccoil
68.9%	0.56	0.42	0.46

Table 2. The enhanced prediction results by using profiles.

Second level RBFNN using profiles			
Prediction accuracy	Cα	Cβ	Ccoil
71.7%	0.64	0.51	0.49

This study also tried to use evolutionary information to improve prediction accuracy. The second step, use PSI-BLAST program to produce profiles that are fed into the neural networks. It can improve the prediction accuracy by several percent. The window size of layer is also decided by experimenting.

5. SUMMARY

This chapter used radial basis function neural networks to develop the protein secondary structure prediction system and has made some achievements. The results show that consulting more evolutionary information can improve the overall performance. The prediction accuracy of the model averaged about 72%, showing the feasibility and validity of the algorithm.

6. ACKNOWLEDGEMENTS

Supported by the National Natural Science Foundation of China (No. 60175024) and Knowledge Innovation Program of CAS (KZCX-SW-117).

REFERENCES

1. D.T. Jones (1999), Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*, 292, 2, pp. 195–202.
2. L.S. Wang and Y. Xu. (2003), SEGID: identifying interesting segments in (multiple) sequence alignments. *Bioinformatics*, 19, pp. 297–8.
3. L.J. McGuffin, K. Bryson, and D.T. Jones (2000), The PSIPRED protein structure prediction server. *Bioinformatics*, 16, pp. 404–5.

4. D. Przybylski and B. Rost (2002), Alignments grow, secondary structure prediction improves. *Proteins*, 46, pp. 197–205.
5. B. Rost (2002), Rising accuracy of protein secondary structure prediction. In: *Protein structure determination, analysis and modeling for drug discovery*, D. Chasman (ed.), Dekker, New York, pp. 207–49.
6. B. Rost and C. Sander (1994), Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins*, 19, pp. 55–72.
7. G.H. Yang, C.G. Zhou, et al. (), Algorithm for predicting protein secondary structure based on improved Bayesian networks. *Progress in Natural Science (Chinese Edition)*, 13, 6, pp. 667–70.

A HYBRID META-HEURISTIC FOR A ROUTING PROBLEM

Jesús Fabián López Pérez
Post-Graduate Program of Management Science, FACPYA UANL, Monterrey, México

Abstract Our problem is about a routing of a vehicle with pickup and delivery of product with time window constraints. This problem requires to be attended with large scale instances (nodes $\geq$ 100). A strong active time window exists ($\geq$90%) with a large factor of amplitude ($\geq$75%). The problem is NP-hard and for that reason the application of an exact method is limited by the computational time. We propose a hybrid methodology which offers good solutions in computational times that do useful its application.

Keywords: logistics, genetic algorithms, time windows, NP-hard.

1. PROBLEM DEFINITION

The delivery and pickup of product with time window constraints can be seen in two principal variants. We have the variant for one single vehicle (SPDP-TW) and by the other hand we have the problem for multiple vehicles (MPDP-TW), Savelsbergh 1995 [1]. Our problem is focused on the first case (SPDP-TW). The PDP-TW is more difficult to be solved than the VRP-TW due to the PDP-TW is a generalization of the VRP-TW, Palmgren 2001 [2]. The objective is to determine the optimal route for a distribution vehicle. A route is defined as the arrival sequence that we require to reach a group of clients. An optimal route is obtained if we achieve to visit all the clients with a minimal cost (or distance or time). The problem constraints are as follows:

a. We have a vehicle leaving from a distribution center. This vehicle attends a group of geographically scattered clients and then returns to the origin point.
b. Each client visited has a requirement of product to be delivered and a load to be collected. We have to observe a finite load capacity for the vehicle all the time.

G. R. Liu et al. (eds), Computational Methods, 1045–1050.

c. We define a cost matrix which identifies the time or distance required to go from each client to all others. The delivery and pickup time for each client is a constant.
d. It is not permitted to arrive before the opening hour neither after the closing hour.

The bibliographical review of routing problems includes Applegate et al. [3]; Dumas & Solomon [4]. The outlined problem is combinatoric in nature and is catalogued as NP-Hard, Tsitsiklis [5]. Regarding routing application aspect, the less investigated variant is the one which has to do with the physical product distribution, Mitrovic [6]. The difficulty for the SPDP-TW depends strongly on the structure of the time windows that are defined for each customer. We refer here the work presented by Ascheuer et al. [7] for the TSP-TW problem. They proved that this problem is particularly difficult to be solved for instances with more than 50% of active nodes with time window constraints. They tested with instances up to 233 nodes. All the instances greater than 70 nodes required more than 5 hours of computational time. Base on their computational experience, they conclude that the instances on the limit up to 70 nodes can be solved to optimality by Branch & Cut algorithms (B&C).

2. METHODOLOGY PROPOSED

Our methodology proposes six routines. We have four preprocessing routines, then the Genetic Algorithm and finally one routine for post-processing. Below we expose:

1. Network topology decomposition phase based on a '*shortest path algorithm (SPP)*:' we consider here the topology corners that are required to model the traffic constraints for the network. We can establish that, if we setup a network with $N1$ nodes, we would obtain only $N2$ nodes, where $N1 \approx 4N2$. The constant '4' related with the quantity of network arcs that we require to model a common street corner. With this in mind, we avoid to include on the network as many nodes as street corners. Instead of the previous thing, we use an SPP algorithm to pre-calculate the optimal sub-tour required to move from each customer to all the others. All these preprocessed sub-tours fill the $N2$ cost matrix that will be used in the next phases.
2. Compressing and clustering phase through a '*neighbourhood heuristic*': the $N2$ nodes are grouped to setup a reduced quantity of $N3$ meta-nodes *(where: $N3 < N2$)*. We require some compatibility for the time window structures of the $N2$ nodes that are considered to be grouped in a meta-node. Starting from a group of nodes to be grouped in a meta-node, the time window structure of this meta-node is defined by the latest opening time and by the earliest

closing time. We use in the algorithm a 50% compression factor for the grouping phase which means that $N2 = 2 * N3$.

3. Discriminate compressing phase through a '*k nearest nodes heuristic*': the network arcs with greater cost are eliminated from the matrix. The logic of the previous thing is because of those arcs have a smaller probability to appear in the optimal solution. For each $N3$ node in the network, we maintain only the *'k'* arcs with the smallest cost, where $k << N3$. We use a conservative 20% discriminate factor in order to maintain the optimal solution within the search space. This empirical assumption means that the matrix that will be transferred to the next phase will be dimensionally reduced and defined by $N4 \times N3$, where $N4 = 20\% * N3$.
4. Aggressive Branch and Cut phase: starting with $N3$ meta-nodes, the objective is to find as quickly as possible, the first feasible solution that cover the time window and vehicle capacity constraints. The logic that we apply here is to iteratively generate cuts within a Branch and Cut scheme. For that purpose we identify in the incumbent solution, the node that has the greater deviation in relation to the time window and/or the vehicle capacity constraint. This node is named '*pivot node*'. Then we verify the nodes of the tour that can be identified as '*related*' in order to re-sequence the position of the *pivot node* within the tour. The logic that we apply here to generate the cut assures that the *pivot node 'k'* use at least one of the arcs that connect it to one of the *related nodes 'j'*. This procedure continues until is found the first feasible solution.

$$
\begin{aligned}
\ni I = \{1..N_3\}&(network\ \ notes)\\
&K \in I(pivot\ \ node)\\
&j \subseteq I\{1..m\}(related\ \ nodes\ \ to\ \ k)\\
&\Sigma_{j=1}^{m}(x_{jk} + x_{kj}) \geq 1 \quad \forall k \subseteq I
\end{aligned}
$$

5. Evolutionary phase: our objective here is to approximate the optimal solution for the compact version of the network. Maintain in the pool of constraints a cut unnecessarily, means to take out the optimal solution or at least a better solution, from the search space. Our computational experience indicates that the quantity of cuts that get to be accumulated in the pool is meaningful (15–40 cuts). The goal is to identify which cuts of the pool are necessary to be eliminated. Identify which cuts must be eliminated, can be seen as a combinatoric sub-problem. We propose an evolutionary strategy to attend this sub-problem. A binary codification permits to represent the elimination (0) and the presence (1) of a cut in the pool. The GA reproduction method applies two random crossing points throughout the chromosome length. The mutation factor is initialized with a 5% value and it is auto-adjusted. Upon increasing the degeneracy level in the population we apply an exponential growth curve in the mutation factor with 50% as an asymptotic limit.

6. Uncompressing phase to disaggregate the proposed route for the original network: Here we have two routines. The first routine is focused in determining the optimal sequence in which the *N3* meta-nodes should be disaggregated to return to the *N2* nodes obtained in phase 2. The previous thing is achieved through a low scale scheduling algorithm which mathematically can be solved to optimality in a feasible time. This problem is referred as a low scale because we treated only with those valid arcs that must be modeled in order to take advantage of the previous meta-nodes sequencing. Starting from a selected meta-node, we construct only the valid arcs to the previous and to the next meta-node. In that sense, if we deal with the first 3 meta-nodes of the network, for example we could identify 2, 5 and 3 nodes that are grouped in those meta-nodes respectively. That means that our math formulation should constraint that the first 2 nodes must be scheduled in the first 2 positions, the next 5 nodes in the next 5 positions and so on. The second routine makes use of the topology information generated in the first phase of our methodology. Its objective is to substitute the sequence of the tour defined by the *N2* nodes according to the cardinal movements that are required to obtain the *N1* nodes of the original network.

3. EXPERIMENTAL DEVELOPMENT AND RESULTS

We propose that our methodology (GA) is viable to find good solutions (***optimality*** $\geq 90\%$) and with a low computational cost (***minutes*** ≤ 5) to solve the SPDP-TW problem. We calculate a 'percentage of optimality' which is measured taking in mind the solution registered by a B&C Algorithm (Control Group) as follows:

$$optimality\ \% = 1 - \frac{GA\ \ solution\ \text{-}\ Lower\ \ Bound}{Lower\ Bound}$$

$$\begin{aligned} where\ \ Lower\ \ Bound = {} & Best\ \ solution\ \ reached\ \ by \\ & B\ \&\ C\ \ within\ \ 5\ \ hours\ \ limit \end{aligned}$$

Our hypothesis will be attended through a 'T' Student test of means differences.

$$\begin{aligned} &Ho : \mu\ Experimental\ \ GA \leq 90\%\ (\mu Control\ \ Group) \\ &Ha : \mu\ Experimental\ \ GA > 90\%\ (\mu Control\ \ Group) \end{aligned}$$

$$t = \frac{\bar{D}-\mu_D}{\frac{S_D}{\sqrt{n}}}, \quad where : \bar{D} = \frac{\sum_{i=1}^{n} D_i}{n}, \quad y\ S_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \bar{D})^2}{n-1}}$$

Table 1. Matrix Computation for the T statistic values.

		Algorithmic instruments to be compared					
		B&C Algrithm (Control Group)	Basic Genetic Algorithm (Evolver)	Basic Genetic Algorithm (Frontline)	Proposed Genetic Algorithm		
					P(x > 90%)	P(x > 92.5%)	P(x > 95%)
Measurements	3th Minute	NA	–0.4039	–0.5581	2.4261	1.0688	–0.0911
made after ‘x’	5th Minute	–2.4260	–0.1159	–0.3068	3.3130	1.5392	0.1111
Computational	8th Minute	–1.2800	0.1616	0.3173	4.8510	4.1050	0.8300
Time	10th Minute	–0.7001	0.4001	0.9030	6.2980	5.5770	1.3277

We applied an ‘*Experimental Design*’ through the use of four instruments:

(a) B&C Algorithm: optimality reference used as control group.
(b) Steady Sate Genetic Algorithm #1: Evolver©engine Ver 6.0 (Palisade).
(c) Generational Genetic Algorithm #2: Solver©engine Ver 4.0 (Frontline).
(d) Proposed Genetic Algorithm #3.

The ‘percentage of optimality’ was applied on minute 3, 5, 8 and 10. Only the B and C instrument was limited up to five hours of computational time. The experimental design was applied for a sample of 40 instances. We define: (1) Mean sampling (m) and (2) Standard deviation sampling (s). The T Student test ‘$P(x > 90\%)$’ applied for each element (m_{ij}, s_{ij}) calculates the probability that the *algorithmic instrument ‘j’* in the *time interval ‘i’* obtains at least a 90% of optimality. We tested instances with more than 70% of active time windows and with a minimal wide of 75%. The dimension of the tested instances are defined by w, where ($100 \leq w \leq 120$). The parameters applied for the implementation of the GA’s # 1, 2 and 3 were fixed empirically to the same values. Table 1 shows the values calculated for the ‘T’ statistic. Table 2 shows the probability coefficients ‘P’.

Table 2. Matrix Computation for the T statistic values.

Algorithmic instruments to be compared					
B&C Algrithm (Control Group)	Basic Genetic Algorithm (Evolver)	Basic Genetic Algorithm (Frontline)	Proposed Genetic Algorithm		
			P(x > 90%)	P(x > 92.5%)	P(x > 95%)
NA	34%	29%	99%	85%	46%
<1%	45%	38%	100%	93%	54%
10%	56%	62%	100%	100%	79%
24%	65%	81%	100%	100%	90%

4. DISCUSSION AND CONCLUSIONS

(a) The B and C obtains better solutions than any of the GA's and when the time is allowed the optimal solution to the problem (38 of 40). We report times below three minutes for instances with less than 70 nodes and with less than 60% of active time windows.

(b) "Out of the shelf" GA's (1 and 2), are inferiors since these never surpass 90% of optimality before the 10th minute. At the end of the 3rd minute GA #1 offers better performance. The GA #2 obtains better solutions before the 10th minute. Our proposed GA #3 obtains better solutions than the two others. We take advantage of the generational reproduction method of the GA # 2 and furthermore our methodology exposed in Section 2 speeds up the generation of good solutions in early iterations.

(c) We establish that the proposed methodology reaches a percentage of optimality $\geq$90% in a computational time $\leq$5 minutes (100% confidence).

(d) The proposed GA offers solutions within an acceptable optimality range and with computational times that make feasible its implementation in business environment:

- ➢ Optimality $\geq$90% in a computational time $\leq$3 minutes (99% confidence).
- ➢ Optimality $\geq$92.5% in a computational time $\leq$5 minutes (93% confidence).

(e) However, our methodology can assure only 54% of confidence when is required to reach a percentage of optimality $\geq$95% in a computational time $\leq$5 minutes.

REFERENCES

1. M. Savelsberg (1995), Local search in routing problem with time windows, Annals of Operations Research, Rotherdam, the Netherlands, *Journal of Heuristics*, pp. 285–305.
2. M. Palmgren (2001), A column generation algorithm for the log truck scheduling problem. Department of Science and Technology, LiTH-MAT-R, Linköping University, Norrköping Sweden, pp. 3–15.
3. D. Applegate, P. Bixby, and V. Chvátal (1998), On the solution of traveling salesman problems, Documenta Mathematica Extra Volume ICM III, Journal der Deutschen Mathematiker-Vereinigung, pp. 645–656.
4. Y. Dumas, J. Desrosiers, and M. Solomon (1995), An algorithm for the traveling salesman problem with time windows. *Operations Research USA*, 43 (2), pp. 367–371.
5. J. Tsitsiklis (1992), Traveling salesman and repairman problems with TW, *Networks*, *USA*, 22, pp. 263–282.
6. Mitrovic, Snezana (1998), Pickup and Delivery Problem with Time Windows. In: *Technical Report SFU CMPT TR 1998–12*, Journées de l'optimization School of Computing Science, University, Burnaby, BC Canada, pp. 38–39.
7. N Ascheuer, M Fischetti, M. Grostchel, (2001), *Solving ATSP with Time Windows by Branch-And-Cut*. Springer-Verlag, Germany. pp. 475–506.

IDENTIFICATION OF GEOMETRIC PARAMETERS OF DRAWBEAD USING NEURAL NETWORKS

G.Y. Li[1], L.F. Han[1,2], X. Han[1] and Z.H. Zhong[1]
[1] *Key Laboratory of Advanced Technology for Vehicle Body Design and Manufactory of M.O.E., Hunan University, Changsha 410082, China*
[2] *College of Mechanical Engineering, Xiangtan University, Xiangtan 411105, China*

Abstract In this paper, a neural network (NN) model was designated to identify the geometric parameters of drawbead in sheet forming processes. The genetic algorithm (GA) was used to determine the neuron numbers of the hidden layers of the neural network, and a sample design method with the strategy of updating training samples was also used for the convergence. The NN model goes through a progressive retraining process and the numerical study shows that this technique can give a good result of the parameter identification of drawbead.

Keywords: drawbead, neural network, genetic algorithm, sample design.

1. INTRODUCTION

In sheet forming processes, the drawbeads play an important role in control of the material flow. Traditional drawbead design mainly depends on designers' experience. Moreover, the drawbead still needs to be adjusted to obtain the proper restraining force during die try-out, which is time-consuming. The precise drawbead design can be solved using the finite element method (FEM). Previous work on the simulation of drawbead in finite element analysis can be classified into two kinds. One is to build the real geometric model of the drawbead. The other is to employ the equivalent drawbead model instead of the real one. The equivalent drawbead model is added to finite element model through a series of discrete nodes along a drawbead curve. These discrete nodes, on which the applied forces were drawbead resistance force (DBRF) and upward

G. R. Liu et al. (eds), Computational Methods, 1051–1056.

force, were attached to die or binder. Many researchers have been working on this subject to find empirical [1] or analytical formulas [2,3] for drawbead restraining forces. Other researchers have also made studies on the influence of geometric parameters [4] and process parameters [5] of drawbead to sheet metal forming.

In this work, the geometric parameters of drawbeads are obtained in an inverse way. The neural network combined with GA is introduced to identify the geometric parameters of drawbead. As we know, the computational cost mainly depends on the structure of NN model and the number of training samples. In the past, many studies focused on the optimization of NN models and proposed pruning algorithm [6], constructive algorithm [7] and evolution neural network [8]. For the problem of samples training, the equations [9,10] were proposed to calculate the needed number of training samples for good generalization. The calculated results are very large. Thus, we proposed a sample design method with the strategy of updating training samples for the fast convergence.

2. FORWARD SOLVER

The corner backstop of the front floor of a truck shown in Figure 1a was considered in this study. The explicit dynamic finite element method was employed as the forward solver. The die used in this process is shown in Figure 1b, in which the circular drawbead was placed. The equivalent drawbead model was used in this study [13]. Figure 2 shows a half model of the drawing process of the corner backstop forming used in the finite element analysis using a commercial package (LS-DYNA 3D). The binder, the die and the punch were modelled as three separate rigid surfaces and the frictional influences between


(a) (b)

Figure 1. The corner backstop of front floor of a truck and its drawing die.

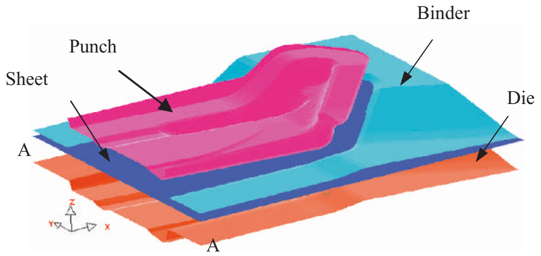


Figure 2. Finite element model of drawing process of corner backstop.

each pair of the contacting surfaces were taken into account. A coulomb friction law was assumed with friction coefficient $\mu = 0.1$ and the thickness of sheet is 1.2 mm. The material was isotropic elastic-plastic following the Barlat yield criterion and isotropic strain hardening. The elastic properties were set as, Young's modulus $E = 207$ GPa, and Poisson's ratio $\nu = 0.28$. The plastic behaviour of the material was modelled using a power law relation with a hardening coefficient K of 650 MPa and a hardening exponent n of 0.26. The binder force was 100 kN and the velocity of the punch and binder was set as 5 m/s.

3. NEURAL NETWORK

3.1 NN structure and initial training samples

In this work, the two hidden layer NN model of three inputs and two outputs was adopted, as shown in Figure 3. The modified back-propagation (BP) algorithm [11] was employed as the learning algorithm. The maximal effective stress, maximal effective strain and maximal thinning ratio of sheet thickness, expressed as x_1, x_2 and x_3, respectively, were used as the inputs for the NN model. The outputs of the NN model were radius of the male bead and radius of the female bead, expressed as y_1 and y_2. The initial neuron numbers of first and second hidden layers, n_1 and n_2, corresponding to the initial training samples were 12 and 12, respectively. In this study, the initial training samples consist of 25 sets of inputs and outputs. The search ranges of the drawbead geometric parameters, were given as follows: the radius of female bead $y_1 = 4 \sim 8$ mm, the radius of male bead $y_2 = 2 \sim 6$ mm. In the initial training samples, the combination of desired outputs of NN, y_1 and y_2, with five levels totalling to 25 set of samples were generated, and simulation results were calculated from a commercial finite element package, LS-DYNA 3D. Thus, the initial training samples have been obtained. The inputs and outputs of all these training samples were normalized from 0.2 to 0.8.

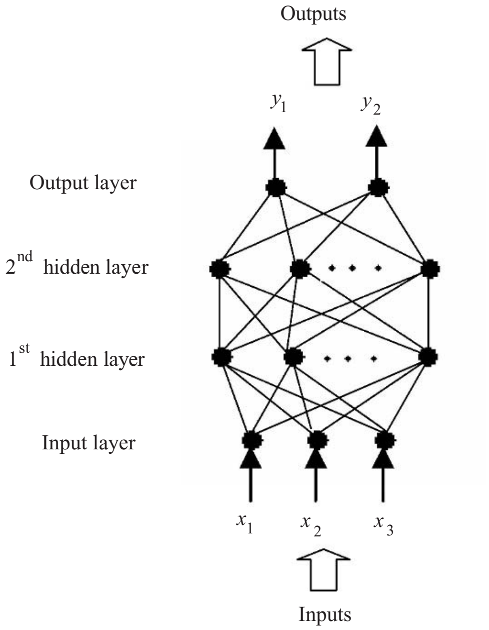


Figure 3. A two-hidden layer NN model.

3.2 Strategy of updating samples and optimization of NN structure

A sample design strategy of updating training samples with optimized NN structure was proposed for the fast convergence in this work. The strategy can be described as follows. The NN model starts to be trained using the initial training samples. According to the training results, the sample with maximum calculated error is removed from the current training samples and another two new samples located at the region near the removed one will be added instead. Thus, an updated training sample is obtained. Then, retraining of NN, with the structure optimized using IP-μ GA [12], starts using the updated training samples. The optimization is aimed to minimize the mean square error (MSE) of NN outputs. The NN model would go through a progressive retraining process till the outputs of the NN model are sufficiently close to the desired ones.

4. RESULTS AND DISCUSSION

Table 1 shows partial training datum using the strategy mentioned above. It could be observed that the calculated error of the NN model decreases as

Table 1. Partial training datum using this strategy.

	Updating number of training samples				
	1st	2nd	3rd	4th	22nd
MSE	1.7E-3	1E-3	3.4E-4	2.9E-4	8E-5

the updating number of training samples increases. Up to the 22nd updating, the calculated error of the NN model was very small. It can be seen that the accuracy of output from the NN model increases with the increasing of the number of retraining cycles. The required accuracy can be, therefore, obtained by repeating the retraining process.

5. CONCLUSIONS

In this paper, we presented a progressive NN model with structure optimized using GA dynamically to inversely identify the radius of the female and male bead. The method of updating training samples by increasing samples at local region gradually, accompanied with optimal structure of the NN model, is used to speed up the training process and improve the simulating degree of the NN model. This method is especially suitable for the case of sample datum obtained by finite analysis automatically. It has been found from the numerical results that the proposed method could be used to identify the geometric parameters of drawbead.

ACKNOWLEDGEMENTS

This work is supported by the National 973 Program under the grant number 2004CB719402 and the Foundation of M.O.E. of China for Excellent Researcher Training Program.

REFERENCES

1. H.D. Nine (1982), New drawbead concepts for sheet metal forming. *Journal of Applied Metalworking,* 2, 3, pp. 185–192.
2. N.M. Wang (1982), A mathematical model of drawbead forces in sheet metal forming. *Journal of Applied Metalworking*, 2, 3, pp. 193–199.
3. D.Y. Li, P. Hu and Y.X. Li (2000), A simple mathematical model for calculating drawbead restraining force. *Chinese Journal of Mechanical Engineering*, 36, 5, pp. 46–49.

4. M. Samuel (2002), Influence of drawbead geometry on sheet metal forming. *Journal of Materials Processing Technology*, 122, pp. 94–103.
5. J.A. Schey (1996), Speed effects in drawbead simulation. *Journal of Materials Processing Technology*, 57, pp. 146–154.
6. R. Reed (1993), Pruning algorithms—a survey. *IEEE Transactions on Neural Networks,* 4, 5, pp. 740–747.
7. T. Kwok and D. Yeung (1997), Constructive algorithms for structure learning in feedforward neural networks for regression problems. *IEEE Transactions on Neural Networks*, 8, 3, pp. 630–645.
8. H. Kitano (1990), Designing neural network using genetic algorithm with generation system. *Complex Systems*, 4, 3, pp. 461–476.
9. P.D. Wasserman (1993), *Advanced Methods in Neural Computing*. Van Nostrand Reinhold, New York.
10. A.B. Haykin (1994), *Neural Networks: Comprehensive Fundamentals.* Van Nostrand Reinhold, New York.
11. D. Xu, F.F. Yap, X. Han and G.L. Wen (2003), Identification of spring-force factors of suspension systems using progressive neural network on a validated computer model. *Inverse Problem in Engineering*, 11, 1, pp. 55–74.
12. G.R. Liu and X. Han (2003), *Computational Inverse Techniques in Nondestructive Evaluation*. CRC Press LLC, Florida.
13. D.Y. Li, P. Hu and Y.X. Li (2000), A Simple mathematical model for calculating drawbead restraining force, *Chinese Journal of Machanical Engineering*, 36, 5, pp. 46–49.

IMPROVED IMMUNE GENETIC ALGORITHM FOR SOLVING FLOW SHOP SCHEDULING PROBLEM

M. Liu, W. Pang, K.P. Wang, Y.Z. Song, and C.G. Zhou
College of Computer Science, Jilin University, Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, Changchun 130012, China

Abstract An improved Immune Genetic Algorithm was proposed to solve the Flow Shop Problem. Basing on the standard Immune Genetic Algorithm, the vaccination technique and a novel method for calculating the affinity between the antibodies was added. Finally the algorithm was test with standard benchmark problems, and the experiment result shows the validity of the algorithm.

Keywords: Immune Genetic Algorithm, vaccination, affinity, flow shop scheduling problem

1. INTRODUCTION

The Flow Shop Scheduling problem is a well-known NP hard problem. It concerns using M equipments to process N products. At any time, one equipment can process only one product. The process sequence of one product on all the equipments was the same, but the time of each product on all the equipments was not necessary the same.

Inspired by the evolution of the life, Genetic Algorithm was designed as a random heuristic global optimization algorithm; simulating the capacity of the Immune System can recognize various antigens, Immune Algorithm was designed as a multi-model searching algorithm [1]. Basically, Genetic Algorithm has not the ability or searching multi-model problem. It is appropriate to adopt the Immune Algorithm for searching the local optima of the multimodel functions. But Immune Algorithm did not perform as well as the Genetic Algorithm concerning the capacity of searching the global optima. Immune Genetic Algorithm was proposed to make up the defects of the above two algorithms.

G. R. Liu et al. (eds), Computational Methods, 1057–1062.

The crossover and mutation operator in the Genetic Algorithm were added to improve the efficiency of the algorithm.

2. SOLVING FLOW SHOP PROBLEM BASED ON THE IMMUNE GENETIC ALGORITHM

2.1 Coding strategy

Flow Shop Scheduling Problem was an ordered problem because the value of the object function was not only related to the value of the solution, but also related to the position in the coding string. Traditional binary coding was not intuitive and increased the difficulty of coding and decoding. So the integer coding method was adopted in this paper. According to the characteristic of the Flow Shop Problem, each character was denoted to a product, each product should appear once and only once in the code string. The order of the appearance of the character in the code string was the order of processing the products.

2.2 Affinity calculation

In the Flow Shop Problem, the affinity between antibody and antigen was the reciprocal of the time when the corresponding solution of the antibody was adopted to fulfill the task, that is, the reciprocal of the time when product numbered as 1 entered the first equipment until the product numbered as N left the M-th equipment [2].

The affinity of Antibody v and Antibody w represents the similarity of the two antibodies. The Information entropy in the classic immune algorithm was not adopted in this paper. A new concept was proposed to describe the affinity between antibodies; it is called permutation times in this paper. One antibody can always become another antibody after some permutation operations. For example, see Table 1.

In [0], the node 3 and node 2 was swapped in antibody 1; after one permutation, in [1] node 6 and node 4 was swapped in antibody 1; so the result after

Table 1. The transformation procedure.

Times	[0]						[1]						[2]					
Antibody 1	5	3	1	6	2	4	5	2	1	6	3	4	5	2	1	4	3	6
Antibody 2	5	2	1	4	3	6	5	2	1	4	3	6	5	2	1	4	3	6

two permutation operations was antibody 1 was transformed into antibody 2 in [2], and the permutation times was 2. The formula of the affinity between antibodies was (v and w are two antibodies):

$$a_{v,w} = \frac{1}{1 + \text{permutation times}} \tag{1}$$

2.3 The Description of the Algorithm

Step 1: Generate the candidate Solution Randomly: A sequence that is made up of the number 1 to N was used to represent a schedule solution. First L such sequences were generated randomly as the candidate solutions, and these sequences were used to fulfill the later operations.

Step 2: The crossover and mutation operator in the classic genetic algorithm was acted on the L antibodies. PMX crossover operator was taken here, and the mutate probability was 0.01.

Step 3: Calculate the affinity between each antibody and antigen; calculate the affinity between antibodies.

Step 4: Judging the offspring antibody generated by the crossover and mutation operator, if the affinity of this antibody to the antigen is lower than a threshold (the minimum value of the affinity to the antigen in the origin parent antibodies namely), the retrogression has happened during the crossover and mutation, and the antibody should be deleted from the offspring antibodies.

Step 5: The antibody that has the best affinity to the antigen in the L antibodies was saved as the bacterin in this generation and was acted on the other antibodies; this is called vaccination (The detailed operation was described in 2.4).

Step 6: Delete the antibodies according to the descending order of the concentration of the antibody, the delete operation was stopped until L antibodies was left in the new antibody population, the left L antibodies composed a new generation population of antibodies. The formula of calculating the concentration of the antibodies was as follows:

$$C_v = \frac{\sum_w a_{v,w} = 1}{L} \tag{2}$$

Step 7: If the number of the iterations has exceeded a predefined integer number, the algorithm stopped, if not, GOTO Step 2.

Step 8: In the memory cell, select the antibody that has the max affinity to the antigen as the global optima solution

Table 2. The vaccination procedure.

bacterin	7	3	5	2	6	9	8	1	4										
order			2	1	3	5	4												
target antibody	2	5	4	1	7	8	9	3	6		2	5	4	1	7	9	8	3	6
order			2	1	3	4	5						2	1	3	5	4		

N = 10, select two random numbers as the start and end point of the acting area: (3,8)

2.4 Vaccination

When Immune Genetic Algorithm was used to solve Flow Shop Problem, if the antibody was created only with the crossover and mutation operation, the new antibody population may not better than the old population with regard to the affinity to the antigen. To solve this problem, it is necessary to select an excellent antibody to guide the generation of the new antibody. This is the destination of the vaccination. Meanwhile, it should be avoided that all the antibodies trend to the direction of the bacterin. It should keep the diversity of the population and prevent the algorithm trapping into local optima.

In this paper, a vaccination method aiming at Flow Shop Problem was, proposed, the experiment shows this method can accelerate the speed of finding the optima, and guarantee the correction of the global optima, for example:

3. EXPERIMENTS AND ANALYSIS

The experiments are made on PC (Pentium IV 2.8 GHz CPU, 512M RAM, WinXP OS, VC++6.0). The results of Flow-shop 12*5(7964)[3] are shown in Figure 1 (L = 50, iteration times = 100), and the shortest time = 8516 (2, 3, 5, 1, 4, 2); the results of Flow-shop 8*9(8815)[3] are shown in Figure 2, and the shortest time without vaccination operator is time = 9455 (9, 4, 6, 1, 2, 5, 3, 7, 8), and with it is time = 9011 (8, 3, 2, 1, 5, 6, 7, 9, 4).

Table 3 shows the performance of four different algorithms for solving four Benchmark problems [3]. IGA denotes the standard Immune Genetic Algorithm, Inc denotes the vaccination operator, NAC denotes the new affinity calculation.

4. CONCLUSION

Flow Shop Problem is often encountered in real world application. For better improving the efficiency, reducing the idle time of the equipments, a

Table 3. Performance comparison.

		IGA	IGA+Inc	IGA+NAC	IGA+Inc+NAC
11*5 (7038)					
	Time	7964	7549	7582	7356
	Deviation	13.16	7.22	7.73	4.52
8*8 (8366)					
	Time	9246	8917	9148	8816
	Deviation	10.52	6.59	9.35	5.38
20*15 (1930)					
	Time	2187	2096	2145	2054
	Deviation	13.32	8.6	11.14	6.42

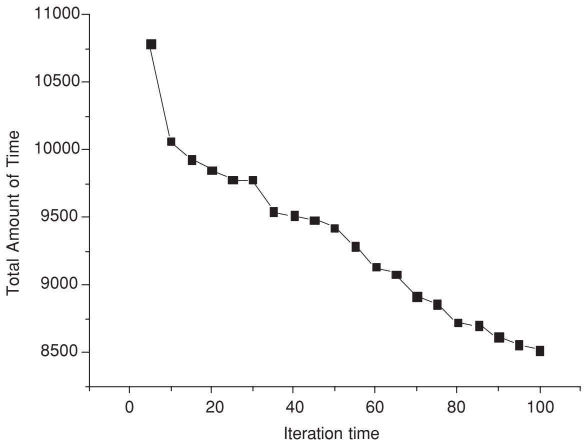


Figure 1. 12*5.

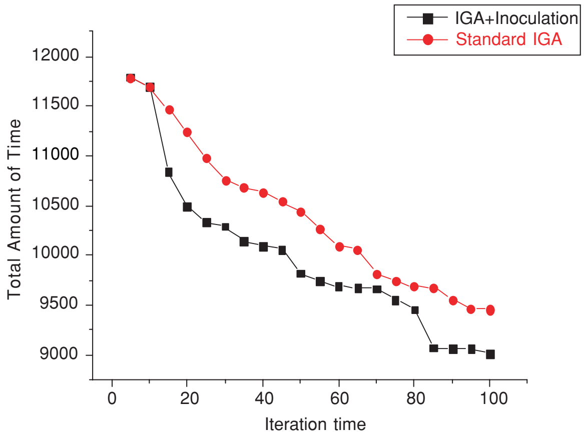


Figure 2. 8*9.

rational process sequence is fairly important, and it plays an important role in improving the economic benefits. Basing on the standard Immune Genetic Algorithm for solving Flow Shop problem, the vaccination techniques and new method of calculating the affinity between antibodies was adopted to improve the Immune Genetic Algorithm. Experiment results show the validity of the algorithm.

REFERENCES

1. L.N. De Castro (2002), Immune, Swarm and Evolutionary Algorithms Part I: Basic Models. In: *Proceedings of the ICONIP Conference (International Conference on Neural Information Processing), Workshop on Artificial Immune Systems*, 3, pp. 1464–1468.
2. D. Dasgupta (1999), Immune-Based Intrusion Detection System: A General Framework. In: *Proceedings of the 2nd national information systems security conference (NISSC)*, Virginia, USA.
3. L.P. Androulakis and V. Venkatasubramanlan (1991), A genetic algorithm framework for process design an optimization. *Computers and Chemical Engineering* 15(4), pp. 217–228.

A DISCRETE PARTICLE SWARM OPTIMIZATION ALGORITHM FOR TRAVELLING SALESMAN PROBLEM

X.H. Shi[1], Y. Zhou[1], L.M. Wang[1,2], Q.X. Wang[1] and Y.C. Liang[1]
[1]*College of Computer Science, Jilin University, Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, Changchun 130012, China*
[2]*Department of Computer Science and Technology, Changchun Taxation College, Changchun 130021, China*

Abstract A discrete particle swarm optimization (PSO)-based algorithm for travelling salesman problem (TSP) is presented by re-designing the 'subtraction' operator. To accelerate the convergence speed, a crossover eliminating technique is also added. Numerical results of some benchmark instances show that the size of solved problems could be increased by using the proposed algorithm compared with those of the existing PSO-based algorithms.

Keywords: particle swarm optimization, travelling salesman problem, swap operator.

1. INTRODUCTION

Particle swarm optimization (PSO) algorithm, originally developed by Kennedy and Elberhart [1], is a method for optimization on metaphor of social behaviour of flocks of birds and/or schools of fish. The PSO system is initialized firstly in a set of randomly generated potential solutions, and then performs the search for the optimum one iteratively. It finds the optimum solution by swarms following the best particle. Presently PSO has attracted broad attention in the fields of evolutionary computing, optimization and many others [2, 3]. Though PSO is developed for continuous optimization problems initially, there have been some reported works focused on discrete problems recently [4].

The travelling salesman problem (TSP) is a well-known and extensively studied benchmark for many new developments in evolutionary computation [5–7]. Furthermore, Hendtlass [8] and Wang et al. [9] proposed different PSO

G. R. Liu et al. (eds.), Computational Methods, 1063–1068.

methods for solving TSP problems. Though the PSO method could be applied to TSP, the size of solved problems reported by Hendtlass [8] and Wang et al. [9] are both smaller than 15. It seems that the size of solved TSP problems using algorithms of PSO is limited.

In this paper, the 'subtraction' operator between two particle positions is modified and a discrete PSO method is constructed for TSP. The crossover eliminating technique is also executed in the proposed method. Numerical results show that the proposed method could improve the size of resolvable TSP problems.

2. PARTICLE SWARM OPTIMIZATION ALGORITHM

First, the standard PSO algorithm is introduced briefly. Suppose that the searching space is D-dimensional and m particles form the colony. The ith particle represents a D-dimensional vector$X_i (i = 1, 2, \ldots, m)$. It means that the ith particle locates at $X_i = (x_{i1}, x_{i2}, \ldots, x_{iD})(i = 1, 2, \ldots, m)$ in the searching space. The position of each particle is a potential solution. We could calculate the particle's fitness by putting its position into a designated objective function. When the fitness is higher, the corresponding X_i is "better". The ith particle's 'flying' velocity is also a D-dimensional vector, denoted as $V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})$ $(i = 1, 2, \ldots, m)$. Denote the best position of the ith particle as $P_i = (p_{i1}, p_{i2}, \ldots, p_{iD})$, and the best position of the colony $P_g = (p_{g1}, p_{g2}, \ldots, p_{gD})$, respectively. The PSO algorithm could be performed by the following equations:

$$X_i(k+1) = X_i(k) + V_i(k+1)\Delta t \tag{1}$$

$$V_i(k+1) = wV_i(k) + c_1 r_1 (P_i - X_i(k))/\Delta t + c_2 r_2 (P_g - X_i(k))/\Delta t \tag{2}$$

where $i = 1, 2, \ldots, m, w$ is the inertia coefficient which is a constant in interval [0, 1]; c_1 and c_2 are learning rates which are nonnegative constants; $r_1$and r_2 are generated randomly in interval [0, 1]; Δt is the time interval, and commonly is set as unit; $v_{id} \in [-v_{\max}, v_{\max}]$, and v_{max} is a designated value. The termination criterion for the iterations is determined according to whether the maximum generation or a designated value of the fitness of P_g is reached.

3. ALGORITHM OF THE DISCRETE PSO ALGORITHM FOR TSP

By adding a memory capacity to each particle in the PSO algorithm, Hendlass [8] applied PSO algorithm to solve small-size TSP problems, and improved its performance. While Wang et al. [9] redefined the PSO operators by introducing the concepts of 'Swap operator' and 'Swap sequence', therefore the TSP problems could be solved by PSO in another way. Coincidentally, both of them selected Burma14 (a benchmark problem in TSPLIB with 14 cities) as simulated samples. That is to say, the size of cities is rather limited in their algorithms. To expand the scale of the solved problem, inspired by the concept of 'Swap operator', we consider adding an uncertainty searching strategy and propose a discrete PSO algorithm for TSP. In the traditional PSO algorithm, the distance of the positions between different particles' or between different moments of the same particle is employed to update the velocity of a particle [see Equation (2)]. Wang et al. [9] expressed the distance by the 'swap sequence'. While in this paper the 'distance' between two particle positions is estimated by their fitness, and the 'distance function' is defined as

$$dis\_fit(X_i, X_j) = |f(X_i) - f(X_j)| \tag{3}$$

where X_i and X_j are the positions of particles, $f(X_i)$ and $f(X_j)$ are their corresponding fitness values, respectively. Denote the coding of X_i as $x_{i1}, x_{i2}, \ldots, x_{im}$ and that of X_j as $x_{j1}, x_{j2}, \ldots, x_{jm}$, respectively, where m is the number of cities, $x_{ik} \in [1,\ m]$. Then X_i represents the travelling circle of $x_{i1} \to x_{i2} \to \cdots \to x_{im} \to x_{i1}$. Similar to the algorithm of Wang et al. [9], we define a 'swap operator' denoted by $SWAP(X_i, X_j, k)$ as follows:

1. Search the kth bit of X_i and the bit of X_j whose value is equal to the kth bit of X_i;
2. Search the first pair bits with different values following them one by one in both X_i and X_j;
3. Swap the bits of X_i whose values are equal to the different values.

The swapping process is described as follows: when updating a position of a particle, an index number k is selected randomly, then the swap operator $SWAP(X_i,\ X_j,\ k)$ is executed, where X_i is the code of the particle, X_j is the corresponding P_g or P_i code. The number of swapping times depends on the value of $dis\_fit\,(X_i, X_j)$ and should be increased with the increasing of the value of $dis\_fit(X_i, X_j)$. Therefore, the number of swap times is selected as

$$number\_swap = \begin{cases} n\_dis & if\ (n\_dis \le n\_\max) \\ n\_\max & else \end{cases} \tag{4}$$

$$n\_dis = int\left(k \cdot dis\_fit(X_i, X_j)/f(X_j) + 1.5\right) \tag{5}$$

where $int(\cdot)$ is the truncation function in C Language, k is a scale parameter, and n_max is a given maximum number for swapping times, respectively.

To speed up the convergence, a process is added to delete the crossover of travelling lines. Denote $line[i]$ representing the travelling line between the ith node and the $(i + 1)$th node of a particle position X. Then the pseudo-code of the delete-crossover process could be illustrated as follows:

```
for each $line[i]$ and $line[j]$
    if $line[i]$ and $line[j]$ is crossed
        for $(k = 0; k < (j - i)/2; k + +)$
            swap$(x_{j-k}, x_{i+k+1})$.
```

In the PSO program, all particles should learn from P_g, therefore the delete-crossover process is performed on P_g in each of iterations. The process is performed only to one particle, therefore, will obviously lead to time-saving. On the other hand, the improvement of P_g affects all particles, so the delete-crossover is fairly effective.

In summary, the proposed discrete PSO algorithm for TSP can be described as follows:

1. Initialize the particle swarm: including the initialisations of each particle's position and the evaluation of the fitness of each particle. At this stage, P_g is also searched and P_i of each particle is set as its initial position.
2. Apply PSO processes to each particle:
    i. Execute the swapping process for each particle;
    ii. Execute the delete-crossover process for P_g;
3. Judge the stop criterion, namely whether the iteration reaches the given number or the best fitness reaches the designated value. If the criterion is satisfied then stop the program, else go to step 2.

4. NUMERICAL RESULTS

To verify the validity of the proposed discrete PSO algorithm, some instances from TSPLIB library [10] are selected for simulations. The experiments are performed on a PC with 2 GHz processor and 128 M memory.

Table 1 presents the numerical results. The first column stands for the names of the test instances, the second for the best known optimal tour length for each

Table 1. Results of the proposed algorithm for TSP problems.

Problem	Opt	Best result	Average	Err(%)
EIL51	426	437	444.6	4.366
BERLIN52	7542	7700	7960.2	5.54
ST70	675	712	733.2	8.62
EIL76	538	580	587.4	9.18
PR76	108159	113505	115144.2	6.46

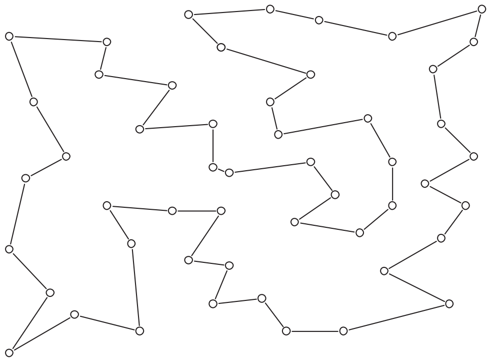


Figure 1. Suggested tour of the benchmark problem EIL51.

problem, the third for the best result obtained, the fourth for the average result and the fifth for the relative error (Err), respectively.

Table 1 shows that the proposed discrete PSO method can be used to solve TSP effectively. From Table 1 it can be seen that among five test problems, the maximum relative error is 9.18% (of the test problem EIL76) and the average relative error of all is easily calculated to be 6.83%. Figure 1 shows the travelling circles for the calculate results of problems EIL51. From Figure 1 it can be seen that the results are fairly good when the problem sizes are ranged from 50 to 80. Whereas the results from the existing PSO-based algorithms of Hendlass [8] and Wang et al. [9] show that the sizes of solved TSP problems were only 14. It is obvious that the proposed algorithm has advantage on the solved problem size compared with the existing algorithms.

5. CONCLUSIONS

Focused on the TSP problem, a novel discrete PSO algorithm is presented by adding an uncertain strategy into the approach. Some benchmark problems are tested to examine the effectiveness of the proposed algorithm. Numerical results show that the proposed algorithm is effective. It has also been shown that the proposed algorithm can solve larger size problems than those solved using the existing algorithms.

ACKNOWLEDGEMENTS

This work was supported by the science-technology development project of Jilin Province of China under Grant No. 20030520 and the doctoral funds of the National Education Ministry of China under Grant No. 20030183060.

REFERENCES

1. J. Kennedy and R.C. Eberhart (1995), Particle swarm optimization. In: Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia, IEEE Service Center, Piscataway, NJ, 4, pp. 1942–1948.
2. P.J. Angeline (1998), Evolutionary optimization versus particle swarm optimization: philosophy and performance differences. *Evolutionary Programming*, 7, pp. 601–610.
3. M. Clerc and J. Kennedy (2002), The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, 6, pp. 58–73.
4. J. Kennedy and R.C. Eberhart (1997), Discrete binary version of the particle swarm algorithm. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Orlando, Florida, USA, 5, pp. 4104–4108.
5. D.E. Goldberg (1989), Messy genetic algorithms: motivation, analysis, and first results. *Complex Systems*, 3, pp. 493–530.
6. J. Bentner, G. Bauer, G.M. Obermair, I. Morgenstern and J. Schneider (2001), Optimization of the time-dependent traveling salesman problem with Monte Carlo methods. *Physical Review E*, 64, 036701.
7. L. Huang, C.G. Zhou and K.P. Wang (2003), Hybrid ant colony algorithm for traveling salesman problem. *Progress in Natural Science*, 4, 13, pp. 295–299.
8. T. Hendlass (2003), Preserving diversity in particle swarm optimization. *Lecture Notes in Computer Science*, 2718, pp. 4104–4108.
9. K.P. Wang, L. Huang, C.G. Zhou and W. Pang (2003), Particle swarm optimization for traveling salesman problem. *International Conference on Machine Learning and Cybernetics*, 3, pp. 1583–1585.
10. G. Reinelt (1991), TSPLIB—A traveling salesman problem library. *ORSA Journal on Computing*, 3, 4, pp. 376–384.

A SEARCH METHOD FOR FINDING A SIMPLE NASH EQUILIBRIUM

S.Y. Sun[1,2], D.Y. Liu[1,2] and Z. Li[3]
[1] *College of Computer Science and Technology, Jilin University, 10 Qianwei Road, Changchun 130012, China*
[2] *Key Lab for Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, 10 Qianwei Road, Changchun 130012, China*
[3] *Business School, Jilin University, 10 Qianwei Road, Changchun 130012, China*

Abstract In this paper, we propose a search method for finding a simple Nash Equilibrium in a 2-player game and execute it on 24 classes of games. The result shows that our algorithm performs better than the classical Lemke–Howson algorithm.

Keywords: game theory, 2-player game, nash equilibrium computing.

1. INTRODUCTION

Non-corporative game theory has been tightly combined with Computational Methods, in which the Nash Equilibrium (NE) is undoubtedly the most important solution concept.

While players face several Nash Equilibriums, their most pragmatic strategy may not be the most optimal one [1], but one they prefer for other reasons. Players may choose a sub-optimal strategy instead of a more optimal but more complex one which might be difficult to learn or to implement. In one word, players prefer strategies as simple as possible.

The problem is whether there exists a simple NE in a 2-player game, and if so, how to find it. To solve this problem, this paper proposes a search method. Comparing with finding a NE, it is much easier to compute whether a NE exists in a special support for each player. Based on this, our algorithm repeats checking the existence of NE in a limited and ordered support space until a NE is found. There are three important steps in our algorithm: first limiting the search space; second ordering the search space; lastly, checking in turn

G. R. Liu et al. (eds.), Computational Methods, 1069–1073.

(according to the order) whether there exists a NE with this support for each player.

We are not the first ones who search support space to find a NE. Dickhaut and Kaplan [2] proposed an algorithm to find all Nash Equlibria. Porter et al. [3] designed an algorithm to find a sample NE. However, the innovation in this paper is on pruning the search space. Our pruning rules are (1) no conditionally dominant strategies in any NE, (2) there exists such a NE whose support is less than $k + 1$(k is the rank of corresponding payoff matrix). These rules ensure that our search space is smaller than that of Porter et al. [3], but the disadvantage is that ranks of payoff matrixes must be computed.

The rest of this paper is structured as follows: first we formulate the problem and give the relevant definitions, and then we describe the algorithm. After that, we compare the execution of our algorithm and the Lemke–Howson algorithm [4]. In the final section, we conclude our work.

2. NOTATIONS AND DEFINITIONS

The necessary notations and definitions are shown in this section.

The set of players is noted as $N = \{1, \quad 2\}$.

The sets of pure strategies of player 1 and 2 are noted as $S_1 = \{s_{11}, s_{12}, \ldots, s_{1m}\}$ and $S_2 = \{s_{21}, s_{22}, \ldots, s_{2n}\}$.

The payoff matrixes of player 1 and 2 are noted as $(U_1)_{m\times n}$ and $(U_2)_{n\times m}$, respectively.

Definition 1. The ith row of matrix $U_{m\times n}$ is *conditionally dominant*, if $\exists\ 1 \le j \le m, \forall k \in \{1, \ldots, n\}$, s.t. $u_{ik} \le u_{jk}$ and $\exists\ 1 \le t \le m$, s.t. $u_{it} < u_{jt}$.

$S_1' = \{s_{1i} \in S_1 |$ the ith row of matrix U_1 is conditionally dominant$\}$

$S_2' = \{s_{2i} \in S_2 |$ the ith row of matrix U_2 is conditionally dominant$\}$

The strategies in S_1' and S_2' are called *conditionally dominant strategies*.

Definition 2. A *mixed strategy* for a player is a probability distribution over the set of his pure strategies and will be represented by a vector $x = (x_1, x_2, \ldots, x_m)$, where $x_i \ge 0$ and $\sum_{i=1}^{m} x_i = 1$. Here x_i is the probability that the player will choose his ith pure strategy. The *support* of x (Supp(x)) is the set of pure strategies that it uses.

Definition 3. A profile of mixed strategies (p, q) for a 2-player game is balanced, if $|\mathrm{Supp}(p)| = |\mathrm{Supp}(q)|$.

Theorem 1. For any Nash Equilibrium $(x^*, \quad y^*)$, $\mathrm{Supp}(x^*) \cap S_1' = \Phi$, $\mathrm{Supp}(y^*) \cap S_1' = \Phi$.

Proof. Using proof by contradiction.

Theorem 2. $(U_1)_{m\times n}$, $(U_2)_{n\times m}$ are payoff matrixes, respectively, for player 1 and 2 in a 2-player game. rank$(U_1) \leq k$, rank$(U_2) \leq d$. $\exists$ a NE $(x,\ y)$, s.t. $|\text{Support}(x)| \leq k+1$, $|\text{Support}(y)| \leq d+1$.

Proof. See Lipton et al. [5].

3. ALGORITHM

The following is a feasible program for finding a NE in a special support for each player, where U_1 and U_2 correspond payoff matrixes of 2 players, and S_1^t, S_2^t correspond the special support for each player.

Vectors $p_1 \in R^m$, $p_2 \in R^n$ correspond Nash Equilibrium (if they exist), and $c_i \geq 0$ are real numbers.

$p_{-1} \in R^m$, $p_{-2} \in R^n$ are vectors satisfying constraints 1, 2 and 4.

Feasibility program 1.
Input: $(U_1)_{m\times n}$, $(U_2)_{n\times m}$, S_1^t, S_2^t.
Output: $p_i,\ c_i, i = 1,\ 2$ s.t.
Constraints: $\forall i = 1,\ 2$

(1) $\sum_{s_{ij}\in S_i^t} p_{\pm i}(s_{ij}) = 1$, and $\begin{array}{ll} 0 < p_{\pm i}(s_{ij}) \leq 1, & \text{if } s_{ij} \in S_i^t \\ p_{\pm i}(s_{ij}) = 0, & \text{if } s_{ij} \notin S_i^t \end{array}$

(2) $\exists s_{ij} \in S_i^t$, st. $p_i(s_{ij}) \neq p_{-i}(s_{ij})$

(3) $p_1 \cdot U_i \cdot p_2^T = c_i$

(4) $p_{(-1)^i\cdot 1} \cdot U_i \cdot p^T_{(-1)^{i+1}\cdot 2} \leq c_i$

See Chvatal [6] for solving feasible program 1.

Definition 4. Suppose that x and y are positive integers. $G = \{1, 2, \ldots, x\} \times \{1, 2, \ldots, y\}$, $(x_1, y_1), (x_2, y_2) \in G$. The simple and balanced order $(G, \leq)$ is defined as follows: $(x_1, y_1) \leq (x_2, y_2)$ if and only if one of following conditions are satisfied:

(1) $|x_1 - y_1| < |x_2 - y_2|$

(2) $|x_1 - y_1| = |x_2 - y_2|$ and $\min\{x_1, y_1\} \leq \min\{x_2, y_2\}$

The following algorithm is for finding a simple NE in a 2-player game.

Search space of Algorithm 1: according to Theorem 1, there are no conditionally dominant strategies in any Nash Equilibrium; according to Theorem 2, there exists a Nash Equilibrium whose support of mixed strategies is less than $k + 1$ (k is the rank of corresponding payoff matrix). The search space is based on these two theorems.

Algorithm 1.
// computing a NE in a 2-player game, 1 and 2 are players

```
Input: $S_1$ and $S_2$ (sets of pure strategies), $(U_1)_{m\times n}$ and $(U_2)_{n\times m}$ (payoff matrixes)
$S_1' \leftarrow \{s_{1i} \in S_1 |$ the $i$th row of matrix $U_1$ is conditionally dominant$\}$
$S_2' \leftarrow \{s_{2i} \in S_2 |$ the $i$th row of matrix $U_2$ is conditionally dominant$\}$
$S_A \leftarrow S_1 - S_1'$, $S_B \leftarrow S_2 - S_2'$
$k \leftarrow$ rank$(U_1)$, $d \leftarrow$ rank$(U_2)$
$x \leftarrow \min\{k+1, |S_A|\}$
$y \leftarrow \min\{d+1, |S_B|\}$
FOR all $(x', y')$ where $x' \in \{1, 2, \ldots, x\}, y' \in \{1, 2, \ldots, y\}$, and $((\{1, 2, \ldots, x\} \times \{1, 2, \ldots, y\}), \leq)$ is a simple and balanced order
FOR all $S_1^t \subseteq S_A$, s.t. $|S_1^t| = x'$ DO
FOR all $S_2^t \subseteq S_B$, s.t. $|S_2^t| = y'$ DO
IF $\exists(p, c)$, which satisfies feasible program 1 for $(S_1^t, S_2^t)$ THEN
RETURN $p$ // $p$ is the found equilibrium
```

4. ASSESSMENT (EVALUATION)

Several comparisons were committed between the performance of Algorithm 1 and that of the Lemke–Howson algorithm implemented in Gambit [7] to assess our algorithm.

We follow the idea of Ryan et al. (2004) and use GAMUT (GAMUT can generate games from a wide variety of classes of games found in the literature) to generate 24 different classes of games. Both algorithms are executed on 100 2-player 300-action games drawn from each class (i.e., 2400 games as sample games). The comparisons are made on median runtimes; percentage of samples solved, and average runtime. It has been shown that both the median runtime and the conditional average runtime of Algorithm 1 are much shorter than that of Lemke–Howson on all the 24 classes. Furthermore, our algorithm finds Nash Equilibrium in more sample games on several distributions, and only the solving percentage on distribution six is a little lower than Lemke–Howson.

5. CONCLUSIONS

In this paper, we present a search method for finding a simple Nash Equilibrium in a 2-player game. We analyse the payoff matrix of each player to prune the support space, and order the support space to ensure the simple and balanced NE can be found as fast as possible. After testing our algorithm on 24 classes of games and comparing it with the classical Lemke–Howson algorithm, the results show our algorithm performs better.

6. ACKNOWLEDGEMENTS

This research benefits from the support of NSFC Major Program/China (60496321), NNSF/China (60373098, 60173006), 863 Hi-Tech Research and Development Program/China (2003AA118020), Major Program of Jilin Province Development Plan for Science and Technology/China (20020303) and General Program of Jilin Province Development Plan for Science and Technology/China (20030523).

REFERENCES

1. A. Rubinstein (1998), *Modeling Bounded Rationality*. MIT Press, Cambridge.
2. J. Dickhaut and T. Kaplan (1991), A program for finding Nash equilibria, *The Mathematica Journal*, 1, 4, pp. 87–93.
3. R. Porter, E. Nudelman and Y. Shoham (2004), Simple search methods for finding a Nash Equilibrium. *American Association for Artificial Intelligence*.
4. C. Lemke and J. Howson (1964), Equilibrium points of bimatrix games. *Journal of the Society for Industrial and Applied Mathematics*, 12, pp. 413–423.
5. R. Lipton, E. Markakis and A. Mehta (2003), Playing large games using simple strategies. *ACM Conference on Electronic Commerce*, pp. 36–41.
6. V. Chvatal (1983), *Linear Programming*. Freeman, New York.
7. R. McKelvey, A. McLennan and T. Turocy (2004), Gambit: software tools for game theory. Available at http:// econweb.tamu.edu/gambit/.

A HERO EVOLUTIONARY ALGORITHM HYBRIDIZING FROM PSO AND GA

D.W. Guo, C.G. Zhou and M. Liu
College of Computer Science and Technology, Jilin University,
Changchun 130012, P.R. China

Abstract Both GA and PSO are typical evolution algorithm with their own advantages. In this paper, a new evolution algorithm is introduced based on GA and PSO. The total population are divided into several tribes, which one Hero individual and several common particles are included in each tribe. The movement velocity of each hero is calculated by global peak point and local best point among this tribe just like PSO. Other common particles will search neighbourhood of hero using recombination method of GA. Hero algorithm will converge fast and escape from local peak inheriting advantages of GA and PSO. These conclusions are proven from experiment of some familiar Benchmark functions.

Keywords: Hero, genetic algorithm, particle swarm optimization

1. INTRODUCTION

The Particle Swarm Optimization (PSO) algorithm was originally introduced by Kennedy and Eberhart [1]. As a new Evolutionary Algorithm, the PSO was inspired by insect swarms. It has been proved to be an effective method to solve some optimization problem.

PSO is a typical global optimization algorithm compared with the traditional genetic algorithm (GA). The population of these two algorithms is initialled randomly. And the searching is carried on with the fitness of all individuals. Relatively speaking, PSO has the characteristics as follows: high speed in convergence, few adjustable variables, is insensitive to the size of population. At the same time, it also has some shortcomings: in the early period of running, low precision, is liable to diverge, is apt to miss the global best solution; In the later stage, similarity of all the particles is very obvious, and the speed of

G. R. Liu et al. (eds.), Computational Methods, 1075–1080.

convergence is slow down, is very easy to sink into local peak point. Generally, PSO performs well in the early iterations, but has problems reaching a near optimal solution.

Some researchers have already made a lot of improvements of PSO algorithm. The inertia weight was brought out by Shi and Eberhart [2], and the constriction coefficient was advanced by Clearc [3]; Spatial neighbourhoods and other improvements were proposed by some scholars in order to increase the diversity of population. Combining PSO and GA, we can synthesize the advantages of these two algorithms, so the effect is brought to. An explicit selection operator in PSO is used by Angeline [4]. It is shown that the capability of local search is improved, but simultaneously the ability of global search is reduced. In addition, recombination and reproduction is also introduced by Lovbjerg [5]. The result shows that, for the function with multi-model local peaks, the speed of convergence is accelerated, and better solution can be found.

A new Hero EA models is constructed in this paper by hybridized merits of PSO and GA algorithms. Some particles are called Hero in the population. And the velocity of movement is similar to the PSO, both of them are decided by the global peak point and local peak point. There are several common particles around each Hero are evolved with recombination of GA. These particles are used to search the neighbourhood of Hero, and the information of local peak point is applied. Thus the aptitude of high speed convergence of PSO is played; simultaneously the capability of global search is assured to a certain degree so that local peak is jumped out.

The next section presents the models of the new Hero EA. Section 3 describes the experimental settings and results of some benchmark functions. The experimental results and conclusion are discussed in Section 4 finally.

2. MODEL

Here, a new model hybridizing from PSO and GA is proposed. In this model, total population is divided into several tribes, while one Hero individual and several common particles are included in each tribe. The speed and direction of each hero's movement is decided by global peak point and local best point among this tribe (includes the points that hero has searched and the points that common particles have searched) just like PSO. On one hand, every common particle evolves in accordance with recombination method of GA; on the other hand, remove with the same speed of the hero in this tribe.

The equation of hero individual movement shows below:

$$\begin{cases} \vec{v}_i = \chi^*(w\vec{v}_i + c_1 * rand() * (\vec{p}_i - \vec{x}_i) + c_2 * rand() * (\vec{p}_g - \vec{x}_i)) \\ \vec{x}_i = \vec{x}_i + \vec{v}_i \end{cases} \tag{1}$$

Inside of tribe, offspring particles are recombined with probabilityp_b. The parent particles are replaced by their offspring particles. Note that the fitness is not used when selecting particles for recombination. After recombination, in order to search the neighbourhood of hero, offspring should move their position according to the velocity of hero.

The arithmetic crossover of positions in the search space is one of the most commonly crossover methods with standard real valued GAs, placing the offspring within the hypercube spanned by the parent particles. The main motivation behind the crossover is that offspring particles benefit from both parents and searches in neighbourhood around hero point. In theory, having two particles on different suboptimal peaks breed could result in an escape from a local optimum or find a new peak, and thus aid in achieving a better one.

At the initial phases, the positions of every hero and common particles are distributed randomly, and the velocity are zero. Along with the evolvement of algorithm, common particles will close to Hero.

The structure of the standard model is illustrated as below.

```
Begin
    initialize
    while (not terminate-condition) do
    begin
        evaluate every particles and heroes.
        Update global best point, local best point of each tribe and the best position of hero.
        calculate new velocity vectors of heroes and move.
        For each tribe do
        Begin
            Recombine particles to make offspring
            Move offspring according to the velocity vector of hero
        end
    end
end
```

3. EXPERIMENTAL SETTINGS

For the sake of validating performance of Hero algorithm, four familiar benchmark problems are tested as the Table 1. The detail formula can be found in other papers. The first and second function are unimodel, and other two are multi-model with many local minima. All these functions are designed such that their global minimum are the origin of the search space. These functions are commonly used in performance evaluation on GA and PSO.

Table 1. Function name, search ranges and dimension of problems.

Index	Function name	Search range	Dimension
f_1	DeJong F1	$-100 < x_i < 100$	100
f_2	DeJong F2	$-30 < x_i < 30$	50
f_3	Griewank	$-600 < x_i < 600$	50
f_4	Rastrigin	$-5.12 < x_i < 5.12$	20, 30, 50

Table 2. The experiment parameters.

Algorithm	Parameter	Value
GA	$p_{\text{crossover}}$	0.7
	p_{mutation}	0.15
PSO	Inertia weight	0.8
	constriction	1.0
ALL	Population size	60
	Iteration turns	1000

Table 3. The best solution.

Function	Dim	GA	PSO	Hero
F1	100	13.27	42.7	23.88
F2	50	493.2	820.7	665.9
F3	50	1.12	1.31	0.95
F4	20	10.25	21.22	3.25
	30	25.39	37.01	26.61
	50	36.98	109.01	122.21

The experiment parameters are shown in Table 2. Notice that in Hero algorithm, there are 20 tribes (or heroes) and three common particles in each tribe just like 60 individuals. Each algorithm will run 10 times independently and record average results.

The best solution found in this experiment of these function is shown in Table 3. Figures 1–4 shows the fitness curves of every problem and every algorithm. From these data and graph, it is shown that the Hero algorithm can converge fast than GA and PSO, specially on high-dimension problem.

4. CONCLUSION

In this paper, a new Hero evolution algorithm is introduced based on GA and PSO. The total population is divided into several tribes, which one Hero

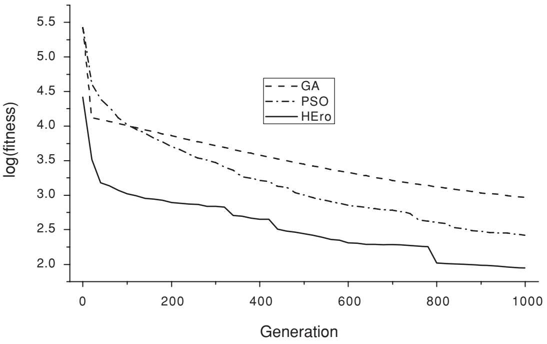


Figure 1. The function F1.

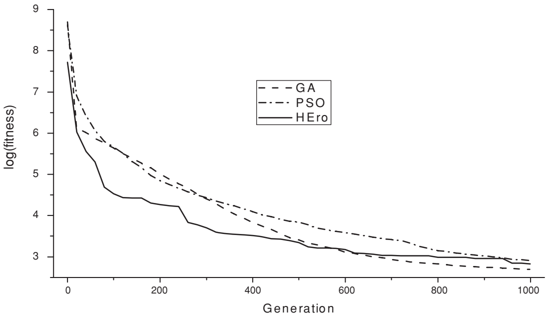


Figure 2. The function F2.

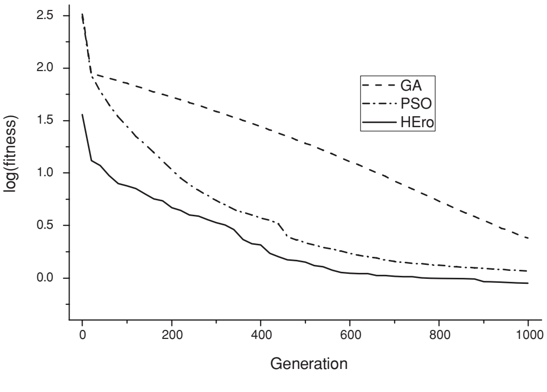


Figure 3. The Griewank function F3.

individual and several common particles are included in each tribe. The movement velocity of each hero is calculated by global peak point and local best point among this tribe just like PSO. Other common particles will search neighbourhood of hero using recombination method of GA. It is shown that the Hero

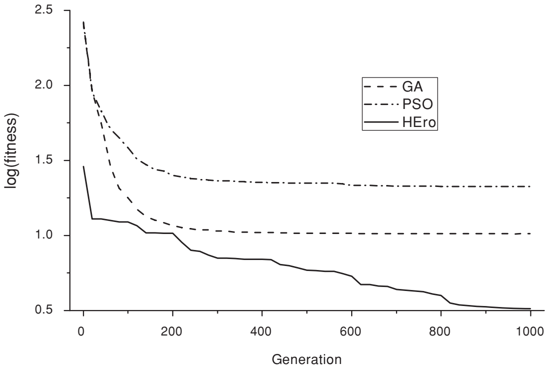


Figure 4. The Rastrigin function F4 ($N = 20$).

algorithm is effective on some high-dimension problem from experiment of some familiar Benchmark functions. In future, this principle will be introduced into other evolutionary algorithm.

REFERENCES

1. J. Kennedy and R.C. Eberhart (1995), Particle swarm optimization. In: Proceedings of the 1995 IEEE International Conference on Neural Networks, 4, pp. 1942–1948. IEEE Press.
2. Y. Shi and R.C. Eberhart (1998), A modified Particle Swarm Optimiser, IEEE International Conference on Evolutionary Computation, Anchorage, Alaska, May 4–9.
3. M. Clerc (1999), The swarm and the queen: towards a deterministic and adaptive particle swarm optimisation. In: Proceedings of the 1999 Congress of Evolutionary Computation, 3, pp. 1951–1957. IEEE Press.
4. P.J. Angeline (1998), Using selection to improve particle swarm optimization. *In 1998 IEEE International Conference on Evolutionary Computation*, Anchorage, Alaska, USA, pp. 84–89.
5. M. Lovbjerg, T.K. Rasmussen, and T. Krink (2001), Hybrid particle swarm optimiser with breeding and subpopulations. In: *Proceedings of the Third Genetic and Evolutionary Computation Conference 2001*, San Francisco, USA.

A DATA COLLECTION MODEL FOR INTRUSION DETECTION SYSTEM BASED ON SIMPLE RANDOM SAMPLING

Liang Hu[1,2], Kuo Zhao[1,2], and Bo Li[1]
[1]*Department of Computer Science and Technology Jilin University, Changchun, China, 130012*
[2]*Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, Jilin University, Changchun, China, 130012*

Abstract Data collection mechanism is the crucial factor that affects the performance of intrusion detection system (IDS). Simple random sampling technique of statistics is introduced to the procedure of data collection of IDS, and a new data collection model for IDS is provided. Formulas used to calculate the sample size of packets under both normal conditions and special conditions where there is a small amount of network traffic with attack signature are presented. Experimental results show the model is able to improve the efficiency of data collection and strengthen the processing performance of IDS with little devaluation of detection precision of IDS.

Keywords: intrusion detection system, data collection, sampling, performance.

1. INTRODUCTION

With the increasing number of computer security events every year, more attention has been paid on network security. Intrusion detection system (IDS) has become an indispensable part of computer security (Denning, 1987). It is necessary for IDS to collect data with an effective and reliable manner. However, it is impossible for IDS to capture all packets due to the limitation of computer resources, especially in the large-scale network with heavy traffic. In addition, the proportion of network traffic with attack (intrusion) signature is commonly small, and capturing the whole traffic will degrade the efficiency of network

G. R. Liu et al. (eds.), Computational Methods, 1081–1085.

bandwidth utilization. Therefore, a more reasonable network traffic collection mechanism should be applied to IDS.

Sampling is an effective technique to monitor network traffic in real time, and has been applied to various network engineering and management applications, such as traffic report (Duffield and Grossglauser, 2001) and attack (intrusion) detection or analyses (Kodialam and Lakshman, 2003). In this chapter, simple random sampling technique of statistics is introduced to the procedure of data collection of IDS, and a new data collection model for IDS is provided.

2. DATA COLLECTION MODEL BASED ON SIMPLE RANDOM SAMPLING

Simple random sampling is the basic sampling technique where we select a group of subjects (a sample) for study from a larger group (a population). Each individual is chosen entirely by chance and each member of the population has an equal chance of being included in the sample.

2.1 Symbol token

In this chapter, all the packets in the network are treated as population, and individual packet is treated as population unit.

Assume the number of population unit is N, and let Y_1, $Y_2 \dots Y_N$ represent the measurements of these population units. The total number of packets with attack signature in the population is A. Let $Y_i = 1$ if the packet ($\mathrm{Y_i}$) with attack signature, otherwise $Y_i = 0 (i = 1, 2 \dots N)$. Then $P = A/N$ represents the population attack strength, and $A = \sum_{i=1}^{N} Y_i$ represents the population attack quantity.

Assume the sample size is n, and let $y_1, y_2 \dots y_n$ represent the measurements of these sample units. The total number of packets with attack signature in the sample is a. Let $y_i = 1$ if the packet (y_i) with attack signature, otherwise $y_i = 0 (i = 1, 2 \dots n)$, then $p = a/n$ represents the sample attack strength, and $a = \sum_{i=1}^{n} y_i$ represents the sample attack quantity.

2.2 Sampling size calculation

It is a very important matter for sampling technique to calculate the sample size. In the same condition, sampling error will decrease with the increasing sampling size. There is an inverse proportion relation between sampling error and the square root of sample size, and sampling error tends to stable after given phases. It is obvious unworthy to reduce sampling error by increasing the sample size. At this point, we can greatly reduce the sample size with slightly degradation of precision. In this chapter, it is the key matter for intrusion detection system to reduce data collection quantity as possible with little degradation of detection precision. The following part presents our effort on how to calculate the sample size of packets. The sampling theorem shows that when the sample size is big enough, the distribution of sampling error will be approximately a standard normal distribution. Give the parameter θ and its estimate $\hat{\theta}$, the variance of this estimate $\hat{\theta}$ is $V(\hat{\theta})$ and the standard deviation is $S(\hat{\theta})$. Let parameter θ is the proportion of packets with attack (intrusion) signature in the whole network traffic, i.e. population attack strength P, and parameter $\hat{\theta}$ is proportion of packets with attack (intrusion) signature in the sample, i.e. sample attack strength p. From sampling theorem, we have

$$V(p) = \frac{PQ}{n}\frac{N-n}{N-1} \tag{1}$$

where $Q = 1 - P = \frac{N-A}{N}$, and the absolute error

$$d = t\sqrt{V(p)} = t\sqrt{\frac{PQ}{n}\frac{N-n}{N-1}} \tag{2}$$

or the relative error

$$r = t\frac{\sqrt{V(p)}}{P} = \frac{t}{P}\sqrt{\frac{PQ}{n}\frac{N-n}{N-1}} \tag{3}$$

so

$$n = \frac{t^2\frac{PQ}{d^2}}{1 + \frac{1}{N}\left(\frac{t^2PQ}{d^2} - 1\right)} \tag{4}$$

or

$$n = \frac{t^2\frac{Q}{r^2P}}{1 + \frac{1}{N}\left(\frac{t^2Q}{r^2P} - 1\right)} \tag{5}$$

we can first calculate

$$n_0 = \frac{t^2PQ}{d^2} \text{ or } n_0 = \frac{t^2Q}{r^2P} \tag{6}$$

if $\frac{n_0}{N} < 0.05$, then n is approximately equal to n_0, otherwise n_0 needs to be modified as

$$n_0 = \frac{t^2 PQ}{d^2} \text{ or } n_0 = \frac{t^2 Q}{r^2 P} \tag{7}$$

If P is in the range of [0.2, 0.8], we can make a conservative estimate of the sample size according as PQ will be maximum value given $P = Q = 0.5$, then calculate the sample size using preceding formula (6) based on t, d and $PQ = 0.25$. If n_0/N cannot be neglected, then we need to modify the sample size.

2.3 Inverse sampling

If the proportion of network traffic with attack (intrusion) signature is very small, i.e. population attack quantity P is very small, relative error r will be better than absolute error d. While we just know that P is likely very small, but we can't reckon the exact range of P. Given the different estimated value of P, the calculated sample size of packets will appear marked difference. To handle this problem, we make use of inverse sampling method as shown in the work by Haldane (1945), that is, determine an integer m ($m > 1$), sample packet one by one until the number of packets with attack (intrusion) signature is equal to m. It can be proved that the mean of sample size n is

$$\mathrm{E(n)} = \frac{m}{P} \tag{8}$$

Therefore, if the proportion of network traffic with attack (intrusion) signature is extraordinary small, the sample size n is actually very big. For example, given $P = 0.00001$, $m = 27$, then the mean of $n = 2700000$.

3. EXPERIMENTS

We use IXIA400T (a special network test facility of IXIA Crop) to generate the large-scale background traffic in 1000M Ethernet, and insert specific attack traffic. The sample size of packets is calculated according to the preceding formulas, and corresponding detection ratio is recorded, compared with the detection ratio by capturing the whole network traffic in order to verify the detection precision of IDS. The population total is around 6,000,000. RealSecure is selected to detect attacks (intrusions) in our experiments, for its important role in security fields, and its detection ratio by capturing the whole network traffic is 93.75%. Considering the randomicity of sampling, we conduct

Table 1. Detection ratio with various sample sizes.

Sample size	Detection ratio		
38,032	87.5%	89.6%	85.4%
76,448	89.5%	87.5%	89.5%
15,3280	91.7%	93.75%	91.7%

sampling three times separately for every sample size. The detection results of RealSecure are as given in Table 1.

4. CONCLUSIONS

This model is able to reduce data collection quantity, improve the utilization ratio of network bandwidth, decrease the detection time, and raise the efficiency of detection with little degradation of detection precision. Especially in the large-scale network with heavy traffic, this model can strengthen the processing performance of IDS by the means of replacing dropping packets passively with sampling packets actively.

REFERENCES

1. D.E. Denning (1987), An intrusion-detection model, *IEEE Transactions on Software Engineering*, SE-13, pp. 222–232.
2. N.G. Duffield and M. Grossglauser (2001), Trajectory Sampling for Direct Traffic Observation, *IEEE/ACM Trans. on Networking*, 9, 3, pp. 280–292.
3. M. Kodialam and T.V. Lakshman (2003), Detecting network intrusions via sampling: A game theoretic approach, *IEEE INFOCOM 2003—The Conference on Computer Communications*, pp. 1880–1889.
4. J.B. Haldane (1945), On a method of estimating frequencies. *Biometrika* 33:222–225.

DESTRUCTIVE EXTENSION RULE IN PROPOSITION MODAL LOGIC K

X. Wu and J.G. Sun
College of Computer Science and Technology, Jilin University, ChangChun 130012,
P. R. China
Key Laboratory of Symbolic Computation and Knowledge Engineer of Ministry of Education,
Jilin University,
ChangChun 130012, P. R. China

Abstract Modal logics are good candidates for a formal theory of agents, the efficiency of reasoning method in modal logics is very important, because it determines whether or not the reasoning method can be used widely in the systems based on agent. Theorem proving based on the extension rule we presented is a new method. Firstly, this paper gives the version of non-clausal extension rule algorithm. Next, we present the extension rule in proposition modal logic K, namely we provide a new proof method for modal logics. And then we give the proof of its soundness and completeness.

Keywords: theorem proving, extension rule, modal logics, destructive.

1. INTRODUCTION

Since the pioneering publications of the 1950s, automated theorem proving has been matured into one of the most advanced areas of computer science. Extension rule based approach is a new theorem proving method we presented recently [1]. For a number of years, modal logics have been applied outside pure logic in areas such as formal methods, theoretical computer science and artificial intelligence. In all of those modal systems, the idea of *context* is explicit or implicit in various systems proposed for modal logics. In a Kripke model, occurrences of the same literal may not mean the same thing in different contexts. One solution has been to introduce notation for context, via prefixes or world paths [2]. Another solution is destructive method [3]. It is called destructive method because it looses information during a context shift. Logics

G. R. Liu et al. (eds.), Computational Methods, 1087–1091.

amenable to such treatment tend to be those Kripke models do not involve symmetry.

In Section 2, we give several reduction rules at first. Moreover, we give a version of non-clausal extension rule algorithm and prove its soundness and completeness. In Section 4, we present the extension rule method in proposition modal logic K and prove its soundness and completeness. In addition, we give the concrete algorithm. A conclusion is drawn in the final part.

2. NON-CLAUSAL EXTENSION RULE

We write P, Q, ... for propositional variables, X, X_1, ... for arbitrary formulas. Formulas are divided into two categories: α formulas and β formulas. In a similar way the modal cases are divided up into ν formulas and π formulas [3]. We will treat $[X_1, \ldots, X_n]$ as synonymous with $X_1 \vee \ldots \vee X_n$ and $\langle X_1, \ldots, X_n \rangle$ as synonymous with $X_1 \wedge \ldots \wedge X_n$. If $X_1, \ldots, X_n$ are formulae, we will refer to $[X_1, \ldots, X_n]$ as a clause. Likewise if each C_i is a clause, we will refer to $< C_1, \ldots, C_n >$ as a clause list. Finally, if $L_1, \ldots, L_n$ are clause lists, we call $[L_1, \ldots, L_n]$ a block.

The extension rule method uses the inverse of resolution together with the inclusion–exclusion principle to solve TP problems. The details can be found in Ref. [1].

Definition 1. Given a clause C and a set M: $C' = \{C \vee a, C \vee \neg a \mid$ 'a' is an atom, $a \in M$, '¬a'and 'a' does not appear in C$\}$. We call the operation proceeding from C to C′ the *extension rule* on C. We call C′ the result of the extension rule.

Definition 2. We say C_1, C_2 is a partition of the clause C if C_1 and C_2 are disjoint and the members of C_1together with those of C_2 are exactly the members of C.

The extension rule algorithm, we give in this paper does not require a previous reduction to clause form. Namely the destructive extension rule method is a non-clause style system, it combine reducing to normal form with checking the unsatisfiability and does them alternately. This will be of more importance when we come to the modal case later on. Also, *blocks* are present, they will become significant when we introduce modal rules in the next section. Now, the rules for deriving one block from another are as follows. *Double Negation Rule (RDN)*: An occurrence of a formula $\neg\neg Z$ may be replaced by an occurrence of Z**.** *Conjunction Rule (RC)*: If an α formula occurs in a clause C, the clause C may be replaced by two clauses, C_1 and C_2, which are like C except that C_1 contains α_1 in place of α and C_2 contains α_2 in place of α. *Disjunction Rule (RD)*: An occurrence of a β formula may be replaced by the two formulas β_1 and β_2. *Special Case Rule (RSC)*: If a clause list L contains a clause C, and C_1,

C_2 is a partition of C, then L can be replaced by two clause lists, L_1 and L_2, where L_1 is like L except that it contains C_1 in place of C, and L_2 is like L but with C_2 in place of C. Let $R = \{RDN, RC, RD\}$. Now we give the non-clause style Extension-Rule algorithm NCER in proposition logic.

Algorithm NCER:

```
1. Set F is the formula we want to prove, $k = 2$
2. while there are no >= k clauses which have no connectives in F
   Loop
     $F_1$ = the clause list got by using R to reduce F
   If $F == F_1$ then return satisfiable
         Else $F = F_1$
   Endloop
3. Let $\Sigma = \{C_1, C_2, \dots C_m\}$ be the set of the clauses having no connectives in
   F, $|\Sigma| = m$. Call the ER1Algorithm to check $\Sigma$'s satisfiability.
4. If $\Sigma$ is unsatisfiable then return unsatisfiable
   Else $k = m + 1$; go to step 2
```

Lemma 1. Let clause list L' be derived from a clause list L by the Double Negation Rule, the Conjunction Rule and the Disjunction Rule. Then L is satisfiable if and only if L' is satisfiable.

Lemma 2. Let B' be derived from the block B by the Special Case Rule. Then B is satisfiable if and only if B' is satisfiable.

Theorem 1. Algorithm NCER is sound and complete for proposition logic theorem proving.

3. EXTENSION RULE IN PROPOSITION MODAL LOGIC K

Definition 3. For each clause C we define two related clauses as follows. $C^{\#}$ consists of all formulas ν_0 such that ν is in C. C^b consists of all formulas π_0 such that π is in C.

The extension rule system we propose includes all the rules of Section 2, including the Special Case Rule, together with one more Reduction Rule. *K-π Rule (RKπ)*: Suppose L is a clause list containing a clause C consisting entirely of π formulas. Then L may be replaced by the clause list L' containing the following clauses: C^b and, for each clause S in L consisting entirely of ν formulas, the clause $S^{\#}$. Now, we give the destructive extension rule algorithm DMKER in proposition modal logic K.

Lemma 3. Let the clause list L' be derived from a clause list L by the K-π Rule. So if L is satisfiable then L' is satisfiable.

Theorem 2. In propositional modal logic K, theorem proving based on the destructive extension rule is sound and complete.

Algorithm DMKER (F):

```
1. Set F is the modal formula we want to prove, $k = 2$.
2. while there are no >= k clauses which have no connectives in F
   Loop
   a. $F_1$ = the clause list got by using R to reduce F
   b. $F_2 = F_1$
   c. If we can use the RSC to reduce $F_1$ then
      a) Denote the returned clause lists by $FS_1$, $FS_2$
      b) If DMKER($FS_1$) = = unsatisfiable and DMKER($FS_2$) = = unsatisfiable then return unsatisfiable
         Else return satisfiable
   d. If $F_1$ includes clauses consisting entirely of $\pi$ formulas then
      a) Choose a clause C
      b) $F_2$ = the clause list got by using the $K\pi$ on the clause C to reduce $F_1$
      c) $F_1$ = delete C from $F_1$
Else if $F_1$ just includes clauses consisting entirely of $\nu$ formulas then $F_2$ = the clause list got by using the $K\pi$ to reduce $F_1$
   e. If F = = $F_2$ and $F_1$ does not include the clause consisting entirely of $\pi$ formulas then return satisfiable
      Else if F = = $F_2$ then $F = F_1$; $K = 2$;go to step d
   f. $F = F_2$;
      Endloop
3. Let $\Sigma = \{C_1, C_2, \ldots C_m\}$ be the set of the clauses having no connectives in F, $|\Sigma| = m$. Call the ER1Algorithm to check $\Sigma$'s satisfiability.
4. If $\Sigma$ is unsatisfiable then return unsatisfiable
   Else $k = m + 1$; go to step 2
```

4. CONCLUSIONS

Agents have become necessary due to the vast extent and scattered nature of the information have landscape. It is accepted that multi-agent systems are important to many areas in computer science, starting from specifying and verifying multi-threading systems and multi-processor computers to modelling populations of robots that have animal or human-like behaviour. Now there has been a lot of deep research on this aspect, including reasoning methods [4], interactions [5], general frameworks [6], etc. Concerning the automated theorem proving in modal logics, there are approximately three ways at present. One is the method based on tableaux [7]. The other is extending the resolution to modal logics directly [8]. The third is the method based on translation [9]. We introduce extension rule into modal logic directly by the destructive method in this paper, because it has an advantage of simplicity, that is, it is easier to understand. We strongly urge a multiplicity of approaches, and the full investigation of all. Limited by the space, we only discuss the things in proposition modal logic K, we will consider the destructive extension rule in modal logic T, K4, S4 and S5 as well as in the first-order modal logics in another paper.

REFERENCES

1. H. Lin, J.G. Sun and Y.M. Zhang (2003), Theorem proving based on extension rule. *Journal of Automated Reasoning*, 31, pp. 11–21.
2. H.J. Ohlbach (1988), A Resolution Calculus for Modal Logics, CADE'88, pp. 23–26.
3. M.C. Fitting (1990), Destructive modal resolution. *Journal of Logic and computation*, 1, pp. 11–21, 83–97.
4. M. Fisher, C. Dixon and M. Peim (2001), Clausal temporal resolution. *ACM Transactions on Computational Logic*, 2, pp. 1–44.
5. A. Artale and E. Franconi (1998), A temporal description logic for reasoning about action and plans. *Journal of Artificial Intelligence Research*, 9, pp. 463–506.
6. U. Hustadt, C. Dixson, R.A. Schmidt and M. Fisher (2000), *Normal Forms and Proofs in Combined Modal and Temporal Logics,* FroCoS'2000, France, pp. 73–87.
7. R.S. Liu, J.G. Sun and X.H. Liu (1998), Epistemic Logic (3): on semantic tableau proof procedure. *Chinese Journal of Computers*, 21, Suppl., pp. 1–8.
8. J.G. Sun, R.S. Liu, and R. Chen (1999), An extension on marked modal resolution. *Chinese Journal of Computers*, 2, pp. 113–119.
9. R.A. Schmidt and U. Hustadt (2003), *A Principle for Incorporating Axioms into the First-Order Translation of Modal Formulae, CADE'2003,* USA, pp. 412–426.

A NOVEL PARTICLE SWARM OPTIMIZATION-BASED APPROACH FOR JOB-SHOP SCHEDULING

H.W. Ge[1], Y.H. Lu[1,2], Y. Zhou[1], X.C. Guo[1] and Y.C. Liang[1]
[1]*College of Computer Science, Jilin University, Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, Changchun 130012, China*
[2]*College of Computer Science, Northeast Normal University, Changchun 130024, China*

Abstract The goal of this article is to develop an efficient scheduling method based a proposed novel discrete particle swarm optimization (PSO) for the job-shop scheduling problem (JSSP). This paper also introduces a novel concept for distance and velocity of particles in the PSO to pave the way for the JSSP. The proposed method effectively exploits the capabilities of distributed and parallel computing systems. Simulation results show that the proposed algorithm can obtain high quality solutions for typical benchmark problems.

Keywords: job-shop scheduling, particle swarm optimization, similarity measure.

1. INTRODUCTION

The job-shop scheduling problem (JSSP) is a very important practical problem in both fields of production management and combinatorial optimization [1]. Because JSSP is among the worst members of the class of NP-complete problems there remains much room for improvement in current techniques and exploitation of new methods. In this paper, a new scheduling algorithm based on the principles of the particle swarm optimization (PSO) is proposed. PSO [2] is an evolutionary computation technique based on swarm intelligence. It follows a collaborative population-based search. PSO effectively exploits the distributed and parallel computing capabilities, and it has great capability of escaping local optima.

G. R. Liu et al. (eds.), Computational Methods, 1093–1098.

2. JOB-SHOP SCHEDULING PROBLEM

In JSSP, n jobs have to be processed on m different machines, where the processing of each job consists of m operations performed on these machines in a specified sequence. The operation of job i has to be performed on machine j with an uninterrupted processing time t_{ij}. A schedule defines the time intervals in which the operations are processed, but it is feasible only if it complies with the constraints that each machine can process only one operation each time and the operation sequence is respected to each job. The goal for optimization is to find the sequence of n jobs to be completed on m machines such that the makespan is minimized. The solution to the JSSP can be represented as the operation permutation of the jobs on each machine. The total number of all possible schedules is $(n!)^m$. Obviously, it is impossible to exhaust all the alternatives for finding the optimal solution. So it is necessary to restrict the searching space and guide the searching process. There are three types of feasible schedules in JSSP, namely inadmissible, semi-active and active schedules. The optimal schedule is guaranteed to be an active schedule. So we only need to search the optimal solution in the set of active schedules.

3. SCHEDULING ALGORITHM BASED ON PSO

3.1 Representation and bilevel encoding

In solving JSSP by PSO, each particle is requested to represent a potential solution of the problem, namely a scheduling. In this paper, an operation-based representation is adopted. By scanning the permutation from left to right, the kth occurrence of a job number refers to the kth operation in the technological sequence of this job. The prominent advantage of the representation is to eliminate the ‘deadlocks’, but notice that, while each individual is related to a unique active scheduling by the process of decoding discussed in the following section, a scheduling can be related to more than one individual. In other words, there is some redundancy in the search space. The job sequence of a machine can be obtained easily by scanning a particle from left to right to find out the job order for the same machine. Repeating the scanning for every machine, we can obtain a new permutation which is called the bilevel encoding of the particle. If the bilevel encoding expressions of two particles are the same, then their corresponding schedules are just the same. If a particle is redundant, we swap two operations of different jobs randomly until it is larruping. The intention of the bilevel encoding is to eliminate the redundancy, and we still use the operation-based representation in the PSO. The initial population is produced using the algorithm proposed by Giffler and Thompson.

3.2 Objective function and fitness function

The objective function most largely used currently for the JSSP is the minimization of the makespan. Mathematically, the JSSP is to search a scheduling such that

$$\min(T(JM)) = \min\{\max[T(1), T(2), \ldots, T(i), \ldots, T(m)]\}, \qquad (1)$$

where $T(i)$ is the final completion time of machining on machine i, $T(JM)$ is the final completion time of all the jobs. The fitness of particle i is evaluated by the formulation that $f_i = 100 \times opt/T_i(JM)$, where opt is the most optimal makespan known. The value of the objective function can be obtained by the decoding process that assigns operations to the machines at their earliest possible starting time by the technological order of each job, scanning the permutation from left to right. But the scheduling acquired is only a semi-active. Then the active decoding is applied, which checks the possible blank time interval before appending an operation at the last position, and fills the first blank interval before the last operation.

3.3 The formulation of proposed novel PSO algorithm

In this section, we describe the formulations of the proposed novel PSO algorithm for the JSSP. Firstly, the concept of the difference is extended. We present the similarity measure of two particles. Let the ith and the jth particles in a D-dimensional space be represented as $X_i = (x_{i1}, \ldots, x_{id}, \ldots x_{iD})$ and $X_j = (x_{j1}, \ldots, x_{jd}, \ldots x_{jD})$, respectively. Define functions

$$s(k) = \begin{cases} 1 & \textit{if } x_{ik} = x_{jk} \\ 0 & \textit{if } x_{ik} \neq x_{jk} \end{cases} \quad \text{and} \quad S(X_i, X_j) = \sum_{k=1}^{D} s(k) \qquad (2)$$

where $S(X_i, X_j)$ is called the similarity measure between particle X_i and particleX_j. The difference of the locations between two particles, namely ‘distance’, is redefined by the following equation:

$$dis(X_i - X_j) = k \cdot [\alpha \cdot |f(X_i) - f(X_j)|/100 + \beta \cdot (D - S(X_i, X_j))/D], \qquad (3)$$

where $f(X_i)$ and $f(X_j)$ are the fitness values of the particles X_i and X_j respectively, D is the dimension of particles, k is an acceleration coefficient, α and β are positive weights. Correspondingly, the concept of velocity is also extended. The velocity is defined the times of ‘adjustment’. The brief outline of the adjustment algorithm is as follows.

Let $X = (x_1, \ldots, x_k, \ldots, x_D)$ and $Y = (y_1, \ldots, y_k, \ldots, y_D)$ be two particles in a D-dimensional space.

Step 1: Select an index k of location and set an indicator $m = 0$, if $x_k = y_k = s$, where s is the number on location k, go to Step 2, else, namely $x_k = s$ and $y_k \neq s$, go to Step 3.
Step 2: Scan set $X' = \{x_i | 1 \leq i < k\}$ and set $Y' = \{y_i \mid 1 \leq i < k\}$ from left to right respectively to find out the times that s appears, and denote them as t_x and t_y respectively. If $t_x > t_y$ go to Step 4; if $t_x < t_y$, go to Step 5; if $t_x = t_y$, go to Step 6.
Step 3: Select a location index j in the particle Y randomly, which satisfies $y_j = s$ and $y_j \neq x_j$. Swap y_j and y_k, and set the indicator $m = 1$, then go to Step 2.
Step 4: Swap s (which appears after location k) and y_j (which satisfies $y_j \neq x_j$, $y_j \neq s$ and $j < k$) until $t_x = t_y$, then go to Step 7.
Step 5: Swap s (which appears in the front of location k) and y_j (which satisfies $y_j \neq x_j, y_j \neq s$ and $j > k$) until $t_x = t_y$, then go to Step 7.
Step 6: If $m = 0$, reselect an index k randomly; if $m = 1$, go to Step 7.
Step 7: Terminate.

Denote '$\oplus$' as the adjustment operation, then the proposed novel PSO algorithm could be performed by the following equations:

$$v_{id} = \text{int}[w v_{id} + c_1 r_1 dis(p_{id} - x_{id}) + c_2 r_2 dis(p_{gd} - x_{gd})], \tag{4}$$

$$x_{id} = x_{id} \oplus v_{id}, \tag{5}$$

where int[$\cdot$] is the truncation function, $dis(\cdot)$ can be computed by Equation (3). Besides, two heuristic operations are used to speed up the convergence. For each particle, swap the locations of λ different elements randomly, and replace the original particle by the best one. If there is no improvement of the best solution for a certain number of iterations, then re-randomize some particles in the population by a certain percentage.

4. NUMERICAL SIMULATION RESULTS

The performance of the proposed PSO-based algorithm for JSSP is examined by the well-known 6 × 6 benchmark problem, which is known that the minimal makespan is 55. The following particle is an optimal solution obtained by the proposed algorithm:

(3,3,3,2,2,2,1,1,4,4,3,5,6,6,6,6,1,4,4,2,5,5,3,3,5,1,1,2,2,4,4,5,5,6,6,1)

and the makespan obtained by the active decoding is equal to the best solution found so far. Table 1 shows the start and end time of each operation of the jobs. Figure 1 shows the Gantt chart of the active scheduling obtained by the particle.

Table 1. An optimal schedule for the 6 × 6 benchmark problem.

	Machine 1		Machine 2		Machine 3		Machine 4		Machine 5		Machine 6	
Job	Start	End	Start	End	Start	End	Start	End	Start	End	Start	End
1	6	9	16	22	5	6	30	37	49	55	42	45
2	38	48	0	8	8	13	48	52	13	23	28	38
3	18	27	27	28	0	5	5	9	30	37	9	17
4	13	18	8	13	22	27	27	30	37	45	45	54
5	48	51	22	25	13	22	52	53	25	30	38	42
6	28	38	13	16	49	50	16	19	45	49	19	28

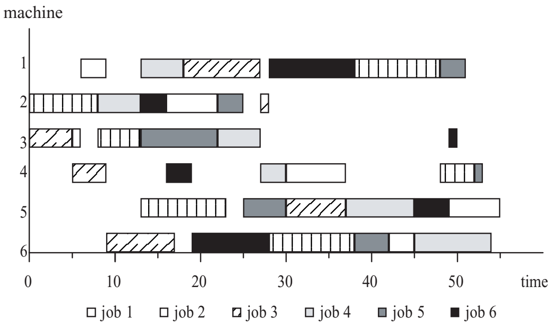


Figure 1. The Gantt chart of an optimal schedule.

5. CONCLUSIONS

A promising novel particle swarm optimization-based algorithm for job-shop scheduling problem is proposed. Numerical experiment results are encouraging and show the real potential of the algorithm. It indicates a novel approach to the combinatorial optimization problems, even if the hard NP-complete problems. Investigation on further testing the performance of the technique as well as theoretical study is in progress.

ACKNOWLEDGEMENT

The authors are grateful to the support of the key science-technology project of the National Education Ministry of China under Grant No. 02090 and the science-technology development project of Jilin Province of China under Grant No. 20030520.

REFERENCES

1. J. Blazewicz, W. Domschke and E. Pesch (1996), The job shop scheduling problem: conventional and new solution techniques. *European Journal of Operational Research*, 93, pp. 1–33.
2. J. Kennedy and R. Eberhart (1995), Particle swarm optimization. *Proceedings of the IEEE International Conference on Neural Networks*, 4, pp. 1942–1948.

UNCALIBRATED ROBOTIC ARM VISUAL SERVO CONTROL

Zhang Qizhi[1], Ge Xinsheng[1] and Liang Yanchun[2]
[1]*Department of Computer Science and Automation, Beijing Institute of Machinery,*
P.O. Box 2865, Beijing, 100085, P.R. China
[2]*College of Computer Science and Technology, Jilin University, Changchun, 130012,*
P.R. China

Abstract This paper focuses on the visual servo control for an uncalibrated robotic arm with an eye-in-hand camera. Without a priori knowledge of the robotic arm's kinematic model or camera calibration, the control system can track a moving object through visual feedback. Two methods are proposed to resolve the output constraint of control. The joint angle feedback is used to correct the calculating values of the joint angles in the first control method. The output of controller is re-scaled to ensure that the output constraint of controller is not violated in the second method. The performances of the proposed control methods are illustrated by the computer simulations.

Keywords: robotic arm, visual servoing, uncalibration, feedback control.

1. INTRODUCTION

The visual servoing control has been an active area of research for many years. There has been significant progress on visual servoing control since 1973 [1]. There are two basic approaches to visual servo control, Position-Based Visual Servoing (PBVS) and Image-Based Visual Servoing (IBVS). In PBVS system, the object's 3D pose is recovered through computer vision and the error between the current pose and the desired pose is computed in the Cartesian task space. In IBVS system, the error is measured between the image features in two images and it is mapped directly to actuator commends. The placement of camera can be put on the robot ("eye in hand") observing the object, or fixed in the environment and observing the object and the robot. A priori knowledge on the kinematical structure and the camera parameters is required in the

G. R. Liu et al. (eds.), Computational Methods, 1099–1104.

model-dependent visual servoing system. Recently, an uncalibrated eye-in-hand visual servoing system has been proposed [2]. This method does not require calibrated kinematics and camera models, so it is robust to the signal noise and calibration errors. The robot control is achieved by quasi-Newton method and estimating the composite Jacobian matrix by Broyden's method at each step. However, the desired joint angles may cause saturation of the speed and/or the torque delivered by the joint actuators. This saturation may occur in the initial stage of the tracking process, in which the end-effector of robot is always far from the moving target point. Large joint angles increments are required to track the moving target. Because no joint angles feedback is used in the visual servoing control algorithm, the joint angles calculated by the control algorithm will be drifted with the saturation of the joint actuators. In this paper, two methods are proposed to improve the performances of visual servoing control system, in which the saturation of the joint actuators is considered. The joint angle feedback is used to correct the calculating values of the joint angles in the first control method. The output of controller is re-scaled to ensure that the output constraint of controller is not violated in the second method.

2. UNCALIBRATED CONTROL FOR EYE-IN-HAND VISUAL SERVOING

Assume that the desired behaviour of the robot is to track a simple target. In the image plane, the moving target is at position $y^*(\theta, t)$, the end-effector of the robot is at position $y(\theta,\ t)$. The error function in the image plane is

$$f(\theta,\ t) = y(\theta,\ t) - y^*(\theta,\ t) \tag{1}$$

where θ represents the joint angle and t represents time. Minimizing the squared error, one has [2]

$$\theta_{k+1} = \theta_k - (\hat{J}_k^T \hat{J}_k)^{-1} \hat{J}_k^T \left(f_k + \frac{\partial f_k}{\partial t} h_t \right) \tag{2}$$

where $f_k = f(\theta_k, t_k)$ is the error at time t_k and $\hat{J}_k$ represent the kth approximation to Jacobian $J_k = \frac{\partial f_k}{\partial t}$ and $h_k = t_k - t_{k-1}$ is the time increment. According to Equation (2), the desired changes in joint angles can be computed. The parameters in Equation (2) can be estimated by Partitioned Broyden's method as follows [2]

$$\tilde{J}_k = \tilde{J}_{k-1} + (\Delta f - \tilde{J}_{k-1}\tilde{h})(\lambda + \tilde{h}^T \tilde{P}_{k-1}\tilde{h})^{-1}\tilde{h}^T \tilde{P}_{k-1} \tag{3}$$

$$\tilde{P}_k = \frac{1}{\lambda}(\tilde{P}_{k-1} - \tilde{P}_{k-1}\tilde{h}(\lambda + \tilde{h}^T \tilde{P}_{k-1}\tilde{h})^{-1}\tilde{h}^T \tilde{P}_{k-1}) \tag{4}$$

Varying the weighting parameter λ between 0 and 1 we can change the memory of the scheme. An algorithm for the combination of the Partitioned Broyden's method and Gauss-Newton controller given by Equations (1–4) can be found in [2]. In every interval, the limitations of the increments of the joint angles are not considered in the algorithms, but the maximum accelerations of joint angles are limited in a real robotic system. If the desired changes of joint angles exceed the limits of the robotic system, the values of the desired joint angles will be drifted from the actual position.

3. CONSTRAINT VISUAL SERVOING CONTROL

The individual joint level controller in Equation (2) can be rearranged as

$$\theta_{k+1} = \theta_k + \Delta\theta \tag{5}$$

where $\Delta\theta = -(\hat{J}_k^T \hat{J}_k)^{-1} \hat{J}_k^T (f_k + \frac{\partial f_k}{\partial t} h_t)$ is the increment of the joint angle.

Firstly, the joint angles feedback can be introduced to solve this drawback of the desired joint angles drifted from the actual position

$$\theta_{k+1} = \theta_k^p + \Delta\theta \tag{6}$$

where θ_k^p are the measurements of the joint angles. Because the joint angles feedback are used, the desired joint angles θ_{k+1} are corrected in each interval, and the drift will not be taken place. The second method is to re-scale the increment of the desired joint angles to ensure the joint actuators not to be saturated

$$\theta_{k+1} = \theta_k + \theta_{max} \Delta\theta / \left\| \Delta\theta \right\| \tag{7}$$

where θ_{max} is the maximum increment of the robotic joint angles in each interval. Because the desired joint angles do not exceed the ability of the joint actuators, the desired joint angles will not be drifted.

4. SIMULATIONS AND RESULTS

The proposed new controllers are implemented on a two link planar robot using the Matlab. Both lengths of link 1 and link 2 are 0.6 m. In the simulations, the target point is given a circular motion in a plane $(x, y, z) = (0.5 + 0.3 \ \cos\ (k\omega/f\text{s}),\ 0.5 + 0.2 \ \sin\ (k\omega/f\text{s}), 1)$. Where k is the iteration number, $\omega = 0.8$ rad/s is the frequency and $f\text{s} = 20\ Hz$ is the sampling frequency. The image is projected using a simulated camera with 640-pixels/m. Uniformly distributed random image noise between ±1 *pixels* is added to the image feature point of the target object. The maximum increment of the robotic

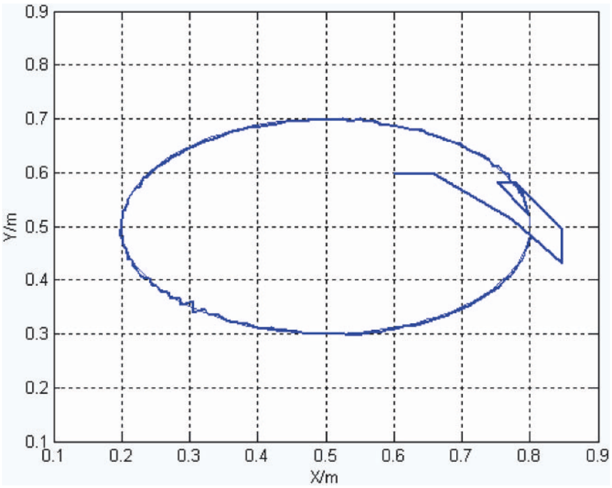


Figure 1. Tracking target by IBVS.

joint angles in each interval is selected as $\theta_{max} = 0.1$ rad. The forgetting factor is selected as $\lambda = 0.99$ in all the simulated examples.

The tracking trajectory using IBVS controller is shown in Figure 1, and the result using the increment re-scaled is shown in Figure 2. The Image error for using IBVS controller is shown in Figure 3, and the result using IBVS controller with the increment re-scaled is shown in Figure 4. Comparing Figures 1–4, it can be seen that, for a moving target, the fast and accurate tracking can be achieved using different controllers. But at the initial stage of the tracking, the

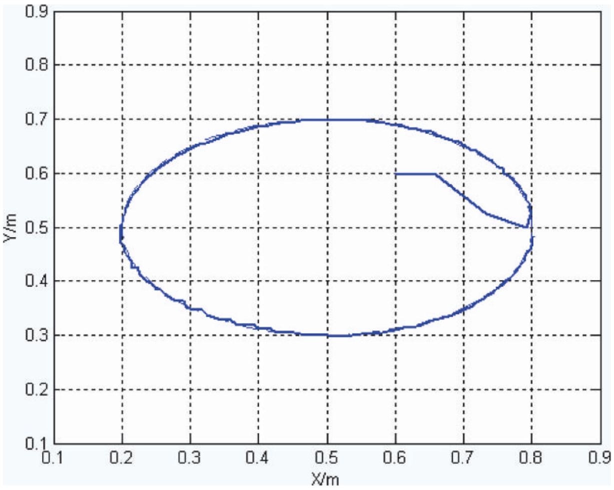


Figure 2. Tracking target by new IBVS.

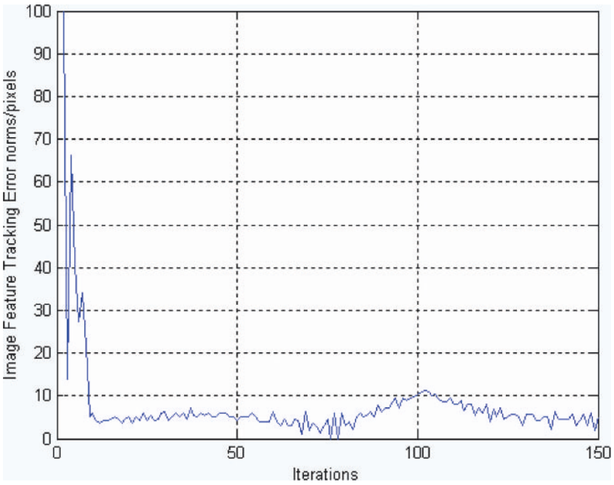


Figure 3. Image error for IBVS control.

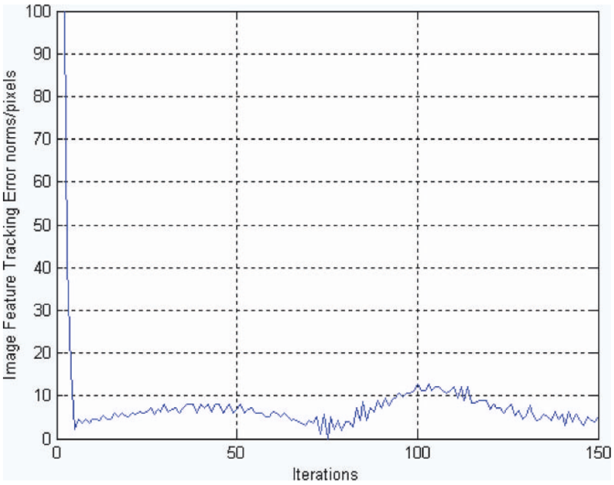


Figure 4. Image error for new IBVS.

image errors for IBVS controller are very large. After 10 iterations, the average RMS tracking image error is about 5 pixels.

5. CONCLUSIONS

In practice, the desired joint angles may cause saturation of the joint actuators, and the joint angles calculated by the control algorithm will be drifted with the saturation of the joint actuators. Two methods are proposed to improve the performances of the visual servoing control system, in which the saturation of the joint actuators are considered. The joint angle feedback is used to correct the calculating values of the joint angles in the first control method. The output of the controller is re-scaled to ensure that the increment of desired joint angles does not exceed the ability of robotic actuators in the second method.

ACKNOWLEDGEMENTS

This research is supported by National Natural Science Foundation of P. R. China (10372014).

REFERENCES

1. N.R. Gans, S.A. Hutchinson and P.I. Corke (2003), Performance tests for visual servo control systems, with application to partitioned approaches to visual servo control. *International Journal of Robotics Research*, 22, 955–981.
2. J.A. Piepmeier and H. Lipkin (2003), Uncalibrated eye-in-hand visual servoing. *International Journal of Robotics Research*, 22, 805–819.

INTERLINEATION AND INTERFLATION FUNCTIONS OF MANY VARIABLES (BLENDING FUNCTION INTERPOLATION) AND ECONOMICAL ALGORITHMS IN THE APPROXIMATION THEORY

Oleg N. Lytvyn
Department of Applied Mathematic Ukrainian Engineering-Pedagogical Academy,
16 Universitatska street, Kharkiv, Ukraine, 61003

Abstract Interflation of the function $f(x_1, \ldots, x_n)$ of the n variables with help of the its traces (and traces of its derivatives of order $\le N$) on the M surfaces the dimension m is recovery (possible, exactly) f. If $m = 0$ this is interpolation on M points (for $n \ge 1$). If $m = 1$ (for $n \ge 2$) it is interlineation (blending function interpolation) on M lines. In this paper the review of last achievements and some applications interflation, interlineation functions and blending approximation functions for construction the economical algorithms in the approximation theory is given.

Keywords: interpolation, interlineation, interflation, blending function interpolation.

1. INTRODUCTION.

Interflation can be used in approximation theory, in methods LIDE and NIDE (methods of the reduction to the systems of the ordinary linear or non linear-integro-differential equations) the solution of the boundary value problems, in cartography, in computer tomography, in signals processing, for the description of surfaces of automobiles, planes, space bodies, etc. References in [1].

2. SOME DEFINITIONS

Let $n,\ M \in N,\ m,\ N \in \overset{0}{N}$—are given numbers; $\Pi_k,\ k = \overline{1, M}$—are given m–dimensional surfaces in R^n $(0 \le m < n)$; $\varphi_{k,p}(x)|_{\Pi_k} = L_{k,p} f(x)|_{\Pi_k}$,

G. R. Liu et al. (eds.), Computational Methods, 1105–1109.

Table 1. Comparison interpolation, interlineations and interflation.

The Approximation Method	Kind of the used information
Interpolation of the functions one or more variables	The values of the $f(x)$ and its derivatives (until fixed order) on some given points
Interlineation of the functions two or more variables	The traces of the $f(x)$ and its derivatives (until fixed order) on some given lines
Interflation of the functions three or more variables	The traces of the $f(x)$ and its derivatives (until fixed order) on some given surfaces

$k = \overline{1, M},\ p = \overline{0, N}$ are given traces of some function $f(x)$ and traces of some operators $L_{k,p} f(x)$. Function f can be the unknown. $L_{k,p} f(x)$ can be partial derivatives or normal derivatives $L_{k,p} f(x)|_{\Pi_k} = \partial^p f(x)/\partial v_k^p|_{\Pi_k},\ p = \overline{0, N}$ for case $m = n - 1$ and so on; $\|\cdot\| = \|\cdot\|_C$.

Definition 1. Operators $O(\{\varphi_{k,p}\};\ x) := O(\{L_{k,p}\},\ \{\Pi_k\},\ \{\varphi_{k,p}\},\ x)$ we shall name the interflation operators if $L_{\ell,q} O(\{\varphi_{k,p}\},\ x)|_{\Pi_\ell} = \varphi_{\ell,q}(x)|_{\Pi},\ \ell = \overline{1, M},\ q = \overline{0, N}$. If $m = 0$, then $O(\{\varphi_{k,p}\},\ x)$ are interpolation operators on M points. If $m = 1, n \geq 2$ then Π_k are lines in R^n, $O(\{\varphi_{k,p}\},\ x)$ are interlineation operators.

Definition 2. Let $O(\{\varphi_{k,p}\},\ x) = \sum_{\ell=1}^{M} \sum_{q=0}^{N} \gamma_{\ell,q}(\{\varphi_{k,p}\},\ x) h_{\ell,q}(x)$, where $h_{\ell,q}(x) = h_{\ell,q}(\{L_{k,p}\},\ \{\Pi_k\}, \{x\}$—some auxiliary functions, which are independence from approximating function $f(x)$ and $\gamma_{\ell,q}(\{\varphi_{k,p}\},\ x) = \gamma_{\ell,q}(\{L_{k,p}\}, \{\Pi_k\},\ \{\varphi_{k,p}\},\ x)$ are linear operators of functions $\varphi_{k,p},\ k = \overline{1, M},\ p = \overline{0, N}$. Then we shall call $O(\{\varphi_{k,p}\},\ x)$ linear, interflation (interlineations, interpolation) operators. In another case there are non-linear interflation (interlineation, interpolation) operators.

Definition 3. Let auxiliary functions $h_{\ell,q}(x) = h_{\ell,q}(\{L_{k,p}\}, \{\Pi_k\}, x)$ are rational, polynomial, trigonometrical or spline—functions or functions, which are using R-functions. Then we shall call $O(\{\varphi_{k,p}\}, x)$ rational, polynomial, trigonometrical, spline and so on interflation (interlineation, interpolation) operators.

Definition 4. If $f(x) \in C^r(R^n),\ r \geq N \geq 1$ and $O(\{\varphi_{k,p}\},\ x) \in C^r(R^n)$ then we shall call that operators $O(\{\varphi_{k,p}\},\ x)$ preserve of the class differentiability $C^r(R^n)$ of the function $f(x)$. In another case we shall call that operators $O(\{\varphi_{k,p}\},\ x)$ not preserve of the class $C^r\ (R^n)$ differentiability of the function $f(x)$.

Definition 5. If $\exists \ell, q : L_{\ell,q} O(\{\varphi_{k,p}\},\ x)|_{\Pi_\ell} \neq \varphi_{\ell,q}(x)|_{\Pi_\ell}$ that $O(\{\varphi_{k,p}\},\ x)$ are rational, polynomial, trigonometrical, spline and so on blending approximation operators.

3. SOME SELECTED RESULTS

1. **Interlineation operators without preserving of the class $C^r\ (R^2),\ r \geq 1$.**

1.1. *Rational interlineation operators on M lines.* Let $n = 2$ and

$$\begin{aligned}
\Pi_k : \omega_k(x) :&= a_k x_1 + b_k x_2 - \gamma_k = 0, k = \overline{1, M},\ a_k^2 + b_k^2 = 1\\
\varphi_{k,s}(x) &= \partial^s f/\partial v_k^s(x)|_{\Pi k} = \partial^s f/\partial v_k^s(x_1, (\gamma_k - a_k x_1)/b_k)\\
&\text{if } b_k \neq 0; v_k = \nabla\omega_k(x) = (a_k, b_k)\\
\text{or } \varphi_{k,s}(x) &= \partial^s f/\partial v_k^s(x)|_{\Pi_k} = \partial^s f/\partial v_k^s((\gamma_k - b_k x_2)/a_k, x_2)\\
&\text{if } a_k \neq 0\\
O_{k,N} f(x) &= \sum_{s=0}^{N} \varphi_{k,s}(x - \omega_k\,(x)\nabla\omega_k(x))\frac{\omega_k^s(x)}{s!}\\
H_k(x) &= \prod_{\substack{i=1\\ i\neq k}}^{M} \omega_i^{N^*}(x) \Big/ \prod_{\ell=1}^{M} \prod_{\substack{i=1\\ i\neq \ell}}^{M} \omega_i^{N^*}(x), \quad N^* = N + 1,\\
&\textit{if } N = 2q + 1, \quad q \in \overset{0}{N} \quad \text{and} \quad N^* = N + 2,\\
&\textit{if } N = 2q, q \in \overset{0}{N};
\end{aligned}$$

Theorem 1. Operator $O_{M,N}(\{\varphi_{k,s}\}, \{\Pi_k\}, x) = \sum_{k=1}^{M} O_{k,N} f(x) H_{k,N}(x)$ has properties

$$\partial^s O_{M,N}(\{\varphi_{k,s}\}, \{\Pi_k\}, x)/\partial v_k^s(x)|_{\Pi_k} = \varphi_{k,s}(x)|_{\Pi_k},\ k = \overline{1,M},\ s = \overline{0,N}.$$

Remark 1. If $\Pi_k : \omega_k(x) = 0, k = \overline{1, M}$ are arbitrary set lines or surfaces in $R^n, n \geq 2$ and $\partial^p \omega_k(x)/\partial v_k^p|_{\Pi_k} = \delta_{0,p,} p = \overline{0, N}$ then statements of the theorem 1 are valid.

1.2. *Polynomial, trigonometrical and spline—interlineations on a set mutually perpendicular straight lines.* Let $G = I^2$, $I = [0, 1]$, $0 = s_{k,0} < \cdots < x_k$, $M_k = 1, k = 1, 2; \partial^{sk} f(x)/\partial x_k^{sk}|_{xk=xk,ik} = \varphi_{k,ik,sk}(x_3 - k)$,

$$B_k f(x) = \sum_{i_k=0}^{M_k} \sum_{s_k=0}^{N} \varphi_{i_k,S_k}(x_{3-k}) h_{M_k},\ N_{,s_k}(x_k),\ h_{M_k,i_k,s_k}^{(q)}(x_{k,j}) = \delta_{q,i_k}\delta_{i_k,j},$$

$q, s_k = \overline{0, N}; i_k, j = \overline{1, M_k}\ h_{M_k,N_{,sk}}(x_k)$ are basic polynomial, trigonometrical or spline interpolation system.

Theorem 2. Operator $Of\ (x) = (B_1 + B_2 - B_1 B_2) f(x)$ has next properties

$$\partial^p\, Of\,(x)/\partial x_k^p = \partial^p f(x)/\partial x_k^p = \varphi_{k,i_k,s_k}(x_{3-k}),\, x_k = x_{k,i_k},$$
$$p = \overline{0, N},\, i_k = \overline{0, M_k},\, k = 1, 2,$$
$$r_{12} f(x) := (I - O) f(x) = (1 - B_1)(1 - B_2) f(x),$$

i.e., $R_{12} f(x) = O(\varepsilon^2)$ if $(I - B_k) f(x) = O(\varepsilon)$, $\varepsilon \to 0, k = 1, 2$. The polynomial case on arbitrary set straight lines (for $n = 2$) and planes (for $n = 3$) [1].

1.3. *Economical polynomial, trigonometrical and spline—interpolation operators with help corresponding interlineation operators*. In general these operators has the form $\overline{O} f(x) = (\overline{B_1} + \overline{B_2} - B_1 B_2) f(x)$. Operators $\overline{B_k} f(x)$ are receiving from $B_k f(x)$ by make substitution $\varphi_{i_k,s_k}(x_{3-k}) \approx \Phi_{i_k,s_k}(x_{3-k})$ in $B_k f(x)$, where $\Phi_{i_k,s_k}(X_{3-k})$ are polynomial, trigonometrical or spline – interpolation operators with properties $\|\varphi_{i_k,s_k}(x_{3-k}) - \Phi_{i_k,s_k}(x_{3-k})\| = o(\varepsilon^2)$.

Theorem 3. The interpolating operator $\overline{O} f(x)$ uses smaller quantity the values of function $f(x)$, than the operator $B_1 B_2 f(x)$ (if they both approximate $f(x)$ with error $O(\varepsilon^2)$).

2. **Interlineation and interflation operators with preserving of the class $C^r(R^n)$, $r \geq 1$.** If $N \geq 1$ then we recommend to using the operators, which are preserving the class $C^r(R^n)$, $r \geq 1$. These operators are using the solutions of the some differential equations with partial derivatives [1].

3. **3D interpolation operators with help 3D interflation operators**. Let

$$f(x) \in C^{r,r,r}(I^3), r = 1, 2,\ u_{k,i_k}(x) = f(x)|_{x_k = I_k/M},$$
$$0 \leq i_k \leq M,\ k = \overline{1, 3},$$
$$L_{k,M} f(x) = \sum_{i_k=0}^{M} u_{k,i_k}(x) h(M_{x_k - i_k}),\ h(t) = (|t-1| - 2|t| + |t+1|)/2$$

Theorem 4. Operator $Of\,(x) = (L_{1,M} + L_{2,M} + L_{3,M} - L_{1,M} L_{2,M} - L_{1M} L_{3,M} - L_{2,M} L_{3,M} + L_{1,M} L_{2,M} L_{3,M}) f(x)$ has properties $Of\ (x)|_{x_k = j_k/M} = f(x)|_{xk = jk/m}$, $j_k = \overline{0, M}$, $k = 1, 3$, $\|f - Of\| = O(M^{-3r}) \forall u \in C^{r,r,r}(I^3)$, $r = 1, 2$

Theorem 5. Let we make substitution

$$u_{1,i_1}(x) = f(i_1/M,\ x_2,\ x_3) \approx \bar{u}_{1,i_1}(x)$$
$$= \sum_{j_2=0}^{M^{3/2}} \sum_{j_3=0}^{M^3} f\left(i_1/M,\ j_2/M^{3/2},\ j_3/M^3\right) h\left(M^{3/2} x_2 - j_2\right) h\left(M^3 x_3 - j_3\right)$$

$$+\sum_{j_2=0}^{M^3}\sum_{j_3=0}^{M^{3/2}} f\left(i_1/M,\ j_2/M^3,\ j_3/M^{3/2}\right) h\left(M^3x_2 - j_2\right) h\left(M^{3/2}x_3 - j_3\right)$$

$$-\sum_{j_2=0}^{M^{3/2}}\sum_{j_3=0}^{M^{3/2}} f\left(i_1/M,\ j_2/M^{3/2},\ j_3/M^{3/2}\right) h\left(M^{3/2}x_2 - j_2\right) h\left(M^{3/2}x_3 - j_3\right)$$

Analogously substitutions make also for another functions two and one variables in *Of*. Then we receiving operator $\overline{O}f(x)$ with properties: 1)$\|f - \overline{O}f\| = O(M^{-3r})$; 2)$\overline{O}f(x)$ used $Q = 6(M+1)(M^{3/2}+1)(M^3+1) = O(M^{5.5})$ values function f. Note, that classical three—linear interpolation operators $L_{1,M^3}L_{2,M^3}L_{3,M^3}f(x)$ has the same error and used $Q_{classic} = (M^3+1)^3 = O(M^9)$ values function f. The statements of the theorem 5 is valid also for splines more high order.

4. **Economical approximation operators with help blending function approximation operators**. Let $f(x) \in C^r(G),\ x = (x_1, x_2),\ B_{k,N}f(x), k = 1, 2; N \geq 2$ are one dimensional approximational operators, which worked to variable x_k and other variable is as parameter. Then $B_N f(x) = (B_{1,N} + B_{2,N} - B_{1,N}B_{2,N})f(x)$ are blending approximation operators with property: $\|f - B_{k,N}f\| = O(\varepsilon),\ \varepsilon \to 0,\ k = 1,\ 2 \Rightarrow \|f - B_N f\| = O(\varepsilon^2)$. The economical approximation operators has the form $(\overline{B}_N f(x) = (\overline{B}_{1,N} + \overline{B}_{2,N} - B_{1,n}B_{2,N})f(x)$. Operators $\overline{B}_{k,N}f(x)$ are receiving from $B_{k,N}f(x)$ by make substitution $\varphi_{i_k,s_k}(x_{3-k} \approx \Phi_{i_k,s_k}(x_{3-k})$ where $\Phi_{i_k,s_k}(x_{3-k})$ are Fourier's, Fejer's, Bernstein's, Haar's and so on approximation operators with properties $\|\varphi_{i_k,s_k})(x_{3-k}) - \Phi_{i_k,s_k}(x_{3-k})\| = o(\varepsilon^2)$.

   Theorem 6. The approximating operators $\overline{B}_N f(x)$ uses smaller quantity of values of Fourier's, Fejer's, Haar's coefficients, than the classical 2D operators $B_1B_2f(x)$ (if they both approximating $f(x)$ with the same error $O(\varepsilon^2)$).

   This results can be extended for case $n \geq 3$.

4. CONCLUSIONS

Author think approximation theory in 21st century will be joint with interflation of the functions many variables. The area of its applications will be expanded.

REFERENCES

1. O.N. Lytvyn (2002), Interlineation of the Functions and Some of Its Applications, (in ukrainian). Kharkiv, Osnova, 544 p.

INDEPENDENT COMPONENT ANALYSIS OF DYNAMIC CONTRAST-ENHANCED IMAGES: THE NUMBER OF COMPONENTS

X.Y. Wu and G.R. Liu
Centre for Advanced Computations in Engineering Science, Department of Mechanical Engineering, National University of Singapore, 9 Engineering Drive 1, Singapore 117576, Singapore

Abstract Independent component analysis (ICA) has attracted much interest in dynamic contrast-enhanced (DCE) neuroimaging, as it allows for blind separation of the brain haemodynamic patterns (components). However, the exact number of components is always unknown in practice. In this work, numerical simulation study was carried out to compare the performance of ICA by using the principle component analysis (PCA) to reduce the number of components. Oscillatory indices method and kurtosis method are also discussed for automatically selecting the components of interest.

Keywords: DCE image processing, ICA, PCA, component of interest.

1. INTRODUCTION

Dynamic contrast-enhanced imaging involves the intravenous injection of a contrast agent (tracer) and sequential imaging to simultaneously monitor the changes of tracer concentration over time. When the tracer reaches the tissues of scanning layer, the captured image will show enhancement due to the presence of tracer. The enhanced signals on different brain tissues represent distinct haemodynamic patterns that caused by the difference of local microvasculature and functions of the brain. Therefore, the measured signals may be looked as a summation of the brain haemodynamic behaviour and some artificial processes, like subtle head movements, physiological pulsations and machine noise. By assuming the spatial independence of those haemodynamic patterns and artificial

G. R. Liu et al. (eds.), Computational Methods, 1111–1116.

processes, independent component analysis is able to separate those mixed signals without any priori knowledge.

Application of ICA on functional neuroimaging (fMRI) was firstly carried out by McKeown *et al* [1] to extract the task-related brain information. While in DEC neuroimaging, as the areas affected by abnormal brain function like tumour should be unrelated to those affected by other artificial factors, it is able to reveal the brain haemodynamic abnormalities by using the spatial ICA technique.

2. METHOD

Each frame of the series of DCE image is first converted into a 1-D signal vector $\mathbf{X}_i$ $(i = 1, \ldots, m)$, here m is the total number of frames. The length of the signal vector p equals the number of pixels per frame. The signal $\mathbf{X}_i$ is considered as a linear combination of the independent components $\mathbf{C}_i$. Assume the number of components is

n, we have :

$$\mathbf{X}_i = \mathbf{M}_1\mathbf{C}_1 + \mathbf{M}_2\mathbf{C}_2 + \ldots + \mathbf{M}_n\mathbf{C}_n \tag{1}$$

The entire image data can then be expressed as:

$$\mathbf{X} = \mathbf{M} \cdot \mathbf{C} \tag{2}$$

where $\mathbf{X}$ is the $m \times p$ matrix of image pixels (data), $\mathbf{M}$ is the $m \times n$ mixing (linear combination) matrix and $\mathbf{C}$ is the n x p component matrix. Each row of the component matrix, $\mathbf{C}_i$, represents one independent component. Both the mixing matrix, $\mathbf{M}$, and the component matrix, $\mathbf{C}$, can be estimated by iterative attempts to minimize redundancy between components (Infomax) and updating the element of weight matrix, $\mathbf{W}$, where $\mathbf{W} = \mathbf{M}^{-1}$. The vector $\mathbf{C}_i$ is then reformed into 2-D to obtain the component map. Those maps are fixed over time, while the relative contribution of each map changes with a unique associated time course (column of mixing matrix, $\mathbf{M}$).

In practice, the exact number of independent components is always unknown. Different settings of the components number can get different ICA results. McKeown and Sejnowski [2] suggested that the component number can be reduced by first preprocessing the data with the PCA technique. However, the cutoff point for the eigenvalues may not be obvious and should be selected carefully. In this work, numerical simulation study was carried out to compare the performance of ICA by using PCA to reduce the number of components. Oscillatory indices method and kurtosis method were also discussed for the selection of components of interest.

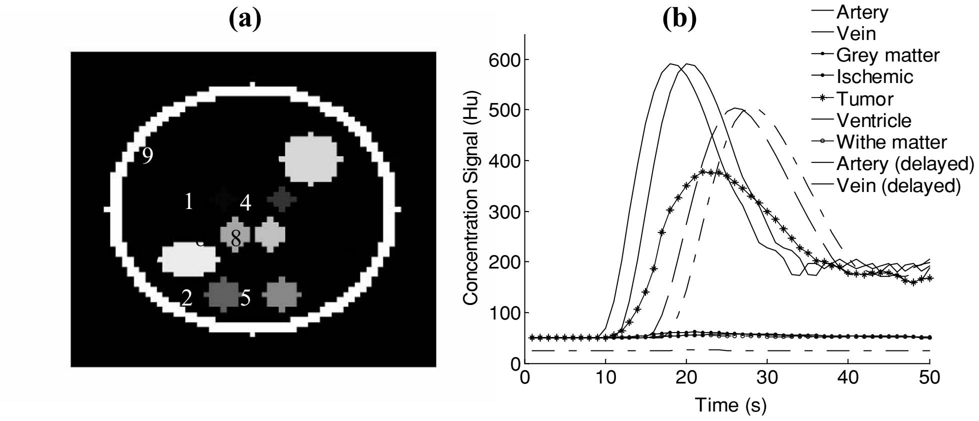


Figure 1. (a) Synthetic brain mask. 1-artery, 2-vein, 3-tumour, 4-artery (delayed), 5-vein (delayed), 6-grey matter, 7-ischaemic area, 8-ventricles, 9-white matter. (b) Simulated concentration signals used in synthetic dynamic images.

2.1 Numerical simulation study

The synthetic dynamic images consisted of 50 images of 63 × 63 pixels with a time interval Δt of 1 s. Totally 9 types of haemodynamic signals were applied on a synthetic brain mask to discriminate their locations, as shown in Figure 1a and b. Those signals were obtained from the DCE CT images of a patient with brain tumour and ischaemic stroke. Smoothing preprocess was applied on those signals before constructing the synthetic dynamic images. Gaussian noise was then added to the synthetic images to generate the signal-to-noise ratios (SNRs) that is typical of the modern imaging modalities. Similar to previous work [3], a preprocessing step of smoothing with a 3 × 3 median kernel was applied to those synthetic images. The Infomax ICA scheme [1] was then used to separate the synthetic images into spatial independent components. The number of components was set to 50 (full ICA) and 5, 10 and 20 with PCA preprocessing.

3. RESULTS AND DISCUSSIONS

In most of our simulation experiments, the tumour, artery, vein and ventricle related component maps were observed. However, the grey matter, white matter and ischaemic area were not clearly seen even when SNR increases to 50. The tumour related component maps show existence of artery and vein partly (Figure 2a). The vein signal and its 2s delayed signal show combination on the

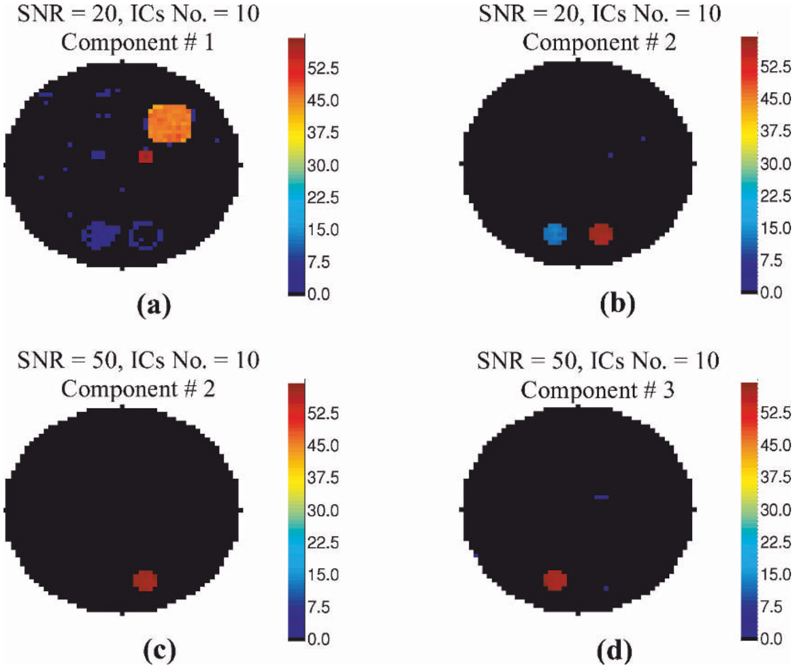


Figure 2. Typical independent component maps at different noise level. (a) Tumour related map at SNR 20; (b) Vein related map at SNR 20; (c) 2 s delayed vein-related map at SNR 50; (d) Vein related-map at SNR 50.

corresponding component map (Figure 2b). However, when SNR increases to 50, those signals were completely separated (Figure 2c,d). The normalized tumour related, artery related and vein related time courses show consistency with their respective original signal even at a high noise level (Figure 3). The interpretable component maps and their corresponding time courses did not change abruptly with the decrease of components amount. Even when the number of components was reduced to 5, the strong signal related components were still recognizable, although more severe combination was observed. In fact, the PCA preprocessing step remains the large variance of the data, so the strong signal related components could be essentially unchanged with the reduction of components amount. The PCA preprocessing step also diminishes the noise effect, which hence improves the ICA result. Firstly, in full ICA, the ventricle signal was split into two components at low noise level, and was missed at high noise level. While with PCA preprocessing, such situation was successfully avoided. Secondly, using PCA to reduce the data dimension saves the computation load. In real brain disease diagnosis, this can be a very meaningful factor. Table 1 shows the average computation time of 12 simulation runs with different components number on a Pentium IV personal computer. The small amplitude

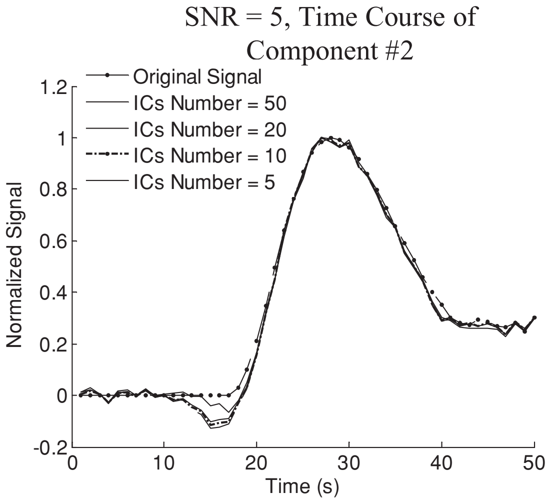


Figure 3. Time course of delayed vein- related component at high noise level

and similarity between each other make the weak signals inseparable in our simulation. Part of the weak signal information that corresponding to the small eigenvalues may be discarded during the PCA preprocessing step. However, in full ICA, those weak components were still inseparable as they were corrupted by the components split and noise effect. One way to solute this problem may be removing the strong components and run the ICA iteratively [4].

In real application, the selection of meaningful component maps is mainly done by human visualization. Such task can be very tedious, as a large majority of component maps are non-interesting especially when the number of components is fairly large. In our previous work [3], an oscillatory indices method was proposed to select the components of interest automatically. The basic idea is that the haemodynamic signals are always steadily, while the noise effect can be very unpredictable. Therefore, the time course of the physiological components should be relative smoother than that of the noise components. The statistical kurtosis of the components may also be useful, as the noise components always have a kurtosis near zero (Guassian distribution). In addition, some components of interest may not have smooth enough time courses when the components split is severe. However, the computation of kurtosis is much more complex than that of the oscillatory indices.

Table 1. Average computation time.

Components number	50	20	10	5
Computation time (s)	99.7	20.3	12.9	7.1

4. CONCLUSIONS

Independent component analysis with dimension reduction by PCA shows a satisfactory performance in our simulation. It not only reduces the noise effect but also saves the computation time.

REFERENCES

1. M.J. McKeown *et al* (1998), Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*, 6, pp. 160–188.
2. M.J. McKeown and T.J. Sejnowski (1998), Independent component analysis of fMRI data: examining the assumptions. *Human Brain Mapping*, 6, pp. 368–372.
3. X.Y. Wu (2004), Functional neuroimaging and analysis. ME Thesis, Nanyang Technological University, Singapore.
4. Y-H. Kao *et al* (2003), Haemodynamic segmentation of MR brain perfusion images using independent component analysis, thresholding, and Bayesian estimation. *Magnetic Resonance in Medicine*, 49, pp. 885–894.

THE GEOMETRIC CONSTRAINT SOLVING BASED ON HYBRID GENETIC ALGORITHM OF CONJUGATE GRADIENT

C.H. Cao, W.H. Li and Biao. Cong
Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, College of Computer Science and Technology, Jilin University, Changchun 130012, P.R. China

Abstract We transform the geometric constraint solving into the numerical optimization solving. A new hybrid algorithm is proposed which combines the merits of global search of the genetic algorithm (GA) and the good property of local search of the conjugate gradient approach. This algorithm uses GA to search the area where the best solution may exist in the whole space, and then performs fine searching. When the algorithm approaches to the best solution and the search speed is too slow, we can change to the effective local search strategy—the conjugate gradient algorithm in order to enhance the ability of the GA on fine searching. It makes the algorithm get rid off the prematurity convergence situation. We apply this algorithm into the geometric constraint solving. The experiment shows that the hybrid algorithm has the effective convergence property and it can find the global best solution.

Keywords: geometric constraint solving, conjugate gradient, genetic algorithm, hybrid algorithm, convergence.

1. INTRODUCTION

Many scholars worked over the constraint solving by numeric computing approach, artificial intelligence approach, degree of freedom approach and graph-based approach. To conclude there are whole solving, sparse matrix solving, joint analysis solving, constraint diffuse solving, symbolic algebra solving and guides solving [1].

In recent years, the genetic algorithm (GA) attracts many researchers' attention in many areas. But in these areas the algorithm may get into the local best

G. R. Liu et al. (eds.), Computational Methods, 1117–1121.

solution because of its poor local searching capability. Conjugate gradient algorithm is a common algorithm used in non-constraint optimization problems. It was also indicated that the more dimensions of the function, the longer time spending on calculating differential and the time spending on every iteration is also much longer.

Integrated the advantages and disadvantages of these algorithms above all, the hybrid GA based on conjugate gradient algorithm is proposed [2]. This algorithm uses GA to search the area where the best solution may exist in the whole space, and then performs fine searching. When the algorithm approaches to the best solution and the search speed is too slow, we can change to the effective local search strategy—the conjugate gradient algorithm in order to enhance the ability of the GA on fine searching. In reference [3], a hybrid algorithm based on dropped method is introduced. It makes the algorithm get rid off the prematurity convergence situation. Because of the operations of coding and decoding all the iterations, the convergence speed of this algorithm becomes very slow. As a try, a new hybrid optimization algorithm is introduced that integrates the conjugate gradient algorithm and the improved GA in this paper. Compared to GA, this new algorithm adds a conjugate gradient, which can improve the ability of the local fine adjustment and avoid getting into early convergence situation. In the other hand, this method adds GA compared with the conjugate gradient algorithm, which has strong global search ability and provides a good original point that improves the calculating efficiency.

2. GEOMETRIC CONSTRAINT SOLVING

Geometric constraint can be defined as the inherent relationships among the geometric elements in a geometric constraint system. It can reflect the shape and location of the geometric body directly. The geometric constraint includes two kinds of geometric constraint, structure constraint and dimension constraint. The former is the geometric topological relationships, which describes the relative location in the space and connection manner. The values of parameter attribute do not change in the process of parameterization. The dimension constraint describes the constraint by the label of the dimension on the graph, such as distance, angle and so on. The label of the dimension in engineering drawing is a direct and natural describer of the geometry. It gives a best modification manner of geometry. The aim of dimension driving is to modify the drawing by the changing label of dimension but maintains the relationships of topological not changed in the whole process. The constraint problem can be formalized as (E, C) [4]' here $E = (e_1, e_2, \ldots, e_n)$, it can express geometric elements, such as point, line, circle, etc; $C = (c_1, c_2, \ldots, c_m)$, c_i is the constraint set in these geometric elements. Usually, one constraint is represented by an algebraic equation, so the constraint can be expressed as follows. $X = (x_0, x_1, \ldots, x_n)$, X_i are some parameters. Constraint solving is to get a solution x to satisfy formula (1).

$$\begin{cases} f_1(x_0, x_1, x_2, \ldots, x_n) = 0 \\ \qquad\vdots \\ f_m(x_0, x_1, x_2, \ldots, x_n) = 0 \end{cases} \tag{1}$$

$$F(X_j) = \sum_{1}^{m} |f_i| \tag{2}$$

Apparently if X_j can satisfy $F(X_j) = 0$ then X_j can satisfy formula (1). Then the constraint problem can be transformed to an optimized problem and we only need to solve min $(F(X_j)) < \varepsilon$, ε is a threshold. In order to improve the speed of the algorithm, we adopt the absolute value of f_i not the square sum to express constraint equation set.

3. HYBRID ALGORITHM

3.1 Improved genetic algorithm

Improved GA can be expressed as follows:

Step 1. Initialization: produce an original group at random.

Step 2. Calculate each individual's adaptation value in according to adaptation function.

Step 3. Choose father generation according to the proportion information of each chromosome. In order to quicken the convergence speed, in every generation the solution that has a maximal adaptation value will be transferred to the next generation compulsively, and not influenced by the choice process.

Step 4. Carry out the crossover operation to two individuals selected at random by the method which associates one-point crossover and two-point crossover, fixed p_c and self-adaptive p_c.

Step 5. Carry out the variation operation to every individual in every generation by the method which associates general variation and big variation, fixed p_m and self-adaptive p_m.

Step 6. If the stop rule can be satisfied, we can acquire the global best solution and the algorithm is convergent, otherwise turn to step 2.

3.2 Conjugate gradient algorithm

A common conjugate gradient algorithm can be written as follows:

Step 1. Choose the initial value $x^{(0)}$ in s and the precision $\varepsilon > 0$;

Step 2. Compute $g_0 = \nabla f(x^{(0)})$, $s^{(0)} = -g_0$, $k = 0$;

Step 3. $x^{(k+1)} = x^{(k)} + \lambda_k s^{(k)}$, $g_{k+1} = \nabla f(x^{(k+1)})$, in the formula $\lambda_k = \arg\min f(x^{(k)} + \lambda_k s^{(k)})$;

Step 4. If, $\|g_{k+1}\| \leq \varepsilon$, the iteration will stop, otherwise turn to step 5;

Step 5. If $k < n-1$, compute $\mu_{k+1} = \|g_{k+1}\|^2/\|g_k\|^2$, $s^{(k+1)} = -g_{k+1} + \mu_{k+1}s^{(k)}$, assume $k = k+1$, turn to step 3, if $k = n-1$, assume $x^{(0)} = x^{(n)}$, turn to step 2.

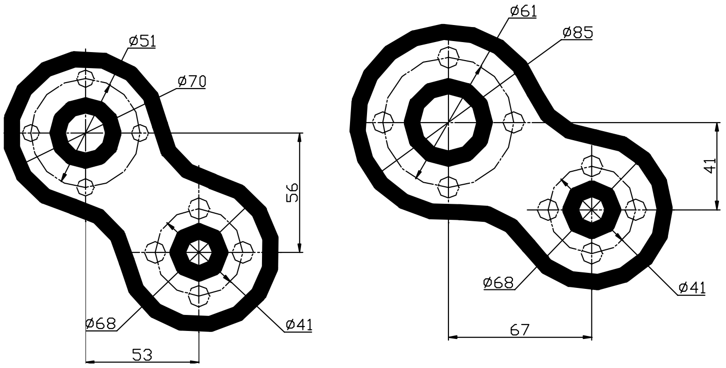
3.3 Hybrid algorithm

Step 1. Initialize the colony;

Step 2. Carry out iteration by the GA, search the best individual. If after some generations the best individual has not appeared, then turn to step 3.

Step 3. If after some steps search by Conjugate Gradient Algorithm there is no better points than, then stop; otherwise assuming to find a point y whose function value is better than, then produce a colony including y, return step 2.

We can realize from Figure 1. that once a user defines a series of relations, the system will satisfy the constraints by selecting proper state after the parameters are modified.



(a) A design instance (b) solving

Figure 1. The result of the experiment result.

4. RESULT

From the study in the chapter, we can come to a conclusion: conjugate gradient algorithm can help the GA avoid the earlier convergence condition. Compared to GA, this new algorithm adds a conjugate gradient, which can improve the ability of the local fine adjustment and avoid getting into early convergence situation. On the other hand, this method adds GA compared with the conjugate gradient algorithm which has strong global search ability and provides a good original point that improves the calculating efficiency. This algorithm can achieve the global optimization solution finally.

REFERENCES

1. Yuan (1999), Research and Implementation of Geometric Constraint Solving Technology, doctor dissertation of Tsinghua University, pp. 1–8
2. M. Pan and G. He (2000), A hybrid genetic of function opitimization based on conjugate gradient algorithm, *Journal of Shandong University of Science*, 19, 4, pp. 10–13.
3. M. Zhao (1997), The hybrid numerical algorithm of the function optimization based on the genetic algorithm and fast decedent algorithm. *System Engineering and Implementation*, 7, 11, pp. 59–64.
4. S. Liu, M. Tang and J. Dong (2003), Two spatial constraint solving algorithms. *Journal of Computer-Aided Design & Computer Graphics,* 15, pp. 1011–1029.

FAST IMAGE MOSAICS ALGORITHM USING PARTICLE SWARM OPTIMIZATION

Y. Zhang, W.H. Li, Y. Meng, Z.J. Tan and Y.J. Pang
Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, College of Computer Science and Technology, Jilin University, Changchun 130012, China

Abstract Image mosaic plays an important role in producing panoramic image. We proposed an automated seamless mosaics algorithm. In our proposal, firstly, we use particle swarm optimization (PSO) to find a certain area which contains sufficient objective characters, then we use pattern matching method to search the matching patch in another image and adjust image; at last, we make use of feathering blending to provide a smooth transition between overlapping areas and get automated seamless mosaics of images.

Keywords: image mosaics, seamless mosaics, particle swarm optimization, feathering.

1. INTRODUCTION

Panoramic image is a method to make use of realistic images to get a full view panoramic space [1–3]. Image mosaics algorithms can be divided into: registration algorithms based on correlation and registration algorithms based on characters [1, 4, 5]. Image automated seamless mosaics involve many fields, such as module identification, optimization and so on. Actually, automatically searching overlapping areas of images can be described as follows: there are two rectangular areas A and B, how to search the position of A_2 in image B, where B contains the area A_2 and area A and area A_2 are identical modules. In this chapter, we make use of Particle Swarm Optimization [6] (PSO) to find an area which contains sufficient objective characters in one image and find corresponding area in another image using module matching, and then adjust these images. At last, we make use of feathering blending to provide a smooth transition between overlapping areas and get automated seamless mosaics of images.

G. R. Liu et al. (eds.), Computational Methods, 1123–1128.

2. MATCHING MOSAICS

2.1 Optimized characteristic block extraction based on PSO

Two adjacent images can be overlapped by each other by 50% because tripod might not be completely horizontal and cameras might be sloping or facing upwards. Thus, the first job of image mosaics is to exactly allocate the overlapping areas of two images. And the second one is to adjust the lightness, since the light intensity of two images from different points might be of great difference.

In the first image, if we can confirm area A, then we can easily get area B using module matching methods in the other image, according to image overlapping theory, taking acceptable range of error into consideration. The more objective characters we are searching in area A, the much difference is required between this area and surrounding areas, and the better. Distance L_2 is simplest and most frequently used distance function to compare the degree of similitude of two areas. The value of L_2 is small when two areas are very similar to each other, and vice versa. For a certain area S, we can calculate four values of L_2 by comparing itself with its surrounding up, down, left and right four areas of the same size, denoted as f_1, f_2, f_3, f_4. The bigger of sum of f_1, f_2, f_3, f_4, the more difference between area S and its surrounding areas. We denote evaluation function of area S as:

$$F = f_1 + f_2 + f_3 + f_4 \tag{1}$$

The distance L_2 of according areas is:

$$f_i = sqrt\left\{\sum_{m=1}^{k}\sum_{n=1}^{l}[(R(p_S^{mn}) - R(p_{S_i}^{mn}))^2 + (G(p_S^{mn}) - G(p_{S_i}^{mn}))^2 + (B(p_S^{mn}) - B(p_{S_i}^{mn}))^2]\right\} \tag{2}$$

Where, S, S_i are target area and its surrounding areas, respectively; k, l are the width and height of the area. $R()$, $G()$, $B()$ are values of three primary colours of pixels. p_S^{mn} is the pixel with the coordinate value (m, n) in area S and $p_{S_i}^{mn}$ is the pixel with the coordinate value (m, n) in area S_i.

For any area S in the right half of the first image, the bigger of value F, the easier we can find an area with sufficient information needed for matching searching. How to find an area S containing sufficient objective characters is a better problem and we just need a satisfied result. Therefore, we can use particle swarm optimization (PSO) [6] to searching for area S.

PSO is an intelligent optimized methods based on iteration and was firstly proposed by Kennedy and Eberhart [6]. In PSO, particle swarm is searching in an n-dimension space for new solution by constantly adjusting its position X, where the position of each particle is a solution of the problem. Each particle can record the best solution it has found, denoted as P_{id}, and the best position of the whole particle swarm ever passed, i.e., the best solution at the present, denoted as P_{gd}.The speed of each particle is denoted as V:

$$V'_{id} = V_{id} + \eta_1 r\,and()(P_{id} - X_{id}) + \eta_2 r\,and()(P_{gd} - X_{id}) \tag{3}$$

Where, V_{id} is the spend of particle i on the dth dimension. η_1, η_2 are parameters to adjust the relative importance of P_{id} and P_{gd}, normally, $\eta_1 = \eta_2 = 2$. Thus, we can calculate the next position of the particle by:

$$X'_{id} = X_{id} + V_{id} \tag{4}$$

From (3) and (4), we can get that the moving direction of particles is determined by three fractions: its velocity V_{id}, the distance between current location and the best location $(P_{id} - X_{id})$, and the distance from the best location of the particle swarm $(P_{gd} - X_{id})$.

When mosaicking two images with overlapping areas, we randomly distributed 10 particles in the right half of the first image. The initial position is the coordinate of several pixels and we define an initial velocity of these particles. Each particle is moving in the solution space. We can find a matching area with certain characters, by adjusting the moving direction and velocity of particles using fitness function. Each particle can decide an area of 20*20, which is used to search for areas with multiobjective characters.

Figure 1(a) demonstrated how to find area S with sufficient characters in one image using PSO algorithm. The area in the green square is the area S we get.

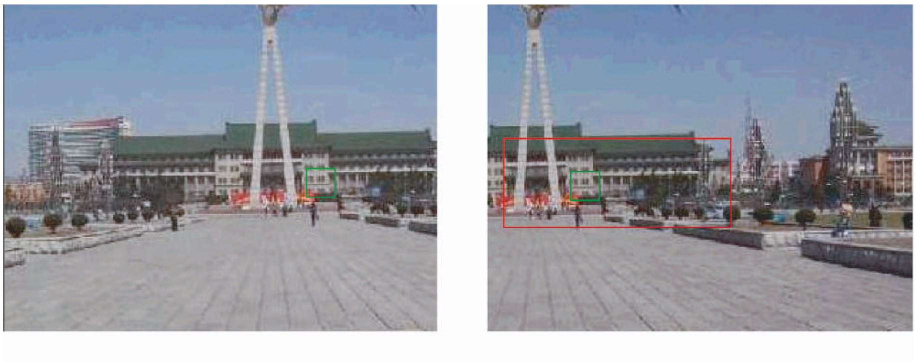

(a) (b)

Figure 1. (a) is characteristic patch we found using PSO algorithm and (b) is the matching patch based on this characteristic patch.

2.2 Image registration

Because the light intensity of the same scene in different photos might be different and the distributions of lightness are not identical, we might not find matching areas when we mosaicking images. So we have to adjust the images to ensure the general identity of the luminance histograms of these images. In this paper, we use the method proposed by Hertzmann [7] to process the first image. We can use the following formula to calculate the lightness of each pixel in the first image:

$$Y(p) = \frac{\bar{\sigma}_{in}}{\sigma_{out}}(Y(p) - \mu_{out}) + \bar{\mu}_{in}; \quad \bar{\sigma}_{in} = \frac{1}{n}\sum_{i=1}^{n}\sigma_i; \quad \bar{\mu}_{in} = \frac{1}{n}\sum_{i=1}^{n}\mu_i \tag{5}$$

Where $\sigma_{i,}\mu_i$ and $\sigma_{out,}\mu_{out}$ are the standard deviations of the luminance and mean luminance in the first image and the second image, respectively.

In the first image, if area S has been determined, we apply range of errors in the second image and then obtain a certain searching space. In this search space, we compare area S with another area of the same size, and the area with the smallest value is the area we are searching for, if we use L_2 distance formula in formula (2) as our evaluation function. Figure 1 demonstrates how to find matching areas, in which, (a) is characteristic patch we found using PSO algorithm at original images with overlapping areas, and (b) is the matching patch based on this characteristic patch at original images with overlapping areas. The area in red square is the matching space we used for searching and the area in green square is the matching patch we get.

3. SMOOTHING PROCESSING

After image allocation, if we simply mosaic two images together, there might be a seam at the splicing tape. To avoid this, we can fix the colours using feathering [1] methods in order to get a smooth combined image with high quality. We can do feathering process in the correcting areas of two images, which is within 10 pixels away from the central matching area. We demote the correcting area of the first image as $R1$, $G1$, $B1$ and that of the second one as $R2$, $G2$, $B2$ and the correcting area after feathering as R, G, B. We use the following formula to feather the pixels in the correcting area:

$$\begin{gathered} R = R1 ^*(1 - p) + R2 ^*p; G = G1 ^*(1 - p) + G2 ^*p; \\ B = B1 ^*(1 - p) + B2^* p \end{gathered} \tag{6}$$

Figure 2. Seamless image mosaics after feathering.

Where, p is the tapering factor, which is related to the overlapping distance to the horizontal direction between images, changing from 1 to 0 in the range of (0, 1). We can achieve seamless image mosaics after feathering, as shown in Figure 2.

4. CONCLUSION

In this chapter, we proposed a new fast image mosaics algorithm based on PSO to achieve full view panoramic image. It combined PSO can speed up image searching and find the areas containing multiobjective characters. We make use of feathering blending to provide a smooth transition between overlapping areas and get automated seamless mosaics of images. The advantage of our algorithms is that ghost phenomenon might occur when there are moving objects in the mosaicking areas, that is, the algorithm is not suitable for the circumstance when characteristic areas are at the boundary of the moving objects.

REFERENCES

1. R. Szliski and H.Y. Shum (1997), Creating full view panoramic image mosaics and environment maps. *Computer Graphics Proceedings of Annual Conference Series, ACM SIGGRAPH*, Los Angeles, CA, pp. 251–258.
2. S.E. Chen (1995), QuickTime VR-An image-based approach to virtual environment navigation. *Computer Graphics Proceedings of Annual Conference Series, ACM SIGGRAPH*, Los Angeles, pp. 29–38.
3. W. Du and H. Li (2002), A novel panoramic representation for dynamic scenes. *Chinese Journal of Computers* (in Chinese), 25, 9, pp. 968–975.
4. L.G. Brown (1992), A survey of image registration technique. *ACM Computing Surveys,* 24, 4, pp. 325–376.

5. D. Cepl and A. Zisseman (1998), Automated mosaicing with super-resolution zoom. *Proceedings of CVPR'98 Santa Barnara*. California, pp. 885–891.
6. J. Kennedy and R.C. Eberhart (1995), Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Networks,* Vol. IV, IEEE, pp. 1942–1948.
7. A. Herzmann, C.E. Jacobs, N. Oliver, B. Curless, and D.H. Salesin (2001), Image analogies. *Computer Graphics Proceedings of Annual Conference Series, ACM SIGGRAPH,* Los Angeles, CA, pp. 327–339.

CHECKING CONSISTENCY IN HYBRID QUALITATIVE SPATIAL REASONING

Haibin Sun[1], Wenhui Li[1] and Yong-jian Zhang[2]
[1]*Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, School of Computer Science and Technology, Jilin University, Changchun 130012, China*
[2]*Department of Information and Electronical Engineering, Shandong Institute of Architecture and Engineering, Jinan 250014, China*

Abstract In this article, we investigate the problem of checking consistency in a hybrid formalism which combines two essential formalisms in qualitative spatial reasoning: topological formalism and cardinal direction formalism. First the general interaction rules are given, and then, based on these rules, an improved constraint propagation algorithm is introduced to enforce the path consistency. The results of computational complexity of checking consistency for CSPs based on various subsets of this hybrid formalism are presented at the end of this article.

Keywords: computational complexity, constraint propagation algorithm, qualitative spatial reasoning, binary constraint satisfaction problem, consistency checking.

1. INTRODUCTION

Combining and integrating different kinds of knowledge is an emerging and challenging issue in Qualitative Spatial Reasoning (QSR). The work by Gerevini and Renz [1] has dealt with the combination of topological knowledge and metric size knowledge in QSR, and the work by Isli et al. [2] has combined the cardinal direction knowledge and the relative orientation knowledge. In this paper, we are devoted to investigating the computational problems in the formalism combining topological and cardinal directional relations. We first introduce the two formalisms of topological and cardinal directional relations, respectively.

G. R. Liu et al. (eds.), Computational Methods, 1129–1133.

1.1 Topology formalism

Topology is perhaps the most fundamental aspect of space. Topological relationships are invariant under topological transformations, such as translation, scaling and rotation. Examples are terms like *neighbour* and *disjoint* [3]. RCC8 is a formalism dealing with a set of eight jointly exhaustive and pairwise disjoint (JEPD) relations, called base relations, denoted as DC, EC, PO, EQ, TPP, NTPP, TPP^{-1}, NTPP^{-1}, with the meaning of DisConnected, Extensionally Connected, Partial Overlap, EQual, Tangential Proper Part, Non-Tangential Proper Part and their converses. Exactly one of these relations holds between any two spatial regions. In this chapter, we will focus on RCC8 formalism.

1.2 Cardinal direction formalism

The work by Goyal and Egenhofer [4] introduced a direction–relation model for extended spatial objects that considers the influence of the objects' shapes. It uses the projection-based direction partitions and an extrinsic reference system, and considers the exact representation of the target object with respect to the reference frame. The reference frame with a polygon as reference object has nine direction tiles: north (N_A), northeast (NE_A), east (E_A), southeast (SE_A), south (S_A), southwest (SW_A), west (W_A), northwest (NW_A) and same (B_A). The cardinal direction from the reference object to a target is described by recording those tiles into which at least one part of the target object falls. We call the relations where the target object occupies one tile of the reference object *single-tile* relations, and others *multi-tile* relations. We denoted this formalism by CDF for brevity.

2. INTERACTION RULES BETWEEN RCC8 AND CDF

The internal operations, including converse and composition, on RCC8 can be found in Ref. [5]. The internal operations on CDF have been investigated in Refs. [6, 7]. In order to integrate these two formalisms, we must investigate interaction rules between them.

To facilitate the representation of the interaction rules, we denote a basic cardinal direction relation by a set *SB* which includes at most nine elements, i.e., the nine single-tile cardinal direction relations. For example, a relation *B:S:SE:SN* can be denoted by $\{B, S, SE, SN\}$. The general cardinal direction relation can be regarded as a set *GB*, whose element is the set *SB*. So we have

Table 1. The interaction table from the atomic relations of RCC8 to CDC relations.

Atomic RCC8 relation	DC	EC, PO, TPP^{-1} or $NTPP^{-1}$	TPP, NTPP or EQ
Induced cardinal direction relation	U	$\forall SB \in GB : B \in SB$	{{B}}

Table 2. The interaction table from the basic relations of CDF to RCC8 relations.

Basic cardinal direction relation (SB)	{B}	$\exists R \in BIN : R \in SB \cap B \in SB$ or BIN	$B \notin SB$
RCC8 relation	DC ∨ EC ∨ PO ∨ TPP ∨ NTPP ∨ EQ ∨ TPP^{-1}	DC ∨ EC ∨ PO ∨ TPP^{-1} ∨ $NTPP^{-1}$	DC

the relation: $SB \in GB$. The universal relation is the set *BIN* = {*B, N, NE, E, SE, S, SW, W, NW*}, and the universe, i.e., the set of all possible cardinal relations, is denoted by U. The interaction rules are presented in Tables 1 and 2.

3. TERMINOLOGIES AND DEFINITIONS

Definition 1. Binary Constraint Satisfaction Problem (BCSP). If every one of the constraints in a Constraint Satisfaction Problem (CSP) involves two variables (possibly the same) and asserts that the pair of values assigned to those variables must lie in a certain binary relation, then the constraint satisfaction problem is called *Binary Constraint Satisfaction Problem*.

Definition 2. We define an *RCC8-BCSP* as a *BCSP* of which the constraints are RCC8 relations on pairs of the variables. The universe of a *RCC8-BCSP* is the set $\Re^2$ of regions. Similarly we can define *CDF-BCSP* as a *BCSP* of which the constraints are CDF relations on pairs of the variables and the universe is the set $\Re^2$ of regions, and *RDF-BCSP* as a *BCSP* of which the constraints consist of a conjunction of RCC8 relations and CDF relations on pairs of the variables and the universe is the set $\Re^2$ of regions.

Definition 3. An n-by-n constraint matrix M is path-consistent if $M \leq M^2$. M is path-consistent just in case $M_{ij} \subseteq M_{ik} \circ M_{kj}$. We must note that path consistency is the necessary, but not sufficient, condition for the consistency of a BCSP. The constraint propagation algorithm in Ref. [8] can be adapted to achieving path consistency by enforcing the above rules, and the computational complexity is obviously $O(n^3)$.

4. COMPLEXITY OF CONSISTENCY CHECKING IN RDF–BCSP

We use T to denote the set of general RCC8 relations, T_b the set of basic RCC8 relations including universal relation, C the set of general CDF relations and C_b the set of basic CDF relations in *RDF–BCSP*.

The work by Renz and Nebel [9] and Renz [10] identified three maximal tractable subsets, i.e., C_8, Q_8 and $\hat{H}_8$, of the relations in RCC8 that contains all base relations and showed that path-consistency is sufficient for deciding consistency for these subsets. The work by Skiadopoulos and Koubarakis [11] have presented the first algorithm for checking the consistency of a set of cardinal direction constraints and proved that the consistency checking of a set of basic cardinal direction constraints can be performed in $O(n^5)$ time while the consistency checking of an unrestricted set of cardinal direction constraints is NP-complete.

Theorem 1. The complexity of checking the consistency of RDF–BCSP based on set $S = T_b \cup C_b$ is O(n^5).

Proof. Checking the consistency of *RCC8-BCSP* based on T_b is polynomial [9], and checking the consistency of *CDF–BCSP* based on C_b has been shown to be $O(n^5)$ [11]. From Table 2, all possible CDF basic relations can only entail $DC \vee EC \vee PO \vee TPP \vee NTPP \vee EQ \vee TPP^{-1}$, $DC \vee EC \vee PO \vee TPP^{-1} \vee NTPP^{-1}$ or DC relations, which belong to the maximal tractable subset $\hat{H}_8$(see Appendix B of [9]) of RCC8. So checking the consistency of *RDF–BCSP* based on the union of T_b and C_b is polynomial. First we can run the improved constraint propagation algorithm to enforce the path consistency, and then the algorithm in Ref. [11] will be employed to check the consistency for *CDF–BCSP* component. Obviously the complexity is $O(n^5)$.

Because checking the consistency of *RCC8–BCSP* based on the set T is ***NP**-Complete* (see Theorem 22 of [9]) and checking the consistency of *CDF–BCSP* based on the set C is ***NP**-Complete* (see Theorem 3 of [11]), we have the following theorem: QED

Theorem 2. Checking the consistency of RDF–BCSP based on the set $S = T \cup C_b$ *or* $T_b \cup C$ *or* $T \cup C$ *is* ***NP**-Complete.*

5. CONCLUSIONS

In this paper, we have combined two essential formalisms in qualitative spatial reasoning, i.e., RCC8 and cardinal direction formalism. The interaction rules have been given and they can be embedded into the propagation algorithm to enforce the consistency of BCSP based on the new hybrid formalism, and then the results for the complexity of checking consistency based on various subsets of this new formalism are given. The complexities for other combinations of formalisms in QSR should be investigated in the future, and the computational problems in Fuzzy QSR should be also interesting.

REFERENCES

1. A. Gerevini and J. Renz (2002), Combining topological and size constraints for spatial reasoning. *Artificial Intelligence (AIJ)*, 137, 1–2, pp. 1–42.
2. A. Isli, V. Haarslev and R. Moller (2001), Combining cardinal direction relations and relative orientation relations in Qualitative Spatial Reasoning. *Technical Report FBI-HH-M-304/01*, Fachbereich Informatik, University Hamburg.
3. M. Egenhofer (1989), A formal definition of binary topological relationships. In: *Third International Conference on Foundations of Data Organization and Algorithms (FODO)*, Paris, France.
4. R. Goyal and M. Egenhofer (2000), Cardinal directions between extended spatial objects. *IEEE Transactions on Knowledge and Data Engineering*. Available at *http://www.spatial.maine.edu/~max/RJ36.html*
5. D.A. Randell, A.G. Cohn and Z. Cui (1992), Computing transitivity tables: a challenge for automated theorem provers. In: *Proceedings CADE 11*, Springer Verlag, Berlin.
6. S. Skiadopoulos and M. Koubarakis (2004), Composing cardinal direction relations. *Artificial Intelligence*, 152, 2, pp. 143–171.
7. S. Cicerone and P. Di Felice (2004), Cardinal directions between spatial objects: the pairwise-consistency problem. *Information Sciences,* 164, pp. 165–188.
8. J.F. Allen (November 1983), Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26, 11, pp. 832–843.
9. J. Renz and B. Nebel (1999), On the complexity of qualitative spatial reasoning: a maximal tractable fragment of the region connection calculus. *Artificial Intelligence (AIJ)*, 108, 1–2, pp. 69–123.
10. J. Renz (August 1999), Maximal tractable fragments of the region connection calculus: a complete analysis. In: *Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI'99),* Stockholm, Sweden.
11. S. Skiadopoulos and M. Koubarakis (2002), Qualitative spatial reasoning with cardinal directions. In: *Proceedings of the 7th International Conference on Principles and Practice of Constraint Programing (CP'02), in Lecture Notes in Computer Science,* Vol. 2470, Springer, Berlin, pp. 341–355.

ICA-SCS DENOISING METHOD FOR WATERMARKING SCHEME

Wenhui Li[1], Dongfeng Han[1] and Yong-Jian Zhang[2]
[1]*Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, College of Compute Science and Technology, Jilin University, Changchun 130012, China*
[2]*Department of Information and Electronical Engineering, Engineering, Shandong Institute Architecture and Engineering, Jinan 250014, China*

Abstract In this paper, we propose a new perceptual digital watermarking scheme based on ICA-SCS denoising and the orthogonal vectors. In the watermarking detection process, the independent component analysis (ICA)-based on sparse code shrinkage (SCS) technique is employed to denoise and making using of the orthogonal vectors' property. By hypothetical testing, the watermark can be extracted exactly. The experimental results show that our method successfully survives image processing operation, image cropping, noise adding and the JPEG lossy compression. Especially, the scheme is robust towards noise adding, image sharpening and image enhancement.

Keywords: ICA-SCS denoising, orthogonal vectors, watermarking, robustness.

1. INTRODUCTION

The digital watermarking has been proposed as a solution to the problem of copyright protection of multimedia data in a networked environment. Several techniques have been developed for watermarking mainly in spatial domain and frequency domain. In Ref. [1], three coding methods for hiding electronic marking in document are proposed. In Refs. [2–4], the watermarks are applied on the spatial domain. The major advantage of the spatial domain method is sample and the major disadvantage of spatial domain watermarking is that a common picture cropping operation may eliminate the watermark. On the other hand, the frequency domain method is more robust than the spatial domain method. So many researchers focus their attentions on this method.

G. R. Liu et al. (eds.), Computational Methods, 1135–1139.

2. OUR APPROACH

2.1 Watermarking permutation and embedding

We use a binary meaningful image of size 64 × 64. In order to resist cropping operation, a fast two-dimensional (2D) pseudo-random number traversing method is used to permute the watermark to disperse its spatial relationship.

During the detection process, we can extract W', and then reverse W' to W.

First we set two orthogonal vectors $P0$ and $P1$:

$$P0 = [0100111100011100] \tag{1}$$

$$P1 = [1011000011100011] \tag{2}$$

$P0$ and $P1$ are selected by the zig–zig scan and then projected to one dimension, which can warrant this method has a good virtue towards JPEG compression. Patterns $P0$ and $P1$ are so selected as to be symmetric and orthogonal to each other. The embedding process follows:

1. For each 8 × 8 block we apply 1-level DWT, then we get three detail sub-bands HL, LH, HH, and one approximate sub-band LL.
2. Selecting one detail sub-band of size 4 × 4, for example HL. The coefficients are modified:

$$\bar{W}(i,j) = \left\{ \begin{matrix} HL(i,j) + alpha * JND(i,j) * P0((i-1)*4+j) if W' = 0 \\ HL(i,j) + alpha * JND(i,j) * P1((i-1)*4+j) if W' = 1 \end{matrix} \right\}$$

$$i, j = 1 \ldots 4 \tag{3}$$

3. Where alpha is a positive number controlling the trade-off between perception and the robust of the watermark. The JND can refer Ref. [5]. For an image of size 512 × 512 we can embed 64 × 64 bits information.
4. Performing the IDWT and forming the watermarked image $\bar{I}$.

2.2 Watermarking detection

The watermarking detection process needs the original image. In the watermark detection process, two steps are used.

1. First the PSNR of a given possibly corrupted image $\tilde{I}$ is calculated. If PSNR is smaller than a given threshold PSNR_0, then the independent component analysis (ICA)-based sparse code shrinkage (SCS) technique [6, 7] is employed to model the denoising problem. Therefore, the noisy image $\tilde{I}$ can be denoted as:

$$\tilde{I} = I + N = As + N. \tag{4}$$

Supposing only the observed data is given, the basis matrix and the ICs can be obtained by first finding a separating matrix W (with $W^{-1} = A$) via sparse coding [6]. Then, can be determined by $s = WX$. After sparse coding, the noisy image $\tilde{I}$ can be transformed by means of W, and a noisy independent component, $s + \tilde{N}$ (in the ICA transformed domain), can finally be derived as follows:

$$W\ \tilde{I} = WI + WN = WAs + WN = s + \tilde{N} \tag{5}$$

Then, we can shrink $s + \tilde{N}$ by means of g and then get the cleaned version of s, which is represented as $\bar{s}$, where

$$\bar{s} = g(s + \tilde{N}) \tag{6}$$

In general, the shrinkage function is explicitly defined [7] based on the sparse density distribution of noisy independent components to have the effects that small arguments are set to zero and the absolute value of large arguments are reduced by an amount depending on the noise level. In the third step, the approximated host image $\bar{I}$ can be derived by an inverse ICA transformation: $\bar{I} = A\bar{s}$. After the estimated host image is determined, it can be used for watermarking detection.

2. After the pre-processing, we then can extract the watermark from image $\tilde{I}$, for each bit the algorithm includes the following steps:
For each 8×8 block of the original image I and the watermarked image $\bar{I}$, we apply 1-level DWT. we get detail sub-bands $HL, LH, HH, \bar{H}L, \bar{L}H, \bar{H}H$ and approximate sub-bands $LL, \bar{L}L$ using:

$$PH0 = ABS(SUM((HL - \bar{H}L) * P0)) \tag{7}$$

$$PH1 = ABS(SUM((HL - \bar{H}L) * P1)) \tag{8}$$

$$\tilde{W}(i, j) = 0 \,\text{if}\, PH0 - PH1 \geq 0 \tag{9}$$

$$\tilde{W}(i, j) = 1 \text{ if } PH1 - PH0 > 0 \tag{10}$$

One bit information can be extracted.
Then we can extract 64×64 bits information. So the extracted watermark $\bar{W}'$ can be obtained by

$$\bar{W}' = InvPermute\ (\tilde{W}, Key(K)) \tag{11}$$

We define the similarity measurement between the referenced watermark W and the extracted watermark $\bar{W}'$ as

$$NC = \frac{\sum_i \sum_j (W(i, j) \oplus \bar{W}'(i, j))}{M \times N} \tag{12}$$

Where $M \times N$ is the size of the watermark.

3. EXPERIMENTAL RESULTS

3.1 Performance comparison

We compare our method with the method proposed by Hsu and Wu [8]. The noise adding experiences of our method can get the value of NC about 0.88, and the method proposed by Hsu and Wu [8] can only get the value of NC about 0.65. The strongly enhanced contrast experiences of our method can get the value of NC is 1, and the method proposed by Hsu and Wu [8] can only get the value of NC is 0.97. When sharpening experiences, the NC value of our method is 1, and the method proposed by Hsu and Wu [8] not mentioned. Our method has a better performance than the method proposed by Hsu and Wu [8], especially under noise adding and signal enhancement attacks.

4. CONCLUSIONS

In this chapter, we propose an adaptive watermarking algorithm based on the ICA-SCS denoising. The algorithm is evaluated from the transparency point of view and the robustness against some common attacks, such as JPEG compression, filtering, noise corruption and cropping. The results show the desirable features of the algorithm, especially for signal enhancement. In addition, the performance of our algorithm is compared with the algorithm proposed by Hsu and Wu [8]. The comparison results show the advantage of our algorithm over the algorithm proposed by Hsu and Wu [8].

REFERENCES

1. J.T. Brassil, S. Low, N.F. Maxemchuk and L. O'Gorman (1995), Electronic marking and identification techniques to discourage document copying. *IEEE Journal on Selection Areas Communication*, 13, pp. 1495–1504.
2. I. Pitas and T.H. Kaskalis (1995), Applying signatures on digital images. In: *Proceedings of IEEE Nonlinear Signal and Image Processing*, pp. 460–463.
3. O. Bruyndonckx, J.J. Quisquater and B. Macq (1995), Spatial method for copyright labeling of digital images. In: *Proceedings of IEEE Nonlinear Signal and Image Processing*, pp. 456–459.
4. S. Walton (1995), Image authentication for a slippery new age. *Dr. Dobb's Journal,* pp. 18–26.
5. M. Barni, F. Bartolini and A. Piva (2001), Improved wavelet-based watermarking through pixel-wise masking. *IEEE Transactions on Image Processing*, 10, 5, 783–791.

6. A. Hyvarinen (1999), Sparse code shrinkage: denoising of nongaussian data by maximum likelihood estimation. *Neural Computation*, 11, pp. 1739–1768.
7. A. Hyvarinen, P. Hoyer and E. Oja (2001), Image denoising by sparse code shrinkage. In: *Intelligent Signal Processing*, S. Haykin and B. Kosko (eds.), IEEE Press, New York.
8. Hsu and Wu JL. (1999), Hidden digital watermarks in images. *IEEE Transactions on Image Processing*, 8, 1, pp. 58–68.

STUDY AND IMPLEMENTATION OF SIMPLIFIED WWW DATA MODEL IN USE FOR WEB MINING

Kanghui, Zhang Xiaoxu and Yang Xiuli
Department of Computer Science and Technology, Jilin University, P.R. China

Abstract Web pages are defined by the HTML. Web pages can be linked by hyper links, so as to construct a hyper media system. For web mining, there must be a data model to describe the rules of interest relation among web pages. In this paper, we simplify the WWW data model that can be used for web mining.

Keywords: WWW, web mining, WWW data model.

1. GENERALIZATION

As the Internet develops increasingly fast, data on the net become richer, but knowledge in the data resources has not been fully mined and used by now. Because web is unstructured and dynamic, the complexity of web pages is much greater than that of text files. There are many index-based web search engines that search the web. However, current keyword-based search engines suffer from several deficiencies. The coverage rate is limited. In addition, they can't provide special services for individual users. A way to solve the problem is to combine traditional data mining technology with web for web mining. Web mining is the application of data mining or other information process techniques to WWW, to find useful patterns. People can take advantage of these patterns to access WWW more efficiently.

Because the web has many semi-structured data and data mining must be based on the structured data, we need to analyse web pages to get a simplified data model.

This paper is to abstract a simplified WWW data model from complex web pages, which can be used as the database for web mining. The data model

G. R. Liu et al. (eds.), Computational Methods, 1141–1146.

abstracts most useful data hidden in web pages, such as page content, page attribute, link, access time of the log and so on, and structurizes them in the database, so as to facilitate web mining.

2. DEFINITION OF SIMPLIFIED WWW DATA MODEL

Web mining is generally defined as abstracting interesting useful models and implicit information from the related www resources and actions.

Web is defined by the HTML. Web pages are linked by hyper links, so as to construct an interlinked hyper media system. To predict user actions, there must be a data model to describe the interest association rules among web pages well. To make web mining easy, the new data model is defined—simplified WWW data model. Firstly the definitions are given as follows:

Definition 1. Page node is represented by three tuples (Pid, P, time), in which Pid identifies with a page node uniquely, P is the attribute set, time is the latest access time,

$$P = \{pi|pi\ is\ the\ attribute, i = 1, 2, \ldots n\}. \tag{1}$$

Page node is the general web page; attribute can be relative URL, type, set of linknodes, content, modifying time and so on. According to the existence and direction of link in the web page, page nodes can be divided into isolated page node, source page node and target page node. Isolated page node is one without any link. Page node with link is called source page node of the link. Page node that the link points to is called target page node of the link. Apparently, for different links, the page node can be both source page node and target page node as well.

Definition 2. Linknode of the page is represented by three tuples (Lid, string, target_node_id), Lid identifies with a linknode uniquely, string describes the linknode's displaying information, target_node_id is the Pid of the target node that the linknode marked by Lid points to.

Definition 3. Linknode in the page is represented by three tuples (source_node, L, target_node), in which source_node is the source page node. L is the linknode of source_node and target_node is the target node.

For the need of data mining, we describe a simplified WWW data model by page nodes, linknodes and links.

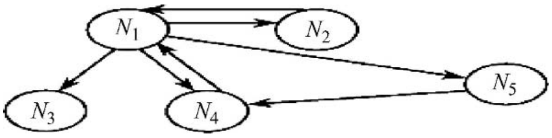


Figure 1. Simplified WWW data model.

Definition 4. Simplified WWW data model can be represented by three tuples (Page_node_set, Page_linknode_set, Link_set), Page_node_set is the set of Page nodes, Page_linknode_set is the set of linknodes, Link_set is the set of links.

As is shown in Figure 1, page nodes N_1–N_5 represent different web pages, respectively, which can be interlinked by oriented edges. These oriented edges represent directly the links between pages.

3. IMPLEMENTATION OF SIMPLIFIED WWW DATA MODEL ON DATABASE

To implement simplified WWW data model on database easily, there should be some modifications about it, so we define three tables on the database as:

Table PageContent (Table 1) stores the content of page node (represented by mainly).

Table Link_node (Table 2) stores linknodes and link information.

Table Page_node (Table 3) stores the node information of the page.

Pid of table Page_node is quoted as foreign key by SourceID of table Link_node and Pid of table PageContent. Page_node.Pid as the unique identifier of page node can be represented as follows: ip of the web site of the page

Table 1. PageContent.

Pid	Keyword
Page node identifier	Keyword of the page node

Table 2. Link_node.

Pid	URL	UpdateTime	AccessTime
Unique identifier of the page node	The relative URL of the page node	The last modifying time of the page	The latest access time of the page

Table 3. Page_node.

SourceID	Lid	String	TargetID
Source page node of the linknode	Linknode identifier	Displaying information of the linknode	Target page node that the linknode point to

and relative route of the page. While Link_node.Lid is represented by number as to its appearance order in the page.

The modified WWW data model—simplified WWW data model is concise, clear and applicable.

4. DATA COLLECTION

Data collection is to read a web file (This program takes Html file as the reading object), then analyse the unformatted data of the file, store useful information in the specific data structure and lastly write it into the database in the form of simplified WWW data model constructed as above.

For writing into the database easily, we define several data structures, which are implemented in the form of class in the programme, the relationship between them is shown in Figure 2.

1. *Class PageInfo* is a class that represents web node (as is shown in Figure 3), whose super class is Object. Its four data members (pageId, url, updateTime, accessTime) correspond to the four arranges of table Page_node, respectively.
2. *Class LinkInfo* is a class that represents linknode (as is shown in Figure 4), whose super class is Object. Its four data members (sourceID, linkID, content, targetID) also correspond to the four arranges of table Link_node,

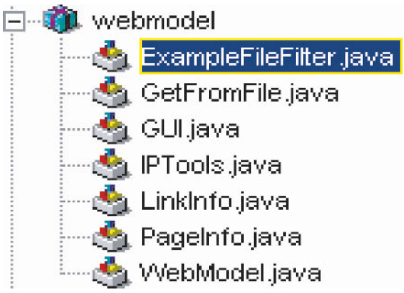
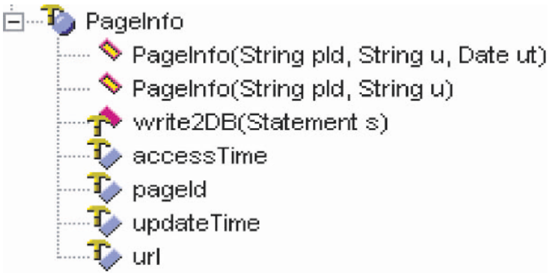


Figure 2. Classes of the program.



PageInfo
PageInfo(String pld, String u, Date ut)
PageInfo(String pld, String u)
write2DB(Statement s)
accessTime
pageId
updateTime
url

Figure 3. Class PageInfo composition.

LinkInfo
LinkInfo(String pID, int lID, String con, String tID)
write2DB(Statement s)
content
linkID
sourceID
targetID

Figure 4. Class LinkInfo composition.

respectively. It has a constructor function and a member function write 2DB(), which has the same function as PageInfo.write2DB described above.

3. The function of *class IPTools* is to get ip of the local computer and some given remote server.
4. *Class WebModel* is the main program of the system, with a main function of main. In the function, firstly it uses function GetFromFile.GetData() described above to read data from a Html file, then calls function processData() to process the data just read from the file, the processed datas are stored in two ArrayList formed by the objects of PageInfo and LinkInfo, then establishes a link to the database, using iterative device to write the content of the two ArrayList into database at a time, lastly closes the link to the database.

5. CONCLUSIONS

From simplifying the WWW data model to constructing the database of a data model and a data collection tool for filling data into the database, this system provides a graphical user interface, which is concise and applicable as one can use the model to do web mining and make high-level management decisions easily.

REFERENCES

1. Han, Meng, Wang, Li (2001), Research on Web Mining. *Journal of Computer Research & Development*, 38, 4.
2. J. Han and M. Kamber (2000), *Data Mining: Concepts and Techniques.* Morgan Kaufmann, San Mateo, CA.
3. S. Chakrabarti (2000), Data mining for hypertext: a tutorial survey. *SIGKDD Exploration*, 1, 2, 1–11.
4. J. Srivastava et al (2000), Web usage mining: discovery and application of usage patterns from web data. SIGKDD Explorations, 1, 2, 12–23.
5. Xie (2001), Web mining technology and designing of the tools. *Journal of Computer Engineering and Applications*.
6. Bruce Eckel (2002), *Thinking in Java*, 2nd ed., translated by Hou. Machine Press, China.

AN EXTENSION OF EARLEY'S ALGORITHM FOR EXTENDED GRAMMARS

Sheng-Jun Wang and Cheng-Zhi Jin
College of Computer Science & Technology, Jilin University, Changchun 130012, P.R. China

Abstract Most programming languages use some variants of EBNF (Extended Backus-Naur Form) to define the grammars of these languages since EBNF can make grammars more compact. Usually, these language parsers can be generated automatically by parser generators such as YACC. However, these parser generators can only work indirectly on EBNF by translating EBNF grammars into equivalent sets of BNF productions. This paper presents an extension of Earley's algorithm which can work on any EBNF grammar directly without grammar conversion. Finally, the time bounds and space bounds are discussed.

Keywords: Earley parsing, extended grammar, ATLAS, operator state machine.

1. INTRODUCTION

The parsing algorithm invented by Earley [1] is a highly practical technique for all context-free grammars (CFGs), especially when handling large languages because the grammars of large languages are often too complex to satisfy the constraints of normal parsing such as LL(1) or LR(1). An excellent parser generator—ACCENT [2] which can generate a parser for an arbitrary context-free grammar is implemented on Earley's algorithm. Later, people improved the algorithm from many aspects, such as: (1) Combine LR method and Earley's method and obtain a faster parser [3]; (2) Give a simple solution to the problem of ε-rule [4].

These algorithms work very well on context-free grammars. But lots of programming languages are defined with EBNFs due to their powerful expression and simple forms [5, 6], and unfortunately, Earley's algorithm can not work on them directly. ACCENT can deal with several extended operators in grammar such as star closure, but in fact it converts extended operators to BNF form first

G. R. Liu et al. (eds.), Computational Methods, 1147–1152.

and then works on the generated BNF, so it is a 'disguised' parser for EBNF grammars. The problem of ACCENT is if there are lots of extended operators in a grammar, it will waste plenty of time to convert them and lead to a lower efficiency. This paper gives a parsing method for EBNF grammars based on Earley's algorithm and a parsing for a concrete grammar of ATLAS-EBNF [7] (a powerful EBNF notation defined in the grammar of ATLAS—a large test language) exemplify the process. The method seems to be useful because we can design a parser for large EBNF grammar directly without grammar translation, thus we can get a faster development of compiler.

2. ATLAS-EBNF GRAMMAR

In this section, ATLAS-EBNF grammar which includes several powerful operators is defined. It will be our working grammar in later sections.

Definition. An **ESE** (Extended String Expression) is defined as below:

(1) $a \in$ ESE, if $a \in V_T$ is a terminal.
(2) $A \in$ ESE, if $A \in V_N$ is a nonterminal.
(3) $\{\kappa^n (S_1, S_2, \dots, S_n)\} \in$ ESE, if $S_i \in$ ESE $(1 \leq i \leq n)$, $\kappa^n \in V_K$ is a n-ary operator. For example, $\{A \mid \{B\#C\}\}$ is a legal ESE. The ESE with operators must enclosed by '{' and '}' to decide the computing order.

Definition. An **extended grammar** G^{ESE} based on some ESE is a quintuple $<V_T, V_N, V_K, S, P>$ where V_T is a set of terminals, V_N is a set of nonterminals, V_K is a set of operators, S is the start symbol of grammar and P is a set of productions. Each production is of the form $A \to X_1X_2\dots X_n$ where $X_i \in$ ESE $(1 \leq i \leq n)$.

Definition. The **language** defined by some extended grammar $G^{ESE} = <V_T, V_N, V_K, S, P>$ is denoted as $L(G^{ESE})$ where

$$L(G^{ESE}) = \{\Gamma(\beta) \mid S \Rightarrow^* \beta, \beta \in (V_T \cup V_K)^*\}$$

where Γ is a function on $(V_T \cup V_N \cup V_K)^* \to (V_T \cup V_N)^*$ whose definition depends on the definition of the corresponding ESE. For some $\beta \in (V_T \cup V_K)^*$, we can conclude $\Gamma(\beta) \in V_T^*$. The symbol $\Rightarrow^*$ is the closure of derivation.

Definition. ATLAS-EBNF grammar: $\mathbf{G}_{atlas} = (V_T, V_N, V_K, S, P)$, where V_T is a set of terminals, V_N is a set of nonterminals, $V_K = \{*, \%, \circ, |, \#, +, \{\}, []\}$, the definition of each ESE on the right-hand side of productions and the corresponding function Γ are shown in Table 1.

Table 1. Definition of ATLAS-EBNF grammar.

$E \to a$	Terminal	$\Gamma(a) = \{a\}$
$\|A$	Nonterminal	$\Gamma(A) = \{A\}$
$\|E^*$	Zero or more repetition of E	$\Gamma(E^*) = \Gamma(E)^*$
$\|E\%$	One or more repetition of E	$\Gamma(E\%) = \Gamma(E)^+$
$\|E_1\|E_2$	E_1 or E_2	$\Gamma(E_1\|E_2) = \Gamma(E_1)\|\Gamma(E_2)$
$\|E_1\#E_2$	List of E_1 separated by E_2	$\Gamma(E_1\#E_2) = \Gamma(E_1)\Gamma(E_2E_1)^*$
$\|E_1 + E_2$	E_1 or E_2 or E_1E_2	$\Gamma(E_1 + E_2) = \Gamma(E_1)\|\Gamma(E_2)\|\Gamma(E_1E_2)$
$\|\{E\}$	To group elements	$\Gamma(\{E\}) = \Gamma(E)$
$\|[E]$	Zero or one occurrence of E	$\Gamma([E]) = \{\varepsilon\}\ \|\Gamma(E)$
$\|E_1 \circ E_2$	Sequence of E_1 and E_2, E_1E_2 for short	$\Gamma(E_1E_2) = \Gamma(E_1)\Gamma(E_2)$

3. STATE MACHINE OF OPERATOR

Definition. An ESE is called a **simple ESE** if it contains only one operator.

Definition. The simple ESE $\{\kappa^n (S_1, S_2, \ldots, S_n)\}$ with a dot at some position is called a **state** of operator κ, which is of the form: $\{X_1 \ldots \bullet X_i \ldots X_m\}$, $X_i \in \{\kappa, S_j\}$ $(1 \leq j \leq n)$. The initial state of κ is a state with a leftmost dot, while the final state is a state with a rightmost dot. For instance, $\{\bullet A\#B\}$ and $\{A\#B\}\bullet$ are states of operator #.

Definition. If A_i is a state of operator κ, the following rules define the state's **transition** of κ when A_i sees a symbol Y (namely Y is the symbol after dot).

1. If $Y \in \{`\kappa', `\{', `\}'\}$, then a transition occurs. We draw an edge labelled Y from A_i to A_j if the state reached is A_j.
2. If Y is a terminal or nonterminal, the state isn't changed but the dot moves over one symbol to indicate that that symbol has been scanned.

Definition. The **state machine** of operator κ is built by following steps: (1) Draw all states of κ; (2) Draw all edges between states according to the meaning of κ.

Definition. The state machines of operators in ATLAS-EBNF grammar are shown in Figure 1 according to their meanings (only list one due to the space limitation).

Where every box except dashed ones represents a state, the digit in left top corner of box with a subscript operator denotes the state number and double-line boxes are final states. Dashed boxes denote substates which can only see terminals and nonterminals, so the transition from one substate to another isn't a state transition. Thus, according to the state machine graph, given a state of some operator, we can compute its successors. If the structure of state is

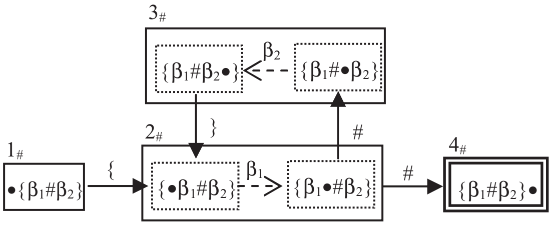


Figure 1. The state machine of #.

SM(#) : (1) Trans(<$1_{\#}$, j>, '{')
= {<$2_{\#}$, j+1>}
(2) Trans(<$2_{\#}$, j>, '#')
= {<$3_{\#}$, j+1>, <$4_{\#}$, j+$|\beta_2|$+2>}
(3) Trans(<$3_{\#}$, j>, '}')
= {<$2_{\#}$, j-$|\beta_1|$-$|\beta_2|$-1}

Figure 2. The equations of state machines.

<Sn, Dot> where Sn denotes state number and Dot denotes a position of dot, a state machine graph can also be described by several transition equations. In Figure 2 (only list one), SM(κ) denotes the state machine of operator κ, Trans is a transition function and $|\beta|$ denotes the number of symbols in β.

4. EARLEY PARSING FOR EXTENDED GRAMMARS

The following is a precise description of the parsing algorithm for input string $X_1 \ldots X_n$ and grammar G. We define the same notation as Earley's in his paper. Number the productions of grammar G arbitrarily $1, \ldots, d-1$, where each production is of the form $D_p \rightarrow C_{p1} \ldots C_{pp'}$ $(1 \leq p \leq d-1)$ where p' is the number of symbols on the right-hand side of the pth production. Add a 0th production $D_0 \rightarrow R \dashv$ where R is the root of G, and $\dashv$ is a new terminal symbol denoting the end of the input.

Definition. An **item** *i* is a quadruple <p, j, f, t>, where p, j and f are integers $(0 \leq p \leq d-1)$ $(0 \leq j \leq p')$ $(0 \leq f \leq n+1)$, t is a stack and i is the number of the item. In fact, p denotes the number of a production, j represents a position in the rule's right-hand side and f is a pointer to item set S_f.

Definition. An item i: <p, j, f, t> is **final** if $j = p'$ (the dot is at the end of the rule). We add a state to a set by putting it last in the ordered set unless it is already a member.

The recognizer of ATLAS-EBNF grammar is a function of two arguments REG(G, $X_1 \ldots X_n|$) computed as follows:

- Let S_i be empty ($0 \leq i \leq n + 1$)
- Add initial item 1: <0, 0, 0, nil> to S_0
- For i ← 0 until n step 1 do

  Begin

  ♦ Process the items of S_i in order, performing the following two steps (1) and (2) on each item m:<p, j, f, t>∈ S_i.

  1. Perform one of the following four operations on item m.
     (a) Predictor: If m is non-final and $C_{p(j+1)}$ is a nonterminal, then for each q such that $D_q = C_{p(j+1)}$, add m':<q, 0, i, nil> to S_i where m' becomes the number of last item in S_i.
     (b) Completer: If m is final, then for each k:<p_1, j_1, f_1, t_1 >∈ S_f such that $C_{p1(j1+1)} = D_p$ add m': <$p_1, j_1 + 1, f_1, t_1$ > to S_i.
     (c) Scanner: If m is non-final and $C_{p(j+1)}$ is a terminal, then if $C_{p(j+1)} = X_{i+1}$, add m': <p, j+1, f, t> to S_{i+1}.
     (d) Extender: If m is non-final, $C_{p(j+1)}$ is neither a terminal nor a nonterminal (namely an operator or '{' or '}') and t = u:t' where u is the top element of t, then for each <sn,dot> ∈ Trans(<u, j>, $C_{p(j+1)}$) add m': <p, dot, f, sn:t'> to S_i.
  2. For each new item m': <p', j', f', t'>, perform the following steps in order.
     (a) If top(t) is a final state of some operator, then pop it.
     (b) If the item processed by step (a) is of the form m'': <p', j', f', t''> and $C_{p(j+1)}$ is '{' which means an ESE of some operator κ is to be processed, then push the initial state of $\kappa - 1_\kappa$ into stack t'', thus m'' is changed to m'': <p', j', f',1_κ: t''>.

  ♦ If S_{i+1} is empty, return rejection.

  ♦ If i = n+1 and S_{i+1} = {<0, 2, 0, nil>}, return acceptance.

End.

5. CONCLUSION

Our parser can work on extended grammars directly. For an arbitrary EBNF, we can design its parser by defining the corresponding operator state machines. However, there exists precedence and associativity between operators in ATLAS-EBNF grammar, which are not considered in this paper, so one of the future works seems to be how to implement them. Another work is how to implement the performance of the semantic actions embedded in an extended grammar. In BNF case, we can get a unique derivation when finish parsing a sentence with Earley's algorithm, thus the performance order of actions can be decided. But how about in EBNF case?

REFERENCES

1. J. Earley (1970), An efficient context-free parsing algorithm. *Communications of the ACM*, 13, pp. 94–102.
2. F.W. Schröer (2000), The ACCENT Compiler Compiler, Introduction and Reference. GMD Report 101, German National Research Center for Information Technology.
3. P. McLean and N. Horspool (1996), A faster Earley parser. In: *Proc. 6th Int. Conf. on Compiler Construction, April 24–26*. *Lecture Notes in Computer Science*, vol. 1060, pp. 281–293.
4. J. Aycock and R.N. Horspool (2002), Practical Earley parsing. *The Computer Journal*, 45, 6, pp. 620–630.
5. N. Wirth (1977), What can we do about the unnecessary diversity of notation for syntactic definitions? *Communications of the ACM*, 20, 11, pp. 822–823.
6. M. Howard Williams (1982), A flexible notation for syntactic definitions. *ACM*, 4, 1, pp. 113–119.
7. 716 C/ATLAS Subcommittee of IEEE Standards Coordinating Committee 20 (1995), Standard Test Language for All Systems—Common/Abbreviated Test Language for All Systems (C/ATLAS).

THE RESEARCH ON POLICY-BASED REGISTRATION MECHANISM OF MOBILE TERMINALS IN MOBILE IP NETWORK

Wei Da, Liu Yan-Heng and Li A-Li
Computer Science and Technology College, Jilin University, China

Abstract With the rapid development of Mobile IP and PBNM applying PBNM in Mobile IP network is a new management blue print of Mobile IP network. The management blue print is hoped to effectively resolve registration management, QoS management and security management questions of Mobile IP network. The registration management of mobile terminals is the base of other managements. Policy-based registration mechanism of mobile terminals in Mobile IPv4 network is different from that in Mobile IPv6 network. The COPS-MIP which has been defined supports policy-based registration of mobile terminals in Mobile IPv4 network. So the paper gives the COPS-MIPv6 which is a new protocol extending from COPS. It is used to support policy-based registration of mobile terminals in Mobile IPv6. There is an analysis about COPS-MIPv6. Also an example is given to explain the application of the COPS-MIPv6.

Keywords: mobile terminals, COPS, COPS-MIP, COPS-MIPv6, PBNM.

1. INTRODUCTION

PBNM [1] (Policy-Based Network Management) is the outcome which adapts the complex and dynamic task object of network management. IETF and DMTF define Policy Core Information Model [2] and its spread, and they raise the frame [3] of PBNM system. PBNM system has begun to take shape. Through the research and development of decades, many products have already come out, and AppDrrvN of the IBM Company is one of that.

Mobile IP [4] makes the mobile hosts maintain communication in the constant movement of different network under the prerequisite of present Internet based network prefix and is a solve scheme of offering mobility support based on network layer in Internet.

G. R. Liu et al. (eds.), Computational Methods, 1153–1157.

The increase of mobile handle official business crowds accelerates the development of Mobile IP network. The users of many mobile equipments hope to maintain the access to Internet and continuous communication in the process of movement, and obtain the same network QoS (Quality of Service) as this of immobile access. This inducts the problem of Mobile IP network management, especially QoS management. Applying PBNM in Mobile IP network will be a new management blue print. The first part mostly discusses policy-based registration mechanism of mobile terminals in Mobile IP network. The second part gives an analysis of COPS-MIPv6 which is extended from COPS. The last part gives an example of COPS-MIPv6's application.

2. POLICY-BASED REGISTRATION MECHANISM IN MOBILE IP NETWORK

The main character of Mobile IP network is that there are many mobile nodes in the immobile network. When a mobile node moves to a new network, it must finish the registration to home network, namely notify home agent its Care-of address. In October of 2000, IETF has raised COPS [5] usage for Mobile IP (MIP) [6] draft. The draft raises policy-based registration mechanism in Mobile IPv4 network. Policy-based registration process of mobile terminals is shown in Figure 1 [6].

The step is interpreted as follows [6]:

1. MN (Mobile Node) sends a Registration Request to FA (Foreign Agent);
2. FPEP (Foreign Policy Enforcement Point) and FPDP (Foreign Policy Decision Point) interact for policy decisions for Registration Request;
3. FA relays Registration Request to HA (Home Agent);
4. HPEP (Home Policy Enforcement Point) and HPDP (Home Policy Decision Point) interact for policy decisions;
5. HA sends Registration Reply to FA;
6. FPEP and FPDP interact for policy decisions for Registration Reply;
7. FA forwards Registration Reply to MN.

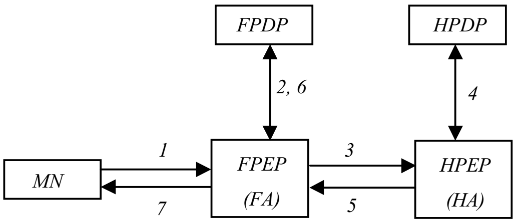


Figure 1. Typical policy control registration in mobile IPv4 network.

Mobile IPv6 needn't configure a special router as foreign agent as Mobile IPv4 does. So mobile terminals of Mobile IPv6 network only has a Care-of address type—Co-located Care-of Address. COPS-MIP doesn't fit policy-based registration of mobile terminals in Mobile IPv6 network. A new protocol must be introduced to solve the question. Chaouchi has extended a new protocol COPS-mt [7]. The key of extending protocol is defining new objects.

3. AN ANALYSIS OF COPS-MIPv6

COPS-MIP is an extension of COPS, and it supports policy-based registration mechanism of mobile terminals in Mobile IPv4 network. For Mobile IPv6 network, we need to extend a new protocol from COPS. COPS not only supports the address format of IPv4, but also IPv6, which can make extension easy. We name the new protocol as COPS-MIPv6. There are several aspects to be discussed as follows:

1. COPS has a very good expansibility. By adopting different client type, it can support much more abundant management functions. Client-type in Common Header of COPS [7] is a 8-bit byte, and identifiers which management information type COPS message is transmitting. COPS regulates that the client type of COPS-RSVP [9] is '1', the client type of COPS-PR [10] is '2' and the client type of COPS-MIP is 'COPS-MIP'. So the client type of COPS-MIPv6 can be defined 'COPS-MIPv6'. All messages between PDP and PEP identify different client type with Client-type value after finishing client open/accepted messages.
2. In the registration progress, both request messages and decision messages of COPS include one or many Context-Object. The object depicts network events that spring the request. The R-Type field of Context-Object depicts base types of events touched off. In the registration process, also two events can be touched off: Incoming-Message request and outgoing-Message request. The other M-Type field of Context-Object identifies which message type registration messages that touches off the request is, whether binding update message or binding acknowledgement message.
3. COPS messages can encapsulate diversified information format. COPS messages can transmit all kinds of length-alterable, self-defined and opaque data. ClientSI (Client Specific Information) of COPS-MIPv6 is one of that. All objects which are in registration message are encapsulated in the ClientSI object. Some objects will be defined, and these objects can be used in the ClientSI object, the Client Specific Information object and the Replacement Data Object.

Table 1. The format of COPS-MIPv6 objects.

Length	S-Num	S-Type
Value......		

The parameter format is the same as that of COPS-MIP. Some objects are shown as follows:

1. Fixed-length Portion Object (FLO) for binding update message. S-Num = 1, S-Type = 1, this object contains the fixed-length portion of the binding update message;
2. Fixed-length Portion Object (FLO) for the binding acknowledgement message. S-Num = 1, S-Type = 3, this object contains the fixed-length portion of the binding update message;
3. Mobile Option Object. S-Num = 2, S-Type = 1, this object is used to take Mobile Option of the registration message;
4. Specific Error Object for Mobile Option. This object is used to show specific error object for Mobile Option. This object indicates that the PDP has encountered a list of mobile option that it does not recognize. The value of parameter is uniform to that of COPS-MIP [6];
5. Error Specific Data object. This object contains error specific information. The definition is uniform to that of COPS-MIP [6]. The error code values that are used in the S-Type are shown as follows: Registration denied by the FPDP [6]; Registration denied by HPDP [11].

4. AN EXAMPLE

We support that the policy information model has been defined. There are HomeAddress class and Care-ofAddress class in the policy information model. There is such a strategy in the policy depository of the foreign subnet: IF HomeAddress > 10.30.84.0) and (HomeAddress < 10.30.84.254) then (Allow). There is such a strategy in the policy depository of the home subnet. IF (Care-ofAddress > 10.60.60.0) and (CareofAddress < 10.60.60.254) then (Allow).

When the mobile terminal (we suppose that the mobile terminal is policy-aware, namely a PEP) whose home address is 10.30.84.130 moves to the foreign subnet. It will obtain a Care-of address—10.60.60.60 and a home agent address. The mobile terminal sends the binding update message. The source address is the Care-of address, and the destination is the home agent. There is home address in the mobile option. The binding update message is sent to the TFPDP through COPS-MIPv6.The TFPDP adopts the corresponding policy and makes a decision that allows the mobile terminal whose home address is 10.30.84.130

accesses the foreign subnet. The mobile terminal (TPEP) forwards the binding update message to the THA (THPEP). The commutation protocol between the THPEP and the THPDP is COPS-MIPv6. The THPDP make an affirmative decision according the upper policy. The THA sends the binding acknowledge to the mobile terminal. The registration process has finished.

5. CONCLUSION

From the paper, we can realize that policy-based registration mechanism of mobile terminals in Mobile IP network can solve the registration problem of mobile terminals to home network. Once the registration succeeded, PBNM system can fleetly and successfully management mobile terminals. The example of the paper distinctly explains the progress of policy-based registration of mobile terminals, at the same time indicates the registration mechanism can effectively assure the registration of mobile terminals. Previously, we have finished the research and realization of PBNM system in immobile network. Now we solve the policy-based registration problem of mobile terminals in mobile IPv6 network. These are the base of making the research on policy-based management of QoS, security aspects, etc. in Mobile IP network. We will make a deep research on policy-based management of QoS, security aspects, etc. in Mobile IP network.

REFERENCES

1. Jean-Christophe Martin (1999), *Policy-Based Networks*.
2. B. Moore, E. Ellesson, J. Strassner and A. Westerinen (2001), Policy Core Information Model—Version 1 Specification, RFC3060.
3. Li Jin-Ping and Gao Dong-Jie (2002), The research and design of policy-based network management software platform. *Computer Engineering and Application*, pp. 177–9.
4. Sun Li-Min, Gan Zhi-Gang, ect. (2003), *Mobile IP Technology*. Electronic Industries book concern, Beijing.
5. J. Boyle (2000), The COPS (Common Open Policy Service) Protocol, D. Durham (ed.), RFC 2748.
6. M. Jaseemuddin and A. Lakas (2000), COPS usage for Mobile IP, Internet draft.
7. H. Chaouchi (2004), A new policy-aware terminal for QoS, AAA and mobility management. *International Journal of Network Management*, 14, pp. 77–87.
8. J. Boyle, R. Cohen, D. Durham, S. Herzog, R. Raja and A. Sastry (2000), COPS usage for RSVP, RFC2749.
9. K. Chan, J. Seligson, D. Durham, ect. (2001), COPS usage for Policy Provisioning (COPS-PR), RFC3084.
10. D. Johnson, C. Perkins and J. Arkko (2004), Mobility Support in IPv6, RFC3775, pp. 141–2.

OPERATIONAL SEMANTIC TO THE EXECUTION OF THE PROCESS MODEL

Fu Yan-ning, Liu Lei and Li Bo
College of Computer Science and Technology, Key Laboratory of Symbolic Computation
and Knowledge Engineering of Ministry of Education of P.R. China
Jilin University, Changchun, 130012, P.R. China

Abstract The procedure of process models executing in workflow engine is described using operational semantic. The formalized description of process instances and activity instances lead directly to the realization of workflow engine. Therefore, it provides a unified processing procedure for the realization of the workflow engine software, which makes it convenient for the cooperation among heterogeneous workflow engines.

Keywords: workflow engine, process model, process instance, activity instance, operational semantic.

1. INTRODUCTION

The so called ‘Process meta-model’ is the model about model which describes the process model [1]. It defines the structures and rules of semantic model, and is mainly used to describe the fundamental components of process model and the mutual quoting relation as well as the those elements’ attributes [2]. The process model shows how to represent a real running business structure as a work procedure [3]. When the process model operates in a workflow engine, in the light of the description file of the process model and other relevant information, the workflow engine generates the corresponding process instances and activity instances and controls their creation, activation, suspension and termination. This is a significant function of the workflow engine. Workflow Management Coalition (WfMC) defines the running state of process instance and activity instance, and it also provides the conditions for transiting such state [4], as Figure1 and Figure 2 illustrate.

G. R. Liu et al. (eds.), Computational Methods, 1159–1164.

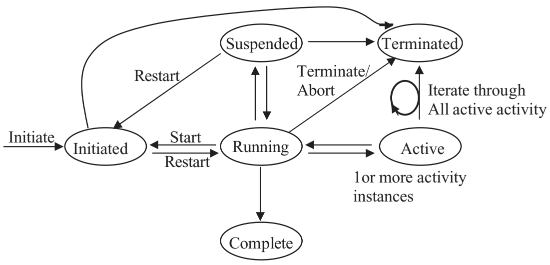


Figure 1. State transitions for process instances.

Therefore, the existent workflow standard provides unified interface and definition file for the product of workflow. But since the overall structures and the realizing methods vary, so do the processing of process instances and activity instances in different workflow engines. The realization of workflow engine will become more explicit if we can completely and specifically describe the processing steps of process instances and activity instances in the workflow engine with a descriptive tool that is irrelevant to the programming language. Then it's possible to exchange these instances among different workflow engines. As a result, coordinating in those different engines will be much easier and faster.

Formalization and exactness form a major characteristic of the operational semantic, which allow for an exact description of the semantic meaning of a certain language [5]. If the definition file of process model can be viewed as a program of a given computer language, and the workflow the abstract engine, we can draw the operational semantic into the description of workflow in order to give a full picture of the execution of process model running in a workflow engine.

2. THE FORMALIZED DESCRIPTION OF WORKFLOW ENGINE

In the first place, design the abstract engine of process instances and activity instances. Furthermore, define the running rules of the process instances. The

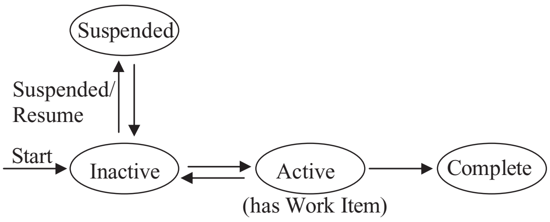


Figure 2. State transitions for activity instances.

initial state of process instance is settled when a certain process model is given. And it will transit in accordance with state transition rules until the terminal state appears.

2.1 Process instance

The process instance can be replaced by sextuple, namely,

(S_0, E_0, ActiSet, CurASet, TranSet, RDSet)

S_0	stands for start activity
E_0	stands for end activity
ActiSet	stands for activity set
CurASet	stands for current activity set
TranSet	stands for transit set
RDSet	stands for relevant data set

Note: 'S_0, E_0, ActiSet, TranSet' belong to the static environment and they can be required by parsing file XPDL. 'CurASet, RDSet' belong to the dynamic environment, and they will change in accordance with the operation of process instances.

RDSet $= \{ \text{id} \rightarrow \text{val} \}$, the relevant data is represented as 'id' in the workflow and the corresponding values is represented as 'val'.

TranSet $= \{ T_{ij} \mid$ conditions to transit from S_i to S_j, $i,j = 0,1,2,\ldots \}$

ActiSet $= \{ S_i \mid i = 0,1,2,\ldots \}$

CurASet $= \{ S_j \mid j = 0,1,2,\ldots \}$

2.2 Activity instances

Activity instances can be described by sextuple, namely, (precondition, postcondition, state, application, followSet, preSet).

Note: Every 'S' that belongs to the current activity set, namely S(S $\in$ CurASet), must be in accordance with the sextuple mentioned above. The corresponding values of each element of sextuple should be:

1. S.precondition = AND/OR
2. S.postcondition = AND/OR

These two attributes stand for the precondition and postcondition of an activity respectively, and their values can be either AND or OR. An activity may denote the action of 'join' and 'split' under a certain confining condition. If the precondition is set as AND, this activity doesn't operate until all the previous parallel activities terminate. On the contrary, if the precondition is set as OR, this activity will operate as long as any one previous activity has finished. Meanwhile, if the postcondition is AND, all the relevant post activities have to be fulfilled, whereas if it is OR, only those activities that comply with the condition can be executed [6].

3. S.state=R(run)/S(suspension)/T(terminate)/W(wait)

As is illustrated by Figure 2, there are four states of activity, namely, R(run), W(wait), S(suspension) or T(terminate).

4. S.application

It refers to the external application that the activity has to invoke.

5. S.followSet

It refers to the post activity set where 'S' shifts to, (i = 0,1,2 . . .), and for every activity in workflow, there is S_i.followSet = { $S_j | t_{ij} \in$ TranSet }.

6. S.preSet

It refers to the pre-activity set whose activities can shift to 'S', (i = 0,1,2 . . .), and for every activity in the workflow, there is S_j.preSet = { $S_i | t_{ij} \in$ TranSet }.

2.3 Rules for transition of states

2.3.1 The initialization of process instances and activity instances

When the description file of workflow is submitted to the workflow engine, the corresponding initial sextuple of process instance is:

(S_0, E_0, ActiSet, CurASet, TranSet, RDSet)

S_0, E_0, ActiSet, TranSet can be parsed through the description file of process model, CurASet = { S_0}, and the value of sextuple in RDSet can be set in reference to the relevant data in the file.

The sextuple corresponding to activity 'S' are:

(precondition, postcondition, state, application, followSet, preSet)

'Precondition, postcondition, application' can be parsed through the description file of process model, and set 'state = W, followSet = { $S_i | t_{0i} \in$ TranSet }, preSet= ={ }'.

2.3.2 Executing changes of process instances and activity instances

When operating, each activity in the current activity set(CurASet)equals to only one item of the workflow task list, that is, for '∀ S∈CurASet', there is:

1 If S.state= =R, for the sextuple of process instances, there is

(S_0, E_0, ActiSet, CurASet, TranSet, RDSet_i) →

(S_0, E_0, ActiSet, CurASet, TranSet, RDSet_j)

Note: When the activity is in the state of execution, only do the relevant data change in response to the external application, namely, RDSet_j=S.application(RDSet_i). While the sextuple of the activity intance takes no change, there is:

(precondition, postcondition, state, application, followSet, preSet)→

(precondition, postcondition, state, application, followSet, preSet)

2 If S.state==S or S.state==W, it shows the activity is in the state of suspension or waiting for resources, therefore, there is no change in neither process instances nor activity instances.

For process instances, there is:

(S_0, E_0, ActiSet, CurASet, TranSet, RDSet)

For activity instance, there is:

(precondition, postcondition, state, application, followSet, preSet)

3 If S.state==T, it shows the external applications that the activity is to invoke are ended, so the workflow engine has to transit the activity in accordance with the relevant data, for process instances, there is:

(S_0, E_0, ActiSet, CurASet_i, TranSet, RDSet)→

(S_0, E_0, ActiSet, CurASet_j, TranSet, RDSet)

If the activity instance is ended, CurASet will change.

Suppose PostSet refers to the next set of activity instances that the activity shifts to, and AddSet refers to the set of activity instances that should be added, then:

PostSet={ S_i | S_i ∈ S.followSet,t ∈ TranSet,t(S,S_i)==true }

Now, make the relevant changes of the activity instance S_i, (S_i ∈ PostSet).

Set S_i.PreSet = S_i.PreSet-{ S },

AddSet={ S_i| S_i ∈ PostSet, (S_i.precondition==AND ∧S_i.PreSet=={ }) ∨(S_i.precondition==OR)}

if S.postcondition==AND, then set S.followSet=S.followSet—PostSet

if S.postcondition==OR or S.postcondition==AND&S.followSet=={ }, then

CurASet_j=CurASet - {S} ∪ AddSet

if S.postcondition==AND&S.followSet !={ },then CurASet_j=CurASet ∪ AddSet

2.3.3 Termination of the process instances

During the operation, the executing of process instances terminate only if CurASet=={ E_0}. Then the whole process instances terminate.

3. CONCLUSION

Operational semantic has its own merits of formalization and exactness [7]. It can accurately depict each step of process instances functioning in a workflow engine. Therefore, it not only guides the realization of workflow engine software, but also makes it more convenient to coordinate among the heterogeneous workflow engines. Next, we will aim at the description of operational semantic to heterogeneous workflow engines, in the hope of making those engines more flexible.

REFERENCES

1. Fan Yu-shun (2001), *Foundation of Workflow Management and Technology*. Tsinghua University Press, Beijing, China.
2. Zhao Wen and Hu Wen-hui (2003), Study and application of a workflow meta-model. *Journal of Software*, 14, 6, pp. 8–14.
3. SUN Rui-Zhi and ShiMei-Lin (2003), A process meta-supporting dynamic change of workflow. *Journal of Software*, 14, 1, pp. 62–7.
4. WFMC (1994), WorkFlow Management Coalition, Workflow Reference Model. WFMC-TC00-1003 Brussels, Belgium.
5. Jin Cheng-zhi (1997), *Theory and Technology of Program*. Jilin University Press, Changchun, China.
6. Wil M.P. van der Aalst (2003), Patterns and XPDL : A Critical Evaluation of the XML Process Definition Language [DB/OL]. http://tmitwww.tm.tue.nl/research/patterns/download/ce-xpdl.pdf.
7. G. Winskel (1993), *The Formal Semantics of Programming Languages: An Introduction*. Massachusetts Institute of Technology.

FUZZY CLUSTERING ON THE WEB IMPLEMENTED BY JSP TECHNOLOGY

Ágnes B. Simon and Dezső Kancsár
College of Nyíregyháza, 31 Sóstói street, Nyíregyháza, Hungary

Abstract We have developed an interactive Web application implemented by Java Server Pages technology for executing the fuzzy c-means clustering method on the input data given by users from client side. On the server side the Tomcat application server dispatches the requests, produces the answers and presents them on the client side. The server machine is not a robust computer (Intel Pentium II category). All of the instaled softwares are under free licence policy.

Keywords: fuzzy c-means clustering, JSP technology, Web application.

1. INTRODUCTION

In general, the cluster analysis (or simply clustering) refers to a lot of methods which try to subdivide a data set X into C subsets (or into C clusters) which are pairwise disjoint, all non-empty, and they reproduce X set via union. In this case, the clusters are termed hard (non-fuzzy) c-partition of set X. The hard methods do not mirror the native substructure of the data set. This problem is resolved by fuzzy clustering. The idea of fuzzy set and membership function was introduced by Zadeh [1]. The values of this function are called memberships. Each individual sample point will have memberships in every cluster. The sum of the memberships for each sample point must be equal to 1. If we have a data set which contains N sample points $\mathbf{x}_k \in R^n$ and we want to separate them into 'C' fuzzy clusters, the result of fuzzy clustering is a membership matrix which has N rows and C columns. The criteria on memberships is formulated in the Fuzzy C-means clustering model formulas (1) and (2). We have to minimize

G. R. Liu et al. (eds.), Computational Methods, 1165–1169.

the extended squared error criterion function J_m (1), subject to the conditions (2).

$$\mathbf{J}_{\mathrm{m}}(\mathbf{U},\mathbf{V})=\sum_{k=1}^{N}\sum_{i=1}^{C}(\mathbf{u}_{\mathrm{ik}})^{\mathrm{m}}\left\|\mathbf{x}_{\mathrm{k}}-\mathbf{v}_{\mathrm{i}}\right\|_{\mathrm{A}}^{2} \tag{1}$$

$$\sum_{i=1}^{C}u_{\mathrm{ik}}=1 \quad 1\leq \mathrm{k}\leq \mathrm{N}, \quad \sum_{k=1}^{N}\mathrm{u}_{\mathrm{ik}}>0 \quad 1\leq \mathrm{i}\leq \mathrm{C} \tag{2}$$

U is the matrix of u_{ik} memberships, **V** denotes the matrix of cluster centres (means), $m > 1$ weighting exponent, $\mathbf{x}_k$ is a feature vector of sample point k in the n-dimensional data set X, $\mathbf{v}_i$ vector denotes the mean vector of sample points in cluster i, A positive definite scaling matrix to induce different norms.

To acquaint people with fuzzy c-means clustering method and give them opportunity to use it via Internet we begun the work on the project. The aim of our project was to develop an interactive Web application based on this model implemented by the JSP technology.

Why did we choose the JSP technology? Prior to this project we created a C language program [3] executing a fuzzy c-means algorithm published by Bezdek et al. [2]. It runs on single PC with OS DOS/Win98. We had practical experience in Java programming and the conversion of the C code was not so hard. The Java technologies are platform-independent. The all softwares we had were under free software licence. To implement the aimed services we needed softwares being free of charge as well.

In the next sections, we describe the development work.

2. PREPARING STEPS

At start of the project the hardware and software tools consist of the PC server machine (Intel Pentium II category), with Debian GNU/Linux 3.0 OS, MySQL3.23.37 database server, Apache1.3.39 Web server, networked on the Internet at the College.

We made the following steps to extend the environment for development and execution:

1. *Downloading and instaling the Java2 Standard edition developmentkit (j2sdk1.3.1) from http://java.sun.com/j2se/.* After running the binary executable we had to set up the environment variables: JAVA_HOME and PATH.
2. *Creating a new database for fuzzy application.* In our software architecture, we created the ‘Fuzzy’ mysql database and its tables granted limited privileges to tomcat@localhost user. In the database, we granted permission to create and manipulate its own data: like users and calculations.

3. *Downloading and instaling Tomcat Application Server (v.3.3.1) from http://jakarta.apache.org.* The instalation was very easy. We had only to unpack the downloaded binary packed file. The key directories inside the TOMCAT_HOME are: /bin; /conf, /log; /doc; /webapps. The next step was a little bit complicated, we needed to configure the global settings of the Tomcat server (conf/server.xml). We configured the Tomcat's database Realm which is the resource of the authentication in the mysql database server and the previously created fuzzy database.
4. *Downloading and Instaling Apache Ant development tool (v.1.4) from http://ant.apache.org*. We notice that this tool was used on Windows platform.

3. DEVELOPING STEPS

3.1 Designing

At the design time we organized our application into private and public sections with '/demo' and '/full' parts to handle our user's rights with the container managed authentication and role. These requirements are fulfilled in this architecture by the code hierarchy and the Web application security standards [4].

3.2 Creating the fuzzy clustering Web application

Writing codes:

- Web components:
- JSP components: *index.html, input.jsp, compute.jsp, validate.jsp, etc.*
- Java Bean components: *Fuzzy.java, Registration.java, Monitor.java*, etc.
- *build.xml, web.xml*

Deployment:

In our case, we chose the packaged file structure for the deployment. We packaged our application into the archive file '*fuzzy.war*' using the Ant *war* task.

Both the compilation and deployment was made by Ant using the build.xml at the development environment.

Later at the execution environment the Tomcat's auto-deploy was used from the *fuzzy.war* file. Before starting the Tomcat we had copied the *fuzzy.war* into the directory $*TOMCAT_HOME/webapps*.

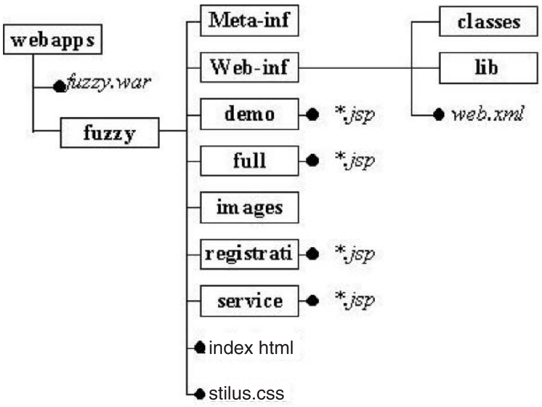
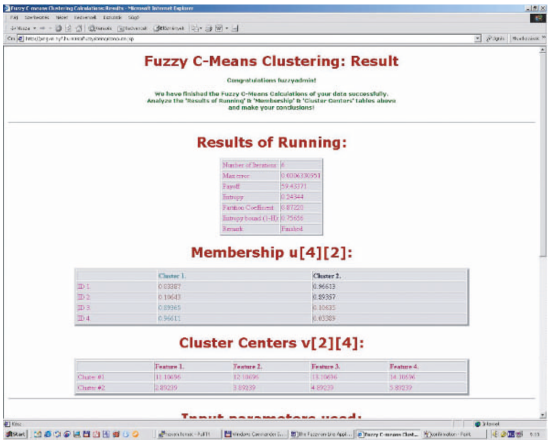


Figure 1. The top hierarchy of the application.

After starting of the Tomcat on the server computer we had got the application file structure under the '/fuzzy' application root directory (Figure 1).

3.3 Running the application

To execute our deployed application you have to use the context*/fuzzy* in the URL. In this case to run the fuzzy clustering Web application you need to write:
http://pingvin.nyf.hu:8080/fuzzy/index.html the URL address on the client side. In the Figure 2, you can see one of the 'faces' of our Web application.



Fuzzy C-Means Clustering: Result

Congratulations fuzzyadmin!

We have finished the Fuzzy C-Means Calculations of your data successfully.
Analyze the 'Results of Running' & 'Membership' & 'Cluster Centers' tables above
and make your conclusions!

Results of Running:

Number of Iterations	6
Max error	0.0006330951
Payoff	59.43371
Entropy	0.24344
Partition Coefficient	0.87220
Entropy bound (1-H)	0.75656
Remark	Finished

Membership u[4][2]:

	Cluster 1.	Cluster 2.
ID 1.	0.03387	0.96613
ID 2.	0.10643	0.89357
ID 3.	0.89365	0.10635
ID 4.	0.96611	0.03389

Cluster Centers v[2][4]:

	Feature 1.	Feature 2.	Feature 3.	Feature 4.
Cluster #1	11.10696	12.10696	13.10696	14.10696
Cluster #2	2.89239	3.89239	4.89239	5.89239

Input parameters used:

Figure 2. Results of the fuzzy clustering.

4. CONCLUSIONS

We managed to develop a modern Web application system by using the JSP technology. The JSP was useful for separating the presentation from the dynamic (business logic) part of the application, creating a scalable, real multi-tier, platform-independent Web application, based on industry standards and freeware software. Additionally, we were able to develop and to serve the system on different platforms being at different towns. From the other aspect the Java technology was available for telework successfully.

REFERENCES

1. L.A. Zadeh (1965), Fuzzy sets. *Information Control*, 8, pp. 338–353.
2. J.C. Bezdek, R. Ehrlich and W. Full (1984), FCM: The fuzzy c-means Clustering Algorithm. *Computer and Geosciences*, 10, 2–3, pp. 191–203.
3. D. Kancsár (1991), *The Fuzzy C-means Clustering Algorithm*. TDK Essay, MGF Nyíregyháza.
4. D. Kancsár and Á.B. Simon (2002), On-line Fuzzy C-means clustering. In: *Proceedings of the Computational Intelligence 3rd International Symposium of Hungarian Researchers,* Budapest, pp. 239–246.

TEXT CLASSIFICATION FOR CHINESE WEB DOCUMENTS

Ming Hu[1,2], Jianchao Xu[1,2] and Liang Hu[2]
[1]*School of Computer Science & Engineering, Changchun University of Technology, Changchun, China, 130012*
[2]*Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, China, 130012*

Abstract Although many methods for feature selection and text classification have been applied to English web documents, relatively few studies have been done on Chinese web documents. This paper introduces a term weighting method based on inverse document frequency, html tags and length of Chinese phrase, provides an algorithm for web text classification based on improving on lattice machine approach. The experiments show this method is effective in feature reduction and text classification.

Keywords: text classification, data-mining, information retrieval, web-mining.

1. INTRODUCTION

Both Web text feature selection and text classification require automatically extracting a set of keywords describing the web text and assigning one or multiple predefined category labels to the web page. A difficult problem with text classification is that there are too many different words occurring in documents to be analysed effectively by learning algorithm, a prior study has proven that the existing feature selection methods have some limitation [1]. Now, extensive ranges of methods have been applied to English text classification [2, 3], relatively few have been benchmarked for Chinese text classification. No approach is known to be the best classifier [3]. Many approaches can assign to fit one to a web page, but some web pages belong to multiple category.

In this paper, we provide an algorithm for web text classification. The idea is as follows: firstly, transform synonyms with a synonym lexicon and compute the weight of terms based on inverse document frequency, html tags and length

G. R. Liu et al. (eds.), Computational Methods, 1171–1175.

of Chinese phrase; secondly, select web text feature based on a messy genetic algorithm; lastly, extend lattice machine approach for assigning web text multiple categories by using these features.

2. EXTENDING LATTICE MACHINE AND TEXT CLASSIFICATION FOR WEB TEXT CLASSIFICATION

Lattice machine [4], is a general framework for supervised learning. In this method, any data-mining algorithms can be regarded as a process of data reduction; data reduction is to find the interior cover of given data in the lattice of universal hyper relations. All data in the interior cover are *equilabelled*. This method has been used to structured data-mining and text classification. Since a web page can belong to multiple categories, we extend this method for web text classification.

We represent web text corpus by the feature vector as in Table 1 where category labels come from "The Method of Classing Chinese Archive" [5]. It can be regarded as a decision system that the cardinality of decision attributes d(Class) is greater than 1. All decision rules in Table 1 are consistent as a text that can be assigned multiple predefined category labels.

In the original lattice machine the partially ordering relation and the least upper bound are given. In the context of web text classification, the documents

Table 1. A feature vector representation of web text corpus.

Doc	Text feature	Class
1	Weather equipment, mapping, earthquake, measure	TB25, TB7, D193
2	Agro technique, aquatic product processing, control flood and improve soil structure, grassland egis, stockbreeding	MA25, MA416, MC424
3	Landform measure, weather equipment, earthquake disaster research, measure	TD3, TB25
4	Aquatic product processing, rural economics, stockbreeding, shelter belt	MA416, JA11111, MA1
5	Democratic parties, congress, mental civilization	A51, A121, FC1212
6	Congress, conference summary, dismiss sb. from his post	A121, FC1212, A1.26
7	Agrotechnique, tending of woods, control flood and improve soil structure, grassland egis	MA25, MB33, MC424
8	Weather equipment, ocean, measure, earthquake	TB25, TB7, HA7145
9	Democratic parties, conference summary, consolidate the party organization and rectify the working style of the party	A51, A46, A1.26

are assigned as the same category labels have large number of common feature elements. We use $\supseteq$ to substitute $\subseteq$ to define the ordering relation and accordingly to substitute $\cap$ for $\cup$. The ordering relation is given as follows:

$$\begin{gathered} doc_1 \leq doc_2 \Leftrightarrow \\ set\_of\_words(doc_1) \supseteq set\_of\_words(doc_2) \end{gathered}$$

where $set\_of\_words(doc_i)$ is the set of features in doc_i, and the least upper bound is given by $doc_1 \vee doc_2 = set\_of\_words\ (doc_1) \cap set\_of\_words(doc_2)$. This amounts to classifying a web text by using a subset of the features in the document. Since the decision entry of a *hyper tuple* is a hyper entry, the concept of *equilabelled* in the original lattice machine is not fit for this kind of text classification, so we extend the concept of decision system as follows.

Definition 1. A decision system for multiple categories $D \leq U, \Omega = C \cup d$, $V_x > x \in \Omega$ is an extending for *information system* as above, where C is a nonempty finite set of condition attributes, d is a decision attribute, $C \cap d = \phi$, V_x is the domain of $x \in \Omega$ and the power set of V_x is P_x,$d : \boldsymbol{R} \longrightarrow P_x$ is a mapping, called a *labelling* of $\boldsymbol{R}$. For all $t \in U$, the *label* of t, $d(t)$, is a hyper entry, denoted all the same by $\{D_1, \ldots, D_K\}$. For all i and j, $D_i \cap D_j$ can be nonempty.

The *decision system for multiple categories* is a natural extending for the *decision system*. if $\forall t \in U$, $\forall x \in \Omega$, $|t(x)| = 1$, it will degenerate into the *decision system*. All elements of dataset $\boldsymbol{R}$ have a label and all elements of V\R have not any label. Since the domain lattice is algebra, we can generalize the concept of *equilabelled* by *intersection labelled* to the elements in the lattice that is covered with $\boldsymbol{R}$ to solve the problem of assigning a label to a new element.

Definition 2. In the domain lattice $\langle T \leq \rangle$, $t \in T$ is labelled with $d(t) \neq \phi$, t is called *intersection labelled* with respect to D_q, if $\forall a_i \in \downarrow t$, $i = 1, \ldots, k$, $d(a_1) \cap d(a_2) \cap \ldots\ldots \cap d(a_K) = d(t) = D_q$. In other words, the intersection of all labelled elements covered by t is a nonempty set, and uses the intersection D_q to label t.

In the domain lattice, the concept of the intersection labelled can be used to label a sublattice. The unlabelled elements in a sublattice can be assumed that have the same labeling as the maximal intersection labelled element in the sublattice as a generalization, so the labelling of the maximal intersection labelled element is contained by labelling of each element in the labelled sublattice. Accordingly, E-set contains all intersection labelled elements that are maximal, and supported by $\boldsymbol{R}$. This generalization can be used to assign a label to a new web page in web text Classification. For document corpus R(Ω,d), where Ω is an attribute denoting the set of features in a document, d is a label attribute. Algorithm, below, discovers all E-set:

Algorithm. Discovery of E-set

```
**Input:** The labelled corpus $R(\Omega,d)$
**Output:** $B$: all E-set
**Method:**
$A \leftarrow R$ □$X \leftarrow R$ □$B = \phi$
While $A \neq \phi$
 {get $a \in A$, $A = A\backslash\{a\}$
  $C = X\backslash\{a\}$, $x = a$
  While $C \neq \phi$
   {get $g \in C$, $C = C\backslash\{g\}$, $s = d(a) \cap d(g)$
    If $s \neq \phi$ and $\Omega(a) \cap \Omega(g) \neq \phi$ then
    {$a = a \vee g$
     for all $p \in C$
If $d(a) \cap d(p) = s$ and $\Omega(a) \cap \Omega(p) \neq \phi$ then $a = a \vee p$
endfor
$B = B\cup(a,s)\}$
 $a = x$
}
}
```

2.1 Experiments

For validating the effect of extended lattice machine, we apply it to those Chinese web corpus in our experiments. We assign one or multiple category labels to the corpus according to The Method of Classing Chinese Archive with Comparison Table of Indexing Descriptors and Classification Labels; then generalization classification rules are discovered with the above algorithm; our experiments show that these rules are validated and this method acts well. Table 2 is the generalization rules from Table 1.

Table 2. The intersection labelled set from Table 1.

From	Text feature	Class
$1\vee3\vee8$	Weather equipment	TB25
$1\vee8$	Weather equipment, measure, earthquake	TB25 □ TB7
$2\vee4$	Aquatic product processing, stockbreeding	MA416
$2\vee7$	Agrotechnique, control flood and improve soil structure, grassland egis	MC424 □ MA25
$5\vee6$	Congress	A121, FC1212
$5\vee9$	Democratic parties	A51
$6\vee9$	Conference summary	A1.26

3. CONCLUSIONS

Feature selection is a NP-complete problem. We introduce term weighting method based on inverse document frequency, html tags and length of Chinese phrase, and report our method to select web text feature based on a messy genetic algorithm. For multi-class multi-label text categorization tasks, we improve lattice machine and provide the concept of *intersection labelled* to extend the concept of *equilabelled*. Our experiments show that the method is efficient. For further research: terms in a document have in coordinate relativity to the category labels of the document, so it should introduce the significance of the term for assigning a label to a document.

REFERENCES

1. W.J. Cohen and Y. Singer (1996), Context-Sensitive Learning Methods for Text Categorization. In: *SIGIR'96:Proc.19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 307–315.
2. C. Apte, F. Damerau and S. Weiss (1994), Automated learning of decision rules for text categorization. *ACM Transactions Information System*, 12, 3, pp. 233–251.
3. Y. Yang and X. Liu (1999), A Re-Examination of Text Categorization Methods. In: *22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'99)*, pp. 42–49.
4. W. Hui, D. Ivo and B. David (1998), Data Reduction Based on Hyper Relation. In: *Proceedings of KDD98, New York*, pp. 349–353.
5. National Archive Office of China (1987), *The Method of Classing Chinese Archive (in Chinese)*. Archive Book Concern, Beijing.

LOGIC-BASED CONSTRAINT HANDLING IN RESOURCE-CONSTRAINED SCHEDULING PROBLEMS

J.Y. Zhang, J.G. Sun and Q.Y. Yang
College of Computer Science & Technology, Jilin University, Changchun, P. R. China 130012

Abstract Resource-constrained scheduling problem is one kind of typical real-life discrete optimization problems, which is one of the strongest application areas of constraint programming. In the constraint programming toolkit 'Mingyue', which embed constraints in the object-oriented language C++, we design a new logic-based method for handling the constraints in the resource-constrained scheduling problem. In this paper, we propose a way of describing those constraints with the discrete-variable logic formula. Based on this model, a resolution algorithm is designed for filtering the discrete variables' domain. Comparisons with other constraint handling approaches and related literature clearly show that our approach can describe the constraints in the high level and solve the resource-constrained scheduling problem in the logic framework.

Keywords: scheduling, constraint handling, logic, discrete optimization.

1. INTRODUCTION

Constraints are powerful tools for handling many real-world problems. Constraint programming (CP) [1] is based on the idea of describing the problem declaratively by means of constraints and, consequently, finding a solution satisfying all the constraints, i.e., assigning a value to each unknown from its respective domain. CP has a big advantage over other frameworks in declarative modeling capabilities. The modeling capabilities of CP are really fascinating and the constraint models are very close to the description of discrete optimization problems. This simplifies the maintenance of the models as well as the introduction of domain dependent heuristics necessary to solve large-scale problems such as Resource-Constrained Scheduling Problem (RCSP), which

G. R. Liu et al. (eds.), Computational Methods, 1177–1186.

is one of the strongest application areas of CP [2]. The reason of such success can be found in a similar character of both scheduling problems and Constraint Satisfaction Problems (CSPs) [3].

Our system 'Mingyue' CB-Scheduler [4, 5] is one kink of CP toolkits, which embed constraints in the object-oriented programming language C++. Other constraint-solving toolkits are also quite popular, such as ILOG solver and scheduler. However, in the typical CP toolkit, constraint handling is to list all the assignment combinations of variables directly in the CSP model. In this paper, we propose handle the constraints in RCSP by using the formula of discrete-variable logic, which is more general and brief.

In this chapter we omit the description of the system 'Mingyue' CB-Scheduler and emphasize the constraint handling method in the system. In Section 2, we first describe the definition of constraint domain and CSPs and then we present a generic CSP model of RCSP. The details about handling constraints in the discrete-variable logic way are showed in Section 3. In Section 4, we give one of the resolution algorithms for filtering the constraint variables' domain in RCSP. Finally, we draw a conclusion on this new method.

2. RESOURCE-CONSTRAINED SCHEDULING CSP MODEL

Partially based upon the scheduling problems we encountered in the industry, we define a generic (and necessarily incomplete) typology of resource-constrained scheduling as follows:

Given are a set of n tasks $T = \{T_1, \ldots, T_n\}$ and a set of *m* resources $R = \{R_1, \ldots, R_m\}$. Each task T_i has its certain time-window $[r_i, d_i]$ (r_i denotes the release-date, d_i denotes the due-date) and needs some amount of resource throughout their execution. Each resource R_i has its certain capacity. The tasks may be interrupted or non-interrupted. There exists temporal constraints C_t and resource constraints C_r between the tasks, which can be binary or n-ary. A solution schedule is a set of integer execution times for each task so that all the temporal and resource constraints are satisfied.

2.1 Constraint domain and CSPs

The legitimate forms of constraint and their meaning is specified by a constraint domain, which specifies the 'syntax' of the constraints. That is, it specifies the rules for creating constraints in the domain. It details the allowed constraints, functions and constraint relations as well as how many arguments are placed. The constraint domain also determines the values that variables can take. Finally, the constraint domain determines the meaning of all these symbols. It determines what will be the result of applying a function to its arguments and whether a constraint relation holds for given arguments.

There are three various constraint domains, including real arithmetic constraints, tree constraints and finite domains [6]. The finite constraint domains are constraint domains in which the possible valued that variable can take are restricted to a finite set, such as Boolean constraints and integer constraints. Finite constraint domains are widely used in CP. Many real-life problems, notably scheduling, routing and timetabling, are simple to express using finite constraint domains, since essentially, they involve choosing amongst a finite number of possibilities. In the artificial intelligence community, satisfaction of constraint problems over finite domains has been studied under the name 'constraint satisfaction problems'.

The finite domains constraint satisfaction problems consists of the following: Given are a set of n variables $Z = \{X_1, \ldots, X_n\}$ with discrete, finite domains $D = \{D_1, \ldots, D_n\}$ and a set of *m* constraints $C = \{c_1, \ldots, c_m\}$ that are predicates ck $(X_i , \ldots, X_j)$ defined on the Cartesian product $D_i \times \ldots \times D_j$. If c_k is true, the valuation of the variables is said to be consistent with respect to ck, or equivalently, ck is satisfied. A solution is an assignment of a value to each variable, from its respective domain, such that all constraints are satisfied.

2.2 Task and resource

A RCSP can be encoded efficiently as a CSP: two variables, st_i and ft_i, are associated with each task T_i; they represent the start time and the finish time of T_i. The smallest values in the domains of st_i and ft_i are called the release-date and the earliest finish time of T_i (r_i and eft_i). Similarly, the greatest values in the domains of st_i and ft_i are called the latest start time and the due-date of T_i (lst_i and d_i). The processing time of the task is an additional variable pt_i, that is constrained to be lower than or equal to the difference between the end and the start times of the task (most often, processing time is known and bound to a value dur_i).

No loss of generality, the scheduling problem we consider is the preemptive scheduling problem (tasks can be interrupted at any time). One can either associate a set variable (i.e., a variable the value of which will be a set) $set(T_i)$ with each task T_i, or define a 0–1 variable $W(T_i, t)$ for each task T_i and time t; $set(T_i)$ represents the set of times at which T_i executes, while $W(T_i, t)$ assumes value 1 if and only if T_i executes at time t. Ignoring implementation details, let us note that:

- the value of $W(T_i, t)$ is 1 if and only if t belongs to $set(T_i)$.
- st_i and ft_i can be defined, in the preemptive case, by $st_i = \min_{t \in set(T_i)}(t)$ and $ft_i = \max_{t \in set(T_i)}(t + 1)$; such variables are often needed to connect tasks

together by temporal constraints. Notice that in the non-preemptive case, $\mathrm{set}(T_i) = [st_i, ft_i)$, with the interval $[st_i, ft_i)$ closed on the left and open on the right so that $|\mathrm{set}(T_i)| = ft_i - st_i = pt_i$.

In the following, capacity (R_i) denotes the constant used to represent the initial capacity of the resource R_i. Based on the value of capacity (R_i), unary resource (capacity $(R_i) = 1$) and discrete resource (capacity $(R_i) > 1$) are distinguished.

2.3 Constraints in RCSP

The constraints appearing in RCSP can be classified into several groups. We classify the constraints using their role in the RCSP into two categories: temporal and resource constraints.

- Temporal constraint (C_t) can be expressed by linear binary constraints between the start and finish variables of tasks. By analysing all C_t, we can construct the temporal constraint network. Note that time-windows $[r_i, d_i]$ of task is special case of temporal constraint.
- Resource constraint (C_r) defines how a given task T will require and affect the availability of a given resource R. It consists of a tuple $\{T_i, R_i, cap\}$ where cap is an integer decision variable or constant defining the quantity of resource R_i required by task T_i.

We introduce constant cap(T, R) to denote the quantity of resource R required by task T. We name this kind of task T C-task (Constant Task). We introduce variable cap (T, R, t) $(\mathrm{cap}(T, R, t) \leq \mathrm{capacity}(R))$ to denote the quantity of resource R required by task T at time-point t $(r_i \leq t \leq d_i)$. We name this kind of task T V-task (Variable Task). Let us consider a set K of tasks, subjected to both release-dates and due-dates, to be sequenced on resource R (unary or discrete) with initial capacity capacity (R) in a RCSP. |K| denotes the number of tasks in set K. energy(R, t) is a variable that represents the availability quantity of the resource R at time t $(r_K \leq t \leq d_K,\ r_K = \{\mathrm{minimum}\ (r_i) \mid T_i \in K\},\ d_K = \{\mathrm{maximum}\ (d_i) \mid T_i \in K\})$. The initial value of energy(R,t) is capacity (R).

3. LOGIC-BASED CONSTRAINT MODELING IN RCSP

Based on the CSP model of RCSP, the general method of modeling constraints in it is to list all the assignment combinations of variables directly. We propose describe the constraint by using the formula of discrete-variable logic. In the framework of discrete-variable logic, the scheduling problem can be solved in the method of multivalent resolution on finite sets.

3.1 Discrete time modeling in RCSP

Many discrete optimization problems are readily expressed using variables that assume a finite number of discrete values. The clearest example is scheduling problem in which tasks or some other set of objects must be arranged in a sequence. The RCSP falls into this class. In the RCSP, the domains of the time variables (st_i and ft_i) are really finite and discrete, which is caused by the discrete time modeling in RCSP.

Task in scheduling deals with time as one of the parameters. Therefore modeling of time is necessary in most scheduling applications. In RCSP, a common modeling situation requires that resource consumption be monitored continuously over time. For example, it may be possible to process several tasks simultaneously, provided that their consumption of resources at no point exceeds a maximum. This is difficult to model using 0–1 variables because one must keep track of which tasks are still in process at any given moment. This is usually done by discretizing time. Our way of modeling time is to divide the time-line into a sequence of discrete time intervals with the same duration. We call such intervals time slices. The duration of the time slice defines the resolution of the schedule.

It is expected that the behaviour of the resource within the time slice is homogenous, i.e., the important events like the change of task appear at the time point between two time slices only. Consequently, if we model tasks using discrete time then the duration of the time slice must respect the duration of all the tasks. More precisely, the duration of the time slice must be a common divisor of the duration of all tasks. If we work with tasks that have no restrictions about their start and finish time then this requirement may lead to a huge number of time slices (high resolution) even if the duration of the tasks is long (low resolution). In the discrete time model the variables describe the situation either at the time points or at the time slices.

3.2 Formulas of discrete-variable logic

We name the time variables x_i in RCSP multivalent variables, which have the finite and discrete domains. An elementary extension of propositional logic can be developed for multivalent variables. In propositional logic, the primitive unanalysed terms are atomic formulas yj. The analysis can be carried slightly deeper by supposing that atomic propositions are themselves predicates that say something about discrete variables $x_1, \ldots, x_n$, which can be regarded as the time variable in RCSP. For instance, a predicate may have the values x_j can assume. Special cases would be $x_j = v$ and $x_j \neq v$, where v is a constant. A number of useful predicates can be defined in terms of more primitive notation, just as equivalence $\equiv$ and implication $\supset$ are defined in terms of $\vee$ and $\neg$ in propositional logic.

The resulting logic is still bivalent in that propositions have one of two truth-values. The variables, however, are multivalent.

Whereas a limited repertory of connectives appear to be useful in propositional logic, multivalued variables multiply the possibilities. The all-different, element, distribute, and cumulative predicates have proved especially useful. The idea of a logical clause is also readily generalized.

4. FORMULAS AND SEMANTICS

The atomic propositions y_j of propositional logic are replaced with predicates $P(x) = P(x_1, \ldots, x_n)$ in discrete variable logic. Predicates can be combined with logical connectives in the same way as logical propositions. one primitive predicate will be sufficient to define all others, namely $P(x) = (x_j \in X_j)$ for $X \subset D_j$.

The semantics are slightly different than in propositional logic. In the latter, the meaning of a molecular formula is given by the Boolean function it represents. In discrete logic, a formula's meaning is given by a truth function f(x) of the discrete variables $x = (x_1, \ldots, x_n)$, where each $x_j \in D_j$. In particular, each predicate is defined by the function f(x) it represents. For example, the function f(x) for $x_j \in Y_j$ takes the value 1 if the value assigned x_j belongs to X. Once the truth-values of the predicates are determined, the truth-values of the formulas containing them are computed in the normal propositional way.

5. MULTIVALENT CLAUSES

Multivalent clauses are a straightforward generalization of propositional clauses and are completely expressive in an analogous sense.

A multivalent clause has the form

$$\bigvee_{j=1}^{m} (x_j \in X_j) \tag{1}$$

where each $X_j \subset D_j$. If X_j is empty, the term $(x_j \in X_j)$ can be omitted from (2), but it is convenient to suppose here that (2) contains a term for each j. If

$X_j = D_j$ for some j, then (2) is a tautology. Note that the literals of a multivalent clause contain no negations. This brings no loss of generality, since $\neg(x_j \in X_j)$ can be written $x_j \in D_j \backslash X_j$.

Any truth function f(x) = f($x_1, \ldots, x_n$) can be expressed as a conjunction of multivalent clauses. This is done simply by ruling out the values of y for which f(x) = 0. Thus, if f(x) = 0 for x = $v_1, \ldots, v_k$, then f(x) is represented by the formula

$$\mathop{\wedge}_{i=1}^{k} \mathop{\wedge}_{j=1}^{n} \left(x_j \neq v_j^i\right) \tag{2}$$

which can be formally written as a multivalent clause:

$$\mathop{\wedge}_{i=1}^{k} \mathop{\wedge}_{j=1}^{n} \left(x_j \in D_j \backslash \{v_j^i\}\right) \tag{3}$$

Because any constraint over finite domains represents such a function f(y), it is equivalent to a finite set of multivalent clauses.

One multivalent clause $\vee_j (x_j \in X_{1j})$ implies another $\vee_j (x_j \in X_{2j})$ if and only if the one absorbs the other; that is, $X_{1j} \subset X_{2j}$ for each j. Equivalent multivalent clauses are identical. Prime implications are defined precisely as for classical clauses.

Any formula of discrete logic can be converted to a conjunction of multivalent clauses by using De Morgan's laws, distribution, double negation, and the fact that $\neg(x_j \in X_j)$ means $(x_j \in D_j \backslash X_j)$.

6. MULTIVALENT RESOLUTION

Resolution is easily extended to the logic of discrete variables. Unit resolution also has an analog. Resolution plays the same role in computing projections as it does in propositional logic. The resolution algorithm for multivalent clauses is related to Cooper's algorithm for obtaining k-consistency for a set of constraints [7]. Given a set of multivalent clauses,

$$\left\{ \mathop{\vee}_{j=1}^{n} (x_j \in X_{ij}) \,|\, i \in I \right\} \tag{4}$$

the resolvent on x_k of these clauses is

$$\left(x_k \in \bigcap_{i \in I} X_{ik} \right) \vee \mathop{\vee}_{j \neq k} \left(x_j \in \bigcup_{i \in I} X_{ij} \right) \tag{5}$$

Ordinary bivalent resolution is a special case. The clauses in the set (4) are the parents of the resolvent (5).

A unit resolution algorithm for multivalent clauses

```
Let S be a set {$C_i$ | $i \in I$} of multivalent clauses,
  where each Ci has the form $\bigvee_{j=1}^{n} (x_j \in X_{ij})$
For all $i \in I$ and j=1,…,n let $X_{ij}=X_{ij} \cap D_{xj}$,
  where $D_{xj}$ is the current domain of $x_j$.
For $i \in I$:
  Let $\chi_i$ be the collection of nonempty sets $X_{ij}$.
  If $\chi_i=\varnothing$ then stop; S is unsatisfiable.
For j=1,…,n let $S_j=\{C_i \mid X_{ij} \in \chi_i\}$.
Let U be the set of pairs($C_k$, $j_k$) for which $C_k$ is a unit clauses;
that is , $\chi_k=\{ X_{kj_k} \}$.
While U is nonempty:
  Remove some pair ($C_k$, $j_k$) from U.
  Remove $C_k$ from $S_{j_k}$.
  For all $C_i \in S_{j_k}$ :
   Let $X_{ij_k} = X_{ij_k} \cap X_{kj_k}$ .
   If $X_{ij_k}=\varnothing$ then:
     Remove $X_{ij_k}$ from $\chi_i$.
    If $\chi_i =\varnothing$ then stop; S is unsatisfiable.
    If $|\chi_i|=1$ then add $C_i$ to U.
```

Figure 1. A unit resolution algorithm for multivalent clauses.

A unit resolution algorithm for multivalent clauses appears in Figure 1. It is very similar to classical unit resolution. When a clause Ck becomes a unit clause $x_j \in X_{kj}$, the domain of x_j can contain only elements that occur in X_{kj}. The clause C_i is deleted and the clauses containing x_j adjusted accordingly.

To speed processing, the algorithm keeps track of which literals remain in each clause. Thus χ_i contains the nonempty sets X_{ij}. It also maintains a list S_j of the clauses that still contain x_j; that is, the clauses C_i for which X_{ij} is nonempty. Whenever X_{kj} in some clause Ck becomes empty, X_{kj} is removed from χ_k. If this makes C_k a unit clause, then every X_{ij} in the constraint set must be updated so that it lies in X_{kj}. The list S_j makes it possible to locate quickly the X_{ij}'s that might be affected. The clause Ck is deleted from the problem and removed from S_j.

Example. The three clauses below

$$\begin{aligned} &(x_1 \in \{1, 4\}) \vee (x_2 \in \{1\}) \\ &(x_1 \in \{2, 4\}) \vee (x_2 \in \{2, 3\}) \\ &(x_1 \in \{3, 4\}) \vee (x_2 \in \{1\}) \end{aligned} \tag{6}$$

resolution of the first and third clauses of (6) on x_1 to yields

$$(x_1 = 4) \vee (x_2 = 1) \tag{7}$$

This clause absorbs three other clauses in (6), resulting in the clause set

$$\begin{aligned} &(x_1 \in \{2, 4\}) \vee (x_2 \in \{2, 3\}) \\ &(x_1 = 4) \vee (x_2 = 1) \end{aligned} \tag{8}$$

Resolution of these on x_2 produces

$$x_1 \in \{2, 4\} \tag{9}$$

This and (7) are the prime implications of (6).

7. CONCLUSION

Our work focuses on handling the constraints in the logic way. In the discrete-variable logic, any constraint over finite domains in RCSP can be represented as a finite set of multivalent clauses. Using the discrete-variable formula to describe constraints can heighten the modeling capabilities of CP. Based on this model, we designed a resolution algorithm in the form of constraint-based search in the logic framework. In sum, logic-based modeling not only heightens the problem modeling capability but also exploits the problem solving method. Scheduling is really a process of getting the constraints right. However, designing a constraint model that can be used to solve real-life large-scale problems is also the biggest challenge of current Constraint Programming.

ACKNOWLEDGEMENT

This paper is supported by the National Natural Science Foundation Key Project of China (Grand No. 60073039, 60273080).

REFERENCES

1. R. Barták (1998), On-line Guide to Constraint Programming. Charles University, Prague, http://kti.mff.cuni.cz/~bartak/constraints/.
2. M. Wallace (1994), Applying constraints for scheduling in constraint programming. In: *NATO ASI Series*, B. Mayoh and J. Penjaak (eds.), Springer Verlag.
3. E.P.K. Tsang (1993), *Foundations of Constraint Satisfaction*. San Diego, Calif.: Academic, pp. 53–63.
4. J.Y. Zhang, X. Li and J.G. Sun (2003), Research on constraint-based scheduling and its implementation. In: *Proceeding of CNCC'03*, Beijing, P.R. China, Tsinghua University Press, pp. 80–85.
5. J.G. Sun and J.Y. Zhang (2004), A generic mechanism for managing resource constraints in preemptive and non-preemptive scheduling. *Processing of SCI'04 Conference*, Orlando, USA.
6. K. Marriott and P.J. Stuckey (1998), *Programming with Constraints: An Introduction*. MIT Press, pp. 133–134.
7. M.C. Copper (1989), An optimal k-consistency algorithm. *Artificial Intelligence*, 41, pp. 89–95.

PARTICLE SWARM OPTIMIZATION METHOD USED IN PIXEL-BASED TEXTURE SYNTHESIS

Y. Meng, W.H. Li, W. Guo and Y.L. Liu
Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, College of Compute Science & Technology JiLin University, Changchun, 130012, China

Abstract In this paper, we suggest to use particle swarm optimization (PSO) in pixel-based texture synthesis. It can search better matching pixels from the sample texture and paste them onto the result texture. High-quality textures can be generated by this method in short time, especially in dealing with narrow textures and arbitrary-shaped images that out of the region of patch-based methods theoretically. The stimulate experiment results shows its efficiency and practicability.

Keywords: pixel-based texture synthesis, patch-based texture synthesis, particle swarm optimization (PSO), L-shaped neighborhood.

1. INTRODUCTION

Many applications of computer graphics, vision and image processing can benefit from texture synthesis algorithm, because textures can describe a wide variety of natural phenomena. In 1999, Efros and Leung [1] developed a very good method by growing texture using non-parametric sampling based on Markov Random Field (MRF). And most recent works on texture synthesis are all based on it. Wei and Levoy [2] improved it by using a multi-resolution image pyramid based on a hierarchical statistical method. Methods mentioned above are all pixel-based methods.

In recent years, patch-based algorithms were presented. Ashikhmin [3] made an intermediate step towards copying patches. Liang, Liu, Xu, Guo and Shum [4] and Efros and Freeman [5] copy patches of input texture at a time. Praun, Finkelstein and Hoppe [6] present a similar approach to create textures covering a 3D model.

G. R. Liu et al. (eds), Computational Methods, 1187–1192.

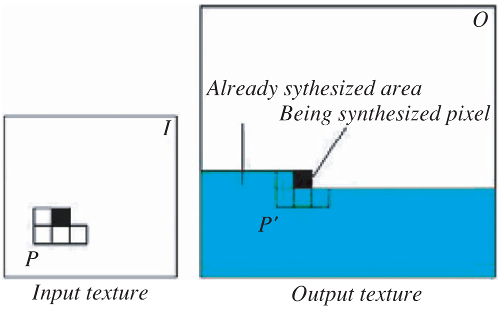


Figure 1. Pixel-based texture synthesis.

2. OUR APPROACH

2.1 Particle Swarm Optimization (PSO)

J. Kennedy and R. C. Eberhart [7] proposed the original idea of PSO. The particle uses the following formulae to update its location and velocity:

$$V(t+1) = V(t) + rand() \times c1 \times (pBest(t) - present(t)) + rand() \times c2 \times (gBest(t) - present(t)) \quad (1)$$

$$present(t+1) = present(t) + V(t+1) \quad (2)$$

2.2 Pixel-based Texture Synthesis

This method is a Markov Random Field based method and it synthesizes new texture one pixel at a time. Figure 1 gives the process of this method: given a sample texture image I, a new image O is being synthesized one pixel at a time. To synthesize a pixel, the algorithm first finds all neighbourhoods N in the sample image that are similar to the pixel's neighbourhood N' in B and then randomly chooses one neighbourhood and takes its center to be the newly synthesized pixel [1].

2.3 Pixel-based Texture Synthesis using PSO

Patch-based methods improved the synthesis speed, but theoretically it has a serious limitation that it can be only used to synthesize broader and rectangle sample texture images. So pixel-based methods is more appropriate to be used to synthesize arbitrary-shaped and narrower sample texture image. We present a pixel-based texture synthesis method by using PSO algorithm to search for an approximate best pixel from a sample image to a result image. This method is based on Efros and Leung's [1] method and its speed is almost comparable with patch-based methods'.

Now we apply the PSO algorithm to pixel-based texture synthesis. First randomly set a number of pixels in input image I and treat them as particles. Each particle determines an *L-shaped* neighbourhood. User can define size of neighbourhood. When the particles travel through the image we compare their neighbourhoods with the being synthesized pixel's and find an approximate best match pixel. Details about how to apply PSO are as below:

All particles have their locations and velocities property. As particles travelling through I, we calculate the *Fitness Function* of each location they pass through. We denote the fitness as the distance between the neighbourhoods of particle and the being synthesized pixel's. Define D as neighbourhood of just being synthesized pixel and S as the neighbourhood of the current particle's location. The formula for calculating fitness between two neighbourhoods D and S is as follows:

$$d(D, S) = \sqrt{\sum_{i=1}^{k}\{[r(D_i) - r(S_i)]^2 + [g(D_i) - g(S_i)]^2 + [b(D_i) - b(S_i)]^2\}} \tag{3}$$

Where $r()$, $g()$, $b()$ are the RGB value of a pixel. D_i means the *i-th* pixel in D, S_i means the *i-th* pixel in S and k means the amount of pixels in the neighbourhoods. We set n particles on the sample texture. Denote P_i as the location of particle i and L_i as the best fitness location of particle i, with best fitness B_i. The global fitness g is the best fitness among all n particles so $g = min(B_1, B_2, \ldots B_n)$. The location of the particle, which has best global fitness, is recorded as G. Each particle will be given an initial velocity vector V_i.

Now we update the particles using the following formulae:

$$\begin{aligned} V_i &= V_i + rand() \times c1 \times (L_i - P_i) \\ &\quad + rand() \times c2 \times (G - P_i), \mathrm{P}_i = \mathrm{P}_i + \mathrm{V}_i \end{aligned} \tag{4}$$

In search process, a *threshold value* function *dtv* should be defined to determine the location. If the difference between neighborhoods of a pixel (particle) and the being synthesized pixel's is less than *dtv* then we get the location, stop the iteration and paste the pixel onto the output texture.

$$dtv = \sqrt{\lambda \times \sum_{i=1}^{k}\left[r(D_i)^2 + g(D_i)^2 + b(D_i)^2\right]} \tag{5}$$

Where k is the number of pixels in the neighbourhood and D_i represents the value of the i-th pixel in the neighbourhood of the just being synthesized pixel. λ is the error tolerance and it is set to 0.2 in our algorithm, this value is most suitable for keeping the randomness while ensuring synthesis quality and avoiding repetition.

The algorithm can be processed into two steps:

First, randomly set value of each pixel in output texture image O, start synthesize at the lower left corner of O. For the pixels on the edge of O that their neighbourhoods are smaller than the size of user defined, we should make up it with the boundary pixels on the other edge of image.

And then, if PSO algorithm found a location with fitness less than dtv then we paste the pixel value at the location to O. Otherwise we terminate the iteration when it exceeds the maximum iteration number and have found the best solution so far. Blend the corresponding pixel and paste it to O; repeat this step until O is filled.

3. RESULTS OF ALGORITHMS AND PERFORMANCES COMPARISON

3.1 Results of Algorithms

Figure 2 gives a group of the synthesis results using the algorithm in this paper and other texture synthesis methods. (a) is the sample image, (b) is result of Efros and Leung's [1] method and (c) is generated by Liang's method. We

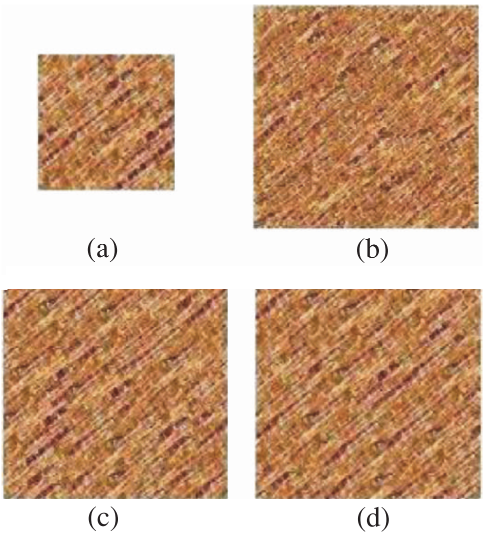

(a) (b)
(c) (d)

Figure 2. Results of several algorithms.

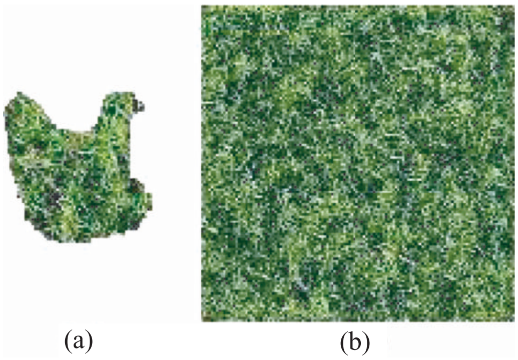

(a) (b)

Figure 3. Synthesis result of an arbitrary-shaped sample.

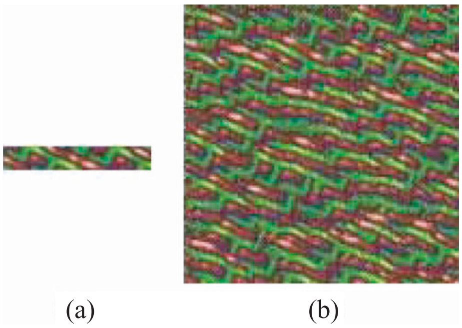

(a) (b)

Figure 4. Synthesis result of a narrow image.

use 40 particles, and obtain result (d). Size of the input sample is 64 × 64. Sizes of the output images are all 128 × 128. From visual inspection, our method is as good as the other two methods. Figures 3 and 4 respectively give synthesis results of an arbitrary-shaped and a narrow image by using our algorithm, which can't be processed by patch-based methods theoretically.

3.2 Performance Comparison

We used several existed methods and ours to synthesize sample textures with size 64 × 64 to new textures with size 128 × 128 and get computing time of them, which is listed in Table 1. They are performed on a DELL PC with 2G CPU. It is clear that our method using 40 particles performs much faster than Efros and Leung's [1]; it is almost comparable with Efros and Freeman's [5].

Table 1. Performances comparison.

Method	Synthesis Time(seconds)
Efros & Leung (1999)	959
Efros & Freeman (2001)	9
our method	14

REFERENCES

1. A.A. Efros and T. Leung (1999), Texture synthesis by non-parametric sampling. *International Conference on Computer Vision*, pp. 1033–1038.
2. L.-Y. Wei and M. Levoy (2000), Fast texture synthesis using tree-structured vector quantization. In: *Proceedings of SIGGRAPH 2000*, pp. 479–488. ISBN 1-58113-208-5.
3. M. Ashikhmin (2001), Synthesizing natural textures. *2001 ACMSymposium on Interactive 3D Graphics*, pp. 217–226. ISBN 1-58113-292-1.
4. L., Liang, C., Liu, Y.-Q., Xu, B. Guo and H.-Y. Shum (2001), Real-time texture synthesis by patch-based sampling. *ACMTransactions on Graphics* 20, 3: 127–150.
5. A. A. Efros and W. T. Freeman (2001), Image quilting for texture synthesis and transfer In: *Proceedings of SIGGRAPH 2001*, pp. 341–346. ISBN 1-58113-292-1.
6. E., Praun, A. Finkelstein, A. and H. Hoppe (2000), Lapped textures. In: *Proceedings of ACM SIGGRAPH 2000, Computer Graphics Proceedings, Annual Conference Series*, ACM Press/ACM SIGGRAPH/Addison Wesley Longman, pp. 465–470. ISBN 1-58113-208-5.
7. J. Kennedy and R.C. Eberhart (1995), Particle swarm optimization. In: *Proceding of IEEE international Conference on Neural Networks*, IEEE, Vol. IV, pp. 1942–1948.

FORMAL METHOD IN IMPLEMENTATION OF ATLAS LANGUAGE*

Guo De-Gui and Liu Lei
Computer Science and Technology College, Jilin University, Changchun, 130012, P.R. China

Abstract ATLAS is a standard test language, which is extensively used in military and electronic tests. In the implementation of ATLAS, we transformed it into C++, and this paper presents the formal methods of the transformation from the core statements of ATLAS into C++ program. The non-signal statements are directly transformed into their semantic equivalent in C++. Meanwhile, other statements, which describe the concrete test and communication process, including the single-action statements, the multiple-action statements and the bus statements, have no equivalent signal-oriented statements in C++. Thus, the single-action statements and the bus statements are transformed into a series of actions that mainly include codes for allocating device, calling drivers and maintaining device states. And we transform each multiple-action signal statement into a sequence of single-action statements, since they are equal in function.

Keywords: ATLAS, test language, program transformation, formal method.

1. INTRODUCTION

A program transformation is a meaning-preserving mapping defined on a programming language [1]. The program transformation methodology provides thinking tools for the development of programs from specifications [2] and program verification [3]. The program transformation techniques provide mechanical tools for program optimization [4, 5], software customization [6] and compilation [7].

ATLAS (Abbreviated Test Language for All Systems) is designed to describe tests in terms that are independent of any specific test system, and has

---

*This research is funded by the Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education of P.R.China.

G. R. Liu et al. (eds.), Computational Methods, 1193–1202.

been constrained to ensure that it can be implemented on ATE (Automatic Test Equipment) [8]. This language is designed to describe tests in terms that are independent of any specific test system, and has been constrained to ensure that it can be implemented on ATE (Automatic Test Equipment). This language has some traits when it is compared with traditional programming languages as follows. (1) Device independency, namely only test requirement will appear instead of any devices in the ATLAS programs written by users. (2) Signal correlation, programs written in ATLAS are all faced to signal. (3) Expansibility, ATLAS provides an extensibility mechanism that allows users to introduce new nouns, noun modifiers and pin descriptors as required accommodating new signal types. (4) Parallelism and timing function, some events appearing in bus statements must occur concurrently, and some test statements should be started at some given time. (5) Syntax is closed to natural language, and the restriction of the grammar is not so strict.

Since the language is very complex, it is difficult to design its compiler directly. Thus, we transformed the language into C++ in the implementation of ATLAS. However, C++ language has no statements to do the operation on bus and signal. Thus the signal and bus statements of ATLAS should be transformed into a series of actions that mainly include codes for allocating device, calling drivers and maintaining device states. According to the characteristics of ATLAS, this paper presents the formal methods of its transformation from ATLAS to C++. This formalization lays the theoretic foundation for the implementation of ATLAS [9].

The remainder of this paper is organized as follows. In Section 2, we provide the abstract syntax definition of the language discussed in this paper. Section 3 is devoted to the formal methods of the transformation from the test language ATLAS to C++. We also give a summary of this paper in section 4.

2. LANGUAGE DEFINITION

ATLAS not only owns the control structure of traditional programming language, but also defines a complete set of signal and bus statements related to test. This leads to the huge complexity of ATLAS. In the following parts, we will only study the selected core statements, which can perform most of the test process.

2.1 Abstract syntax

```
ATLAS ::= DECL STML
STML ::= STM | STML $ STM
STM ::= SETUP | CONNECT | DISCONN | ARM | FETCH | CHANGE |
        RESET | APPLY | ENABLE | DISABLE | DO | ...
```

We will separately provide the abstract syntax of each discussed statement in the following formal specifications.

2.2 Definition of transformation function

```
  Tp:
  DEC*×STM* →CDEC*×CSTM*
  Td: DEC→CDEC*
Ts: STM→CSTM*
Te: EXP→CEXP
```

Where Tp is the function that transforms an ATLAS program into C++ whose specific work is implemented by Td, Ts and Te; Td transforms ATLAS declarations into C++ declarations; Ts transforms ATLAS statements into C++ statements and Te transforms ATLAS expressions into C++ expressions. The notation ⊕ in the following transforming specifications denotes concatenation of strings and ⊙ denotes the operation of adding an element into a list.

3. TRANSFORMATION METHODS

Converting the main program structure into the main() function of C++ as follows:

```
  Tp 【DECL STML】 = main() {Tdl【DECL】; Tsl【STML】}
  Tdl【DECL】 = (DECL=null →"", let dec ⊙ DL=DECL in Td【dec】 ⊕ Tdl【DI】)
  Tsl【STML】 = (STML=null→"", let stm ⊙ SL=STML in Ts【stm】 ⊕ Tsl【SL】)
```

The non-signal statements of ATLAS have their equivalent in C++, thus their transformation is only the structural equivalent mapping. Due to the space limitation, only the transformation rules of signal and bus statements are provided in the following sections.

3.1 Single-action statements

Single-action signal statements relate to the commuting signal in the test process and there are no statements in C++ performing the similar function. Thus, the transformed statements should include the codes of device allocation, calling driver and state transition.

Simply stated, device allocation means finding a suitable device in ATE system to test the UUT (Unit Under Test). In the following transforming descriptions, the predefined function DevAlloc() is used to show the allocating process. In the transformed C++ program, sequences of device drivers are combined to perform the function of a single-action signal statement of ATLAS.

And, with the execution of the ATLAS program, the state of the active device in the ATE system will change and this lead to different available operations for the device. Therefore, the state transition of active device should be taken into consideration.

3.1.1 SETUP statement

The SETUP statement is used to allocate a virtual resource and to control them so as to generate or receive a defined signal. The permitted SETUP path for sources goes through the CONNECTED state to the PREPARED state, while the permitted path for loads or sensors goes through the UNALLOCATED state to the SET state.

```
Ts【SETUP, [(MCML),] NOUN, SCL, CONN】 =
  let device = DevAlloc(NOUN, SCL, CONN) in
  let mcmFunclist = Tfunc【device, [MCML]】 in
  let stmcFunclist = Tfunc【device, SCL】 in
  let StateStm = Tstate【device】 in mcmFunclist ⊕ stmcFunclist ⊕
    StateStm
Tfunc【device, list】 = let (id, devinfo)=device in case devinfo of
  (SOURCE, CONNECTED):
  (SENSOR/LOAD, UNALLOCATED): Tlist【id, list】
  other: "" end
Tlist【id, list】 = (list=null→"",
    let elem ⊙ list1=list in SetupDrive(id, Te【elem】) ⊕ Tlist【id,
      list1】)
Tstate【device】 = let (id, devinfo)=device in let (kind, state)=
  devinfo in case devinfo of
  (SOURCE, CONNECTED): state= PREPARED
  (SENSOR/LOAD, UNALLOCATED): state= SET
  other: "" end
```

Where SetupDrive() is the device driver of SETUP.

3.1.2 CONNECT statement

The CONNECT statement fastens or joins together the interface points of the UUT and the interface points of the source, sensor or load device and to initiate the signal between the virtual resource and the UUT. The permitted CONNECT path for sources goes through the UNALLOCATED or CONNECTED state; while the permitted CONNECT path for loads goes through the PREPARED or SET state. The path for sensors is similar to loads' except that the sensors can only go through the SET state.

```
Ts【CONNECT, [(MCML),】 NOUN, [SCL,] CONN]=
  let (id, devinfo) = DevAlloc(NOUN, [SCL,] CONN) in
```

```
let (kind, state)=devinfo in case devinfo of
(SOURCE, UNALLOCATED/CONNECTED):
ConnectDrive(id, CONN) ⊕ state = CONNECTED
(LOAD, SET/PREPARED):
ConnectDrive(id, CONN) ⊕ state = PREPARED
(SENSOR, SET): ConnectDrive(id, CONN) ⊕ state = PREPARED
other: "" end
```

Where ConnectDrive() is the device driver of CONNECT.

3.1.3 DISCONNECT statement

The DISCONNECT function is to disconnect the interface points of the UUT and the interface points of the source, sensor or load device. Multiple connections of a source or a load resource to separate UUT interface points are permitted. Thus, DISCONNECT actions will keep the load resource in the PRAPARED state unless all UUT interface connections of the resource are disconnected – in which case the state changes to the SET state. Likewise, the source resource will stay in the CONNECTED state, where multiple connections could have been made until the last connection is opened – at which time the state changes to the UNALLOCATED state. IsEntirelyDisconn() in the following description is a predefined function which decides whether all UUT connections are disconnected.

```
Ts【DISCONNECT, [(MCML),] NOUN, [SCL,] CONN】 =
  let (id, devinfo)=DevAlloc(NOUN, [SCL,] CONN) in
  let (kind, state)=devinfo in
  case devinfo of
    (SOURCE, CONNECTED):
      let bRet=IsEntirelyDisconn(id, CONN) in
      DisconnDrive(id, CONN) ⊕
      (bRet=true→state=UNALLOCATED, state=CONNECTED)
  (SENSOR, PREPARED):
    let bRet=IsEntirelyDisconn(id, CONN) in
    DisconnDrive(id, CONN) ⊕
    (bRet=true→state=SET, state=PRAPARED)
  (LOAD, PREPARED): DisconnDrive(id, CONN) ⊕ state=SET
  other: ""
end
```

Where DisconnDrive() is the device driver of DISCONNECT.

3.1.4 ARM statement

The ARM statement is used only with sensor type resource and induces a measurement cycle by a sensor during which the results of the measurement will be determined before transfer by the associated FETCH statement. The state of active sensor changes from PRAPARED to MEASURED with the execution of ARM statement.

```
Ts【ARM, (MCML), NOUN, SCL, CONN】 =
  let (id, devinfo)=DevAlloc(NOUN, SCL, CONN) in
  let (kind, state)=devinfo in
  case devinfo of
      (SENSOR, PREPARED): ArmDrive(id) ⊕ state=MEASURED
      other: ""
  end
```

Where ArmDrive() is the device driver of ARM.

3.1.5 FETCH statement

The FETCH statement function is to wait for the completion of a measurement cycle induced by an ARM statement and to transfer the measured value(s) into any id in the MC branch. This statement is available for sensor type resource.

```
Ts【FETCH, (MCL), NOUN, SCL, CONN】 =
  let (id, kind, state)=DevAlloc(NOUN, SCL, CONN) in
  case (kind, state) of
    (SENSOR, MEASURED): Tmcl【id, MCL】 ⊕ state=PRAPARED
    other: ""
  end
Tmcl【id, MCL】 = let mc ⊙ MCL1=MCL in let (mcm, var)=mc in case MCL1
of
    null: FetchDrive(id, &var)
    other: FetchDrive(id, &var) ⊕ Tmcl【id, MCL1】
end
```

Where FetchDrive() is the device driver of FETCH.

3.1.6 CHANGE statement

The CHANGE statement is used to alter some of the characteristics of a test requirement that are active in the testing of a UUT. This statement causes a virtual resource (source, load or sensor) to remain in the same state.

```
Ts【CHANGE, [(MCML),] NOUN, SCL, CONN】 =
  let (id, kind, state)=DevAlloc(NOUN, SCL, CONN) in
  (state=PRAPARED→Tscl【id, SCL】 , "")
Tscl【id, SCL】 = (SCL=null→"",
    let sc ⊙ SCL1=SCL in ChangeDrive(id, sc) ⊕ Tscl【id, SCL1】
```

Where ChangeDrive() is the device driver of CHANGE.

3.1.7 RESET statement

The RESET statement causes a signal to return to its quiescent state and to make its virtual resources available for re-use in a SETUP or multiple-action signal statement. The permitted RESET path for sources goes through the PRAPARED state to the CONNECTED state, while the permitted path for loads or sensors goes through the SET state to the UNALLOCATED state.

```
Ts【RESET, [(MCLM),] NOUN, SCL, CONN】 =
  let (id, devinfo)=DevAlloc(NOUN, SCL, CONN) in
  let (kind, state)=devinfo in
  case devinfo in
    (SOURCE, PRAPARED): ResetDrive(id) ⊕ state=CONNECTED
    (SENSOR/LOAD, SET): ResetDrive(id) ⊕ state=UNALLOCATED
  other: ""
end
```

Where ResetDrive() is the device driver of RESET.

3.2 Multiple-action statements

Multiple-action statements are functionally equivalent to sequences of single-action signal statements and other non-signal procedural statements, and therefore we can transform a multiple-action statement into a sequence of single-action statements. For different type resources, the sequence of every single-action statement varies. The transformation of APPLY and MEASURE statements exemplifies the process.

```
Ts【APPLY, NOUN, [SCL,] CONN】 =
  let ConnStm=Ts【CONNECT, NOUN, NULL, [SCL,] CONN】 in
  let SetupStm=Ts【SETUP, NOUN, NULL, [SCL,] CONN】 in
  let kind=GetDevKind(APPLY, noun) in
  case kind of
    SOURCE: ConnStm ⊕ SetupStm
    SENSOR/LOAD: SetupStm ⊕ ConnStm
  end
```

Where GetDevKind() is the predefined function that returns the kind of device.

3.3 Bus statements

3.3.1 ENABLE EXCHANGE statement

This statement's function is to enable a set of protocols and to initialize them for the subsequent statements and to connect them to the correct UUT pins.

Making the transformation of this statement is to allocate several databuses according to these protocols and do the initialization. Meanwhile, the mapping relationship between these allocated databuses and the identifiers exc-config should be established to enable the upcoming bus statements find the needed bus. Predefined function BusAlloc() is used to show this process.

```
Ts【ENABLE, EXC-CONFIG exc-config, PROTOCOLL】= case PROTOCOLL of
  null: ""
  other: let protocol ⊙ PROTOCOLL1=PROTOCOLL in
    let (b, p, conn)=protocol in
    let id=BusAlloc(exc-config, b, p, conn) in
    BusInit(id) ⊕ BusConn(id, conn) ⊕
    Ts【ENABLE, exc-config, PROTOCOLL1】
  end
```

Where BusInit() is the driver of initializing databus.
BusConn() is the driver that connects UUT to databus.

3.3.2 DISABLE EXCHANGE statement

This statement's function is to render an active exc-config inactive, and BusClose() driver does this job in the following descriptions. The predefined function GetBusDev() used in transformation descriptions means finding a suitable one from the devices allocated by the ENABLE statement.

```
Ts【DISABLE, EXC-CONFIG exchange-config】=
  let buslist=GetBusDev(exchange-config, all) in
  let clause=Tlist【buslist】in clause
Tlist【buslist】=(buslist=null→"",
    let id ⊙ buslist1=buslist in BusClose(id) ⊕ Tlist【buslist1】)
```

3.3.3 DO statement

Several statements support the testing of UUTs that utilize buses. Herein, the DO statement permits the actual traffic on the bus in terms of exchanges and specifies the type of test device participation in each exchange. The transformation of DO statement is to make the decision of performing bus receiving or sending data according to the different roles and/or modes and call the bus receiving or sending driver function to transmit data on the bus. Meanwhile, we decide whether the bus communication needs to be monitored according to the value of TRANS field.

```
Ts【DO, EXCHANGE, USING exc-config, PEXCL】=
  let pexc ⊙ PEXCL1=PEXCL in let (protocol,exc)=pexc in
  let id=GetBusDev(exc-config, protocol) in
  (PEXCL1=null→Texc【id, exc】,Texc【id, exc】⊕ Ts【exc-config, PEXCL1】)
```

```
  Texc 【id, MODE, talker, listener, comm, data, status, ROLE, TRANS】 =
    let result=SendOrRecv(MODE, talker, listener, ROLE) in
    let bRet=NeedMonitor(TRANS) in case (result, bRet) of
      (recv, true): let (mc, md, ms) = TRANS in
        ReceiveDrive(id, Te 【comm】, Te 【data】, Te 【status】) ⊕
        ReceiveDrive(id, Te 【mc】, Te 【md】, Te 【ms】)
    (recv, false):ReceiveDrive(id, Te 【comm】, Te 【data】, Te 【status】)
    (send, true): let (mc, md, ms) = TRANS in
      SendDrive(id, Te 【comm】, Te 【data】, Te 【status】) ⊕
      ReceiveDrive(id, Te 【mc】, Te 【md】, Te 【ms】)
      (send, false):SendDrive(id, Te 【comm】, Te 【data】, Te 【status】)
end
```

Where predefined function GetBusDev() means getting a suitable one from the allocated devices, SendOrRecv() is a predefined function, which decides whether bus receive or send data. NeedMonitor() is a predefined function, which decides whether bus need to monitor or not. ReceiveDrive() and SendDrive() are drivers of databus.

4. SUMMARY

In this paper, we present the formal transformation methods from the test language ATLAS to C++. Since there are equivalents in C++ for the non-signal statements of ATLAS, their transformation is only the one-to-one mapping from ATLAS to C++. The statements relating to signal and bus are transformed into a series of actions that mainly include codes for allocating device, calling driver and state transition in C++. According to the formal methods, we implemented a transformation system from ATLAS to C++. Although it is a complicated process to convert the signal and bus statements into C++, this paper solves most of this kind problem. And the future work lies in study of complex signal, event-based signal and the digital test.

REFERENCES

1. R. Paige (1997), Future directions in program transformations. *ACM SIGPLAN Not*, 32, 1, pp. 94–97.
2. D. Sands (1996), Proving the correctness of recursion-based automatic program transformations. *Theoretical Computer Science*, 167, 1–2, pp. 193–233.
3. L. de Alfaro and Z. Manna (1996), Temporal verification by diagram transformations. CAV'96, LNCS 1102, pp. 288–299, Springer.
4. G. Jin, Z. Li and F. Chen (2001), A theoretical foundation for program transformations to reduce cache thrashing due to true data sharing. *Theoretical Computer Science*, 255, 1–2, pp. 449–481.

5. P. Wadler (1990), Deforestation: Transforming programs to eliminate trees. *Theoretical Computer Science*, 73, 2, pp. 231–248.
6. H. Yang and Y. Sun (1997), Reverse engineering and reusing COBOL programs: A program transformation approach. IWFM'97, Electronic Workshops in Computing.
7. A. Pnueli, O. Shtrichman and M. Siegel (1998), The code validation tool CVT: Automatic verification of a compilation process. *STTT*, 2, 2, pp. 192–201.
8. IEEE (1995), Standard Test Language for All Systems—Common/Abbreviated Test Language for All Systems (C/ATLAS). 716 C/ATLAS Subcommittee of IEEE Standards Coordinating Committee 20.
9. R.A. Kemmerer (1990), Integrating formal methods into the development process. *IEEE Software*, 7, 5, pp. 37–50.

GENETIC ALGORITHM FOR EVALUATION METRICS IN TOPICAL WEB CRAWLING

T. Peng, W.L. Zuo and Y.L. Liu
College of Computer Science and Technology, Jilin University, Changchun 130012, P. R. China

Abstract A topic driven crawler chooses the best URLs to pursue during web crawling. It is difficult to evaluate what URLs downloaded are the best. This paper presents some important metrics and an evaluation function for ranking URLs about pages relevance. We also discuss an approach to evaluate the function based on GA. The best combination of the metrics' weights can be discovered by GA evolving process. The experiment shows that the performance is exciting, especially about a popular topic.

Keywords: topic driven crawler, genetic algorithm, real-coded, relevant pages.

1. INTRODUCTION

A crawler is a program that retrieves web pages for a search engine, which is widely used today. The WWW information is distributed, also the information environments become complex. Because of limited computing resources and limited time, topic driven crawler (also called focussed crawler, retrieve web pages relevant a topic) has been developed. Topic driven crawler carefully decides which URLs to scan and in what order to pursue based on previously downloaded pages information. Some evaluation methods for choosing URLs [1] and several special crawlers, Naive Best-First crawler and DOM crawler [2] do not have satisfying adaptability. In this paper, we present an approach to evaluate a function about pages' relevance based on genetic algorithm (GA). We use GA to evolve some weights of the metrics. GAs are general purpose search algorithms which use principles inspired by natural genetic populations to evolve solutions to problems [3, 4]. In our approach, not as usual, an individual is a combination of the real-coded metrics' weight, and it's more natural to represent the optimization problem in the continuous domain.

G. R. Liu et al. (eds.), Computational Methods, 1203–1208.

2. THE EVALUATION FUNCTION (METRICS)

Not all pages which crawler observed are ‘relevant’ during crawling. For instance, if a crawler builds a specialized database on a particular topic, then pages referring to that topic are more important, and should be visited as early as possible. Similarly, If a page points to lots of authority pages, then the page is a high *hub* score page [5]. If the crawler cannot visit all the pages, then it is better to visit those ‘important’ pages, since this will give the end-user higher ranking results. We define the evaluation function $I(u)$, u is the URL that the crawler will be pursued, p is the parent page of u. The evaluation function is a weighted combination of followed metrics:

1. *sim*(p, q) (similarity of page p to topic q): A topic q drives the crawling process, and *sim*(p, q) is defined to be the textual similarity between p and q, this is shown in the work by the authors [6].

$$sim(p, q) = \frac{\left(\sum_{j=1}^{r} q_j \times l_j \times \omega_j\right)}{(|\omega'| \times |p|)} \tag{1}$$

   where $\omega' = (\omega_1 \times l_1, \ldots, \omega_r \times l_r)$; ω_j: the weight of the jth word;
   l_j: the inverse document frequency (*idf*) of the jth word;
2. *hub*(p) (the evaluation of hubs property): *Hub* pages are defined to be web pages which point to lots of ‘important’ pages relevant a topic.

$$hub(p) = \frac{|L_p|}{\frac{\sum_{i=1}^{N} |L_i|}{N}} \tag{2}$$

   where $|L_p|$: the number of outlinks of page p;
   $\frac{\sum_{i=1}^{N} |L_i|}{N}$: the average number of outlinks of the pages that are already downloaded.
3. *bc*(p) (backlink count): the number of links to p

$$bc(p) = \frac{|P_p|}{M} \tag{3}$$

   where M is a parameter provided by user.
4. *uts*(u, q) (similarity of URLs text to topic p):

$$uts(u, q) = sim(u, q) + thesaurus(u, q) \tag{4}$$

   where *thesaurus*(u, q) is uses the thesaurus dictionary of topic q, this experimentation does not take the metric(the future work I will do).

5. *pagerank*(p): a page p that is pointed at by pages $t_1, \ldots, t_n$, and c_i is the number of links going out of page t_i. d is a damping factor [1].

$$pagerank(p) = (1 - d) + d\left(\frac{pagerank(p_1)}{c_1} + \cdots + \frac{pagerank(p_n)}{c_n}\right) \tag{5}$$

$$\begin{aligned} I(u) = {} & \omega_1 \times sim(p, q) + \omega_2 \times hub(p) + \omega_3 \times bc(p) + \omega_4 \times uts(u, q) \\ & + \omega_5 \times pagerank(p) \end{aligned} \tag{6}$$

where $\sum_{i=1}^{5} \omega_i = 1$. A high $I(u)$ value indicates u links more relevant page to the topic.

3. EVOLVE THE WEIGHTS WITH GENETIC ALGORITHM

The weights of the metrics (6) are evolved with genetic algorithm. The individuals are real-coded, because that the representation of the solution could be very close to the natural formulation of our problem. Since the amount of the weights equals to 1, the weight ω_i is coded into the gene, c_i, and c_i is defined by

$$c_i = \sum_{j=1}^{i} \omega_j, i = 1, \ldots, 4 \tag{7}$$

Each individual in the population is the combination of $c_1, \ldots, c_4$. Obviously, there must be a restriction of any individual, $x_i \geq x_{+1}$, to ensure the individual could be decoded into the weights.

The standard fitness proportional model, which is also called roulette wheel selection, is used as the selection method to select the individuals for reproduction. The probability of an individual to be selected is $P_i = f_i / \sum_{j=1}^{n} f_j$, n is the population size [7]. Individuals are crossed using simple crossover method [8]. We can assume that $C_1 = (c_1^1, \ldots, c_5^1)$ and $C_2 = (c_1^2, \ldots, c_5^2)$ are two chromosomes selected for the crossover operator. The single crossing position $j \in \{1, \ldots, 3\}$ is randomly chosen and the two new chromosomes are built as

$$C_1' = \left(c_1^1, c_2^1, \ldots, c_i^1, c_{i+1}^2, \ldots, c_5^2\right) \quad C_2' = \left(c_1^2, c_2^2, \ldots, c_i^2, c_{i+1}^1, \ldots, c_5^1\right)$$

Of course, due to the restriction of the individual which is referred before, the genes of C_1', C_2' must be sorted according to the sort ascending. If an individual is chosen for the mutation operator, one of the randomly chosen gene c_i will change to $c_i' \in (c_{i-1}, c_{i+1})$ which is a random value, and we assume that

$c_0 = 0$, $c_5 = 1$. Finally, all of the individuals, including new and old ones, are sorted by their fitness, and the best fit individuals become the new population in the next generation. We set the probability of crossover to be 0.8, and the probability of mutation to be 0.05. After 50 generations, we finish the evolving process, choose the individual with the highest fitness of the population, and decode the genes to the weights as the result.

4. THE EXPERIMENTS AND RESULTS

In the experiment, we built a virtual crawler, and crawled the web starting from 200 relevant URLs (Seed URLs) simultaneously. After crawling 18,000 pages, all the crawler threads stop. We choose the best individual based on the metric of *fitness*.

$$fitness = \frac{1}{N} \cdot \sum_{i=1}^{N} x_i, \quad x_i = \begin{cases} 1, sim(p_i, q) \geq \theta \\ 0, sim(p_i, q) < \theta \end{cases} \tag{8}$$

θ is a parameter provided by user, N is the number of crawled pages.

Table 1 shows the best weight combination of the different topics. Figure 1 shows the percent of the relevant pages on the topic of 'computer'. The dashed represents every weight $\omega_i = 0.2$ fixed. The real line represents the best individual weight. Figure 2 on the topic of 'drama'.

Table 1. The best weight combination.

Topics	w_1	w_2	w_3	w_4	w_5
Computer	0.218	0.207	0.106	0.353	0.116
Drama	0.175	0.226	0.103	0.371	0.125

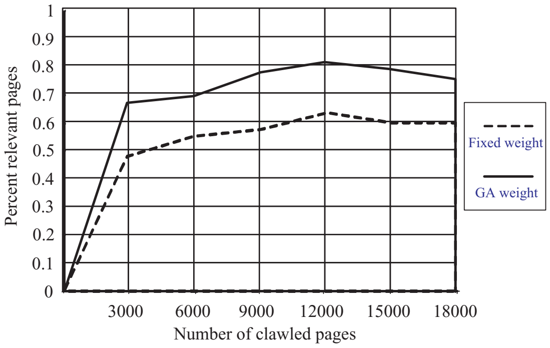


Figure 1. Topic of computer.

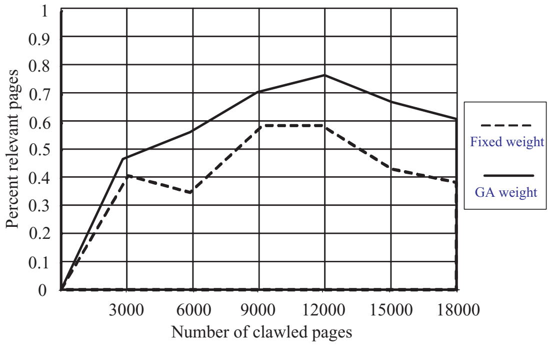


Figure 2. Topic of drama.

5. CONCLUSIONS

In this paper, a topic driven crawler based on GA has a good performance during crawling. According to the two figures above, popular topic 'computer' has a better *fitness* than topic 'drama'. After GA evolving, the results of above experiment is significant, and this is just what we expected. We will perfect our work, metrics $thesaurus(u, q)$ etc., in future.

6. ACKNOWLEDGEMENTS

This work is sponsored by the National Natural Science Foundation of China under grant number 60373099.

REFERENCES

1. J. Cho, H. Garcia-Molina and L. Page (1998), Efficient crawling through URL ordering. In: *Proceedings of the 7th World Wide Web Conference*.
2. G. Pant and F. Menczer (2003), Topical crawling for business intelligence. In: *Proceedings of the 7th European Conference on Research and Advanced Technology for Digital Libraries (ECDL)*.
3. J.H. Holland (1975), *Adaptation in Natural and Artificial Systems*. The University of Michigan Press.
4. D.E. Goldberg (1989), *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, New York.
5. J. Johnson, K. Tsioutsiouliklis and C.L. Giles (2003), Evolving strategies for focused Web crawling. In: *Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003)*, Washington, DC.
6. B.W. Xu and W.F. Zhang (2001), *Search Engine and Information Retrieval Technology*. Tsinghua University Press, Beijing, China, pp. 147–150.

7. C.G. Zhou and Y.C. Liang (2001), *Computational Intelligence*. Jilin University press, Changchun, China.
8. F. Herrera, M. Lozano and J.L. Verdegay (1998), Tackling real coded genetic algorithms: operators and tools for behavioural analysis. *Artificial Intelligence Review*, 12, pp. 265–319.
9. G. Salton (1983), *Introduction to modern information retrieval*, 1st ed. McGraw-Hill.

A BOUNDARY METHOD TO SPEED UP TRAINING SUPPORT VECTOR MACHINES

Y. Wang[1], C.G. Zhou[1], Y.X. Huang[1], Y.C. Liang[1] and X.W. Yang[2]
[1]*College of Computer Science and Technology, Jilin University, Key Laboratory for Symbol Computation and Knowledge Engineering of the National Education Ministry, Changchun 130021, China*
[2]*Department of Applied Mathematics, South China University of Technology, Guangzhou 510640, China*

Abstract In this paper, we propose a boundary method to speed up constructing the optimal hyperplane of support vector machines. The boundary, called key vector set, is an approximate small superset of support vector set which is extracted by Parzen window density estimation in the feature space. Experimental results on Checkboard data set show that the proposed method is more efficient than some conventional methods and requires much less memory.

Keywords: support vector machines, Parzen window, key vector, feature space.

1. INTRODUCTION

Support vector machine (SVM) is a powerful machine learning method based on Statistical Learning Theory, which is introduced by Vapnik et al. [1]. The main difference between SVM and the traditional learning method lies in that: SVM learning approaches to map the input data into a higher dimensional feature space, and then attempts to construct the optimal separating hyperplane using the Structural Risk Minimization rule (SRM) in order to make less mistakes. This optimization problem is known to be challenging when the number of data points exceeds thousands [2]. Though having no theory to guarantee samples in higher dimensional feature space would be certainly linear division, a large number of experimental results have indicated the assumption that searching optimal separating hyperplane in higher dimensional feature space is feasible [3]. As a traditional feature selection method, Parzen window density estimation has been widely applied to many pattern recognition problems [4].

G. R. Liu et al. (eds), Computational Methods, 1209–1213.

In this paper, a small approximate superset of support vectors set, called Key Vector set, is presented in order to speed up training SVM. Extracting key vector set is based on the distribution character of large number of training samples in feature space by Parzen window density estimation.

2. STANDARD SVM

Let $(x_i, y_i), i = 1, \dots, m$ be the training set, where $x_i \in R^n, y_i \in (-1, 1)$. Training a SVM is equivalent to solving a linearly constrained quadratic programming problem: $\min_{\alpha} \; W(\alpha) = \frac{1}{2}\alpha^T Q\alpha - e^T\alpha, \; s.t. \; \alpha^T y = 0, 0 \le \alpha \le C$, where C is a regularization constant, the matrix Q is defined by $Q_{ij} = y_i y_j K(x_i, x_j), \; i, j = 1, \dots, m$. Then the separating rule, based on the optimal hyperplane, is the indicator function: $f(x) = sgn\left[\sum_i \alpha_i y_i K(x, x_i) + b\right]$, where K is a kernel function, x_i with nonzero α_i are the support vectors.

In kernel designs, we employ the idea to transform the input patterns into the reproducing kernel Hilbert space (feature space) by a set of mapping functions $\varphi(x)$, and $\varphi(x_i) \cdot \varphi(x_j) = K(x_i \cdot x_j)$ must satisfy the Mercer condition, usually a RBF Gaussian kernel: $K(x, x') = \exp(-||x - x'||^2/2\sigma^2)$ [1].

3. PARZEN WINDOW DENSITY ESTIMATION

The Parzen window density estimation can be used to approximate the probability density $p(x)$ of a vector of continuous random variables X. Given a set of n dimensional samples $(x_1, \dots, x_n)$, the probability density estimate by the Parzen window is given by $p(x) = \sum_i^n \omega(x - x_i, h)$, where ω is a window function and h is the window width parameter. Parzen showed that $p(x)$ converges to the true density if ω and h are selected properly [4].

4. KEY VECTOR SET

Definition 1. The vector located on the boundary of the same clusters in feature space is called **Key Vector** (KV).

Definition 2. The set of key vectors is called **Key Vector Set**.

As the definition of key vector, it can be assumed that the probability density of a KV must be low. So the proposed key vector set is extracted as follows:

1. Calculate the feature space distance D_{ij} between two vectors. For linear classification we have $D_{ij} = ||x_i - x_j|| = \sqrt{x_i \cdot x_i + x_j \cdot x_j - 2^*(x_i \cdot x_j)}$ and

for nonlinear kernel classifier we have

$$D_{ij} = ||\varphi(x_i) - \varphi(x_j)|| = \sqrt{\varphi(x_i)\cdot\varphi(x_i) + \varphi(x_j)\cdot\varphi(x_j) - 2^*(\varphi(x_i)\cdot\varphi(x_j))}.$$

Note that $\varphi(x_i)\cdot\varphi(x_j) = K(x_i \cdot x_j)$, it follows that

$$D_{ij} = \sqrt{K(x_i\cdot x_i) + K(x_j\cdot x_j) - 2^*K(x_i\cdot x_j)}\,.$$

2. Calculate $D_{Limit} = \frac{1}{n^2R}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ij}$, where constant R restricts the range of density.
3. Estimate the intensity $\omega(D_{ij})$ of two points using Parzen window function

$$\omega(D_{ij}) = \begin{cases} 1 & D_{ij} \le D_{Limit} \\ 0 & D_{ij} > D_{Limit} \end{cases}, j \in [1, \ldots, m],$$

   which is a rectangle function to get sample density, or $\omega(D_{ij}) = e^{-D_{ij}^2/2\rho^2}$, $D_{ij} \le D_{Limit}$, $j \in [1, \ldots, m]$, which is a Gauss function to get probability density. And set $\rho = \frac{D_{Limit}}{2}\cdot\frac{1}{(2\ln 2)^{1/2}}$ to make the points in $\frac{D_{Limit}}{2}$ ranges to influence its density greatly [5].
4. Calculate density $DENS_i$ of each vector:

$$DENS_i = \sum_j \omega(D_{ij}),\ i, j \in [1, \ldots, m]\,.$$

5. Make up the key vector set by the corresponding vector x_i, which $DENS_i$ is smaller than the threshold or proportional value Par. In our following Checkboard experiment, we set Par and $DENS_i$ as $DENS_i \le \min(DENS_j) + Par^*(\max(DENS_j) - \min(DENS_j))$.

5. TRAINING SVM USING KEY VECTOR SET

As the assumption that samples in the feature space would be linear division [3], the KV set would be an approximate superset of support vectors set. Training SVM using Key Vector set could be implemented as follows:

1. The key vector set can be extracted by the proposed algorithm.
2. Then to obtain the optimal hyperplane by SVM from the key vector set using the same related kernel function.

6. IMPLEMENTATION AND RESULTS

In order to examine the speed and effectiveness of the presented algorithm, we apply the proposed algorithm to Checkerboard data set, which has 1000 samples in two clusters [6]. From Table 1, it can be seen that the running time

Table 1. Comparison of different Par values for Checkerboard, R = 3, $\sigma = 20$, C = 100, RBF.

Par value	Key vector	Running time (s)	Training accuracy	SV (>1E−6)
1(standard SVM)	500 + 500	1976.37	98.9%	126(81)
0.3	230 + 224	162.64	97.7%	97(63)
0.2	133 + 117	14.39	94.0%	78(59)
0.1	44 + 37	0.33	71.2%	49(46)

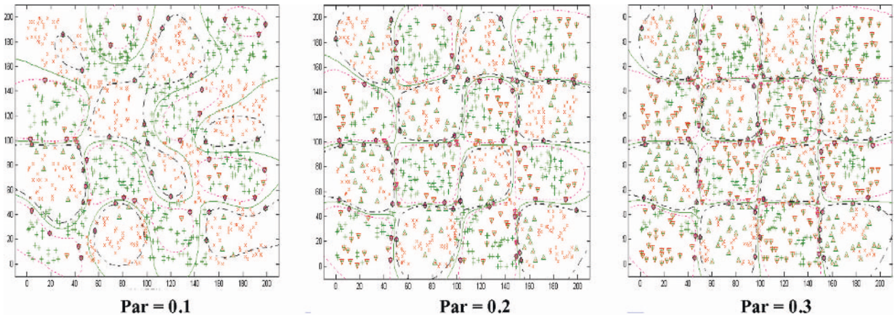


Figure 1. Checkerboard experiment results of several Par values (Δ, ∇ key vector, O support vector,—optimal hyperplane).

of the proposed method is shorter greatly than that of the conventional method (spends almost 2000 seconds), the training accuracy is preferable too. And the experimental results can be seen from Figure 1. The experiments have been performed on a PC (P3 1GHz, 128M, Matlab6.5) and take about 40 seconds to get KV set.

7. CONCLUSIONS

In this paper, a boundary method to speed up SVM is presented. We extract the key vector set, an approximate superset of support vector set, to construct the optimal hyperplane and have obtained better results. Future research of this topic is to find a more effective method of extracting key vector set and to choose the parameters according to the information of kernel function.

ACKNOWLEDGEMENT

This work was supported by the National Natural Science Foundation of China under Grant No. 60175024 and 60433020, and the Key

Science-Technology Project of the National Education Ministry of China under Grant No. 02090.

REFERENCES

1. V. Vapnik (1995), *The Nature of Statistical Learning Theory*. Springer-Verlag, New York.
2. S.F. Zheng (2004), Unsupervised clustering based reduced support vector machines. *Computer Engineering and Application*, 14, pp. 74–6 (in Chinese).
3. M.H. Yang and N. Ahuja (2000 June), A geometric approach to train support vector machines. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2000 (CVPR 2000)*, Hilton Head Island, pp. 430–7.
4. N. Kwak and C.H. Choi (2002 December), Input feature selection by mutual information based on Parzen window. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 24, 12, pp. 1667–71.
5. Z.Q. Bian and X.G. Zhang (2000), *Pattern Recognition*, 2nd ed., Tsinghua University Press, Beijing.
6. URL (1999), *Check Board*, http://www.cs.wisc.edu/∼olvi/data/check1.txt.

A FAST ALGORITHM FOR GENERATING CONCEPTS USING AN ATTRIBUTE TABLE

Ming Lu, Chengquan Hu, Hong Qi, and Liang Zhao
The College of Computer Science and Technology, Jilin University, China 130012

Abstract As an unsupervised learning technique for conceptual clustering, Formal Concept Analysis has been widely used in many areas including machine learning and data mining. This paper organizes the search space for concepts as a prefix tree, and employs a new data structure called Attribute Table to map and prune the tree. Thus, the procedures of conceptual clustering can be carried out only in a few valid subspaces. A fast algorithm for generating concepts is proposed in this paper. Experimental evidence shows that our algorithm performs very well for generating concepts on both dense and sparse contexts.

Keywords: formal concept analysis, conceptual clustering, prefix tree, attribute table.

1. INTRODUCTION

As an unsupervised learning technique for conceptual clustering, Formal Concept Analysis has been widely used in many areas including machine learning [1] and data mining [2]. Several algorithms have been proposed to generate concepts or a lattice structure, called concept lattice [3–6]. Some researches show that the algorithms based on the computation of closures and the test for canonicity will represent better performance than others when the contexts (or databases) are large and dense [6]. Unfortunately, such algorithms are still very time-consuming.

Our work is to design a new fast algorithm for the use of generating concepts for large and dense formal contexts. Our method is to organize the search space for concepts as a prefix tree (also called a trie) named *Concept Trie*. A new data structure called *Attribute Table* is used to map the information of the trie and further to prune the trie. By these means, the procedures of conceptual

G. R. Liu et al. (eds.), Computational Methods, 1215–1219.

clustering are carried out only in some valid subspaces. In each valid subspace, at least one canonical concept can be found while there is no canonical concept in any invalid subspace. As a result, concept lattice is generated in a most small subset of the original search space because it is just composed of all canonical concepts.

Readers can learn from [7] for the details of Formal Concept Analysis, such as formal context, Galois connection, closure operator, formal concept and canonical closure.

2. CONCEPT TRIE

Considering a linear order on M that $m_1 < m_2 < \ldots < m_{|M|}$, each subset of M could be represented as a string. The whole search space for concepts can be organized as a prefix tree, also called a trie, and named *Concept Trie*. Each subtree is exactly a subspace. A subtree is called an *invalid* subtree if and only if there are no canonical concepts (or closures) on it, or a *valid* one. Invalid subtrees can be cut off.

Let $A = B \cup \{r\}$ and $r \in M \backslash B$, we can directly draw an obvious conclusion that A'' is a canonical closure if and only if A'' is one node of the subtree whose root is A. Instead of computing A'', we give a more convenient method to check the validity of a subtree.

Conclusion 1. Let $A = B \cup \{r\}$ and $r \in M \backslash B$, we write $r_{B'} = B' \cap \{r\}'$. The subtree with root A can be cut off if any of two statements mentioned below becomes true:

1. $\exists m \in B, r_{B'} \cap m_{B\prime} = \emptyset$, and $\exists\, n \in M \backslash B, n < r$.
2. $\exists m \in M \backslash B$, fulfilling $m < r$ and $r_{B'} \subseteq m_{B'}$.

Proof 1. $m \in B \Rightarrow A' \subseteq m_{B'} \Rightarrow A' \subseteq mB' \cap r_{B\prime} \Rightarrow A' = \emptyset \Rightarrow A'' = M \Rightarrow n \in A'' \Rightarrow A''$ is not canonical; 2. $(A' \subseteq r_{B'}) \wedge (r_{B'} \subseteq m_{B'}) \Rightarrow A' \subseteq mB' \Rightarrow A' \subseteq \{m\}' \Rightarrow m \in A'' \Rightarrow A''$ is not canonical.

Meanwhile, considering performance, we generate canonical closures just in company with subtree validity checking according to the conclusion below.

Conclusion 2. Let $A = B \cup \{r\}$ and A'' be canonical. For each attribute m that $r < m, m \in A''$ if and only if $r_{B'} \subseteq m_{B'}$.

Proof 2. $r_{B'} \subseteq m_{B'} \Rightarrow (B \cup \{m\})'' \subseteq (B \cup \{r\})'' \Rightarrow m \in A''; 2. (m \in A'') \wedge (B' \cap \{r\}' \subseteq B') \Rightarrow \{B \cup \{r\}\}' \subseteq \{m\}' \Rightarrow B' \cap \{r\}' \subseteq \{m\}' \Rightarrow B' \cap \{r\}' \subseteq B' \cap \{m\}' \Rightarrow r_{B'} \subseteq m_{B'}$.

Conclusion 3. If A'' is canonical, then $A'' = B \cup \{r\} \cup \{m|(r < m) \wedge (r_{B'} \subseteq m_{B'})\}$.

3. ATTRIBUTE TABLE AND THE ALGORITHM

We design an a *ttribute table* to prune the Concept Trie and extract concepts at the same time. The information held in the table map a subset of the Trie. Given a formal context $K = (G, M, R)$, the Attribute Table is $AT = (B, r, C, T)$. $B \subset M$ is the *base set* (initialized as G'), and is the father of the current subtree. $r \in B$ is a *reference attribute*. $B \cup \{r\}$ is the root of current subtree. $C = M \backslash B$ is the *candidate set* (initialized as M). The *body* of the table is T: $C \times C \rightarrow 2^{|G|}$, and $AT[m_1, m_2] = m_{1B'} \cap m_{2B'}$. We call the set consisting all $AT\ [m_i, m_i](\forall m_i \in C)$ the *diagonal* of the table.

We can perform two operations on the Attribute Table. Algorithm UpdateAT (Table 1) is used to update an original Attribute Table, by which means we can cut off some invalid subtrees and get some concepts. Another operation is DeriveAT (see algorithm 2 in Table 2). By this operation, a new Attribute Table is created for the used of mapping a valid subtree that will be considered

Table 1. Algorithm UpdateAT.

```
Algorithm 1 UpdateAT
Input: an original Attribute Table
Output: a filled Attribute Table and some concepts
FOR EACH $m_j \in AT.C$ that $m_j < AT.r$ DO
   FOR EACH $m_i \in AT$ that $AT.r < m_i$ DO
   IF $AT\ [m_j, m_j] \subseteq AT\ [m_i, m_i]$ THEN remove $m_j$ from $AT.C$ ENDIF
   ENDFOR
   Remove $m_j$ from $AT.C$.
ENDFOR
$C := AT.C$;
WHILE there is unmarked attribute in $C$ DO
   $m_i$ := the smallest unmarked attribute in $C$;
   $Intent_i := AT.B \cup \{m_i\}$
   FOR EACH $m_j \in C$ fulfilling $m_i < m_j$ and $AT[m_j, m_j] \cap AT\ [m_i, m_i] \neq \emptyset$ DO
   IF $AT[m_i, m_i] = AT[m_j, m_j]$ THEN $Intent_i := Intent_i \cup \{m_j\}$, remove $m_j$ from $C$
   ELSEIF $AT[m_i, m_i] \subset AT[m_j, m_j]$ THEN $Intent_i := Intent_i \cup \{m_j\}$
   ELSEIF $AT\ [m_j, m_j] \subset AT\ [m_i, m_i]$ THEN mark $m_j$ENDIF
   ENDFOR
   ($AT\ [m_i, m_i]$, $Intent_i$) is a concept
   $C := C \backslash \{m_i\}$;
ENDWHILE
```

Table 2. Algorithm DeriveAT and Table.

Algorithm 2 DeriveAT

```
Input: an unmarked $m_i$ and its $Intent_i$
Output: $DAT$, mapping the subtree whose root is $Intent_i$
$DAT.r = m_i$; $DAT.B = Intent_i$; $DAT.C = \emptyset$;
FOR EACH $m_j \in AT.C$ and $m_j \neq m_i$ DO
    IF $AT[m_i, m_j] \neq \emptyset$ THEN
        $DAT.C := DAT.C \cup \{m_j\}$
        $DAT[m_j, m_j] = AT[m_i, m_j]$
    ENDIF
ENDFOR
```

Algorithm 3 TABLE

```
Input: an original Attribute Table related to the context $K$
Output: all concepts on $K$
$UpdateAT(AT)$
FOR EACH unmarked attribute $m$ in $AT.C$ DO
$DAT = DeriveAT(m)$
$TABLE(DAT)$
ENDFOR
```

next. The complete algorithm for generating concepts is named TABLE (see algorithm 3 in Table 2).

4. CONCLUSIONS

We implemented our algorithm using C++ and compared it with the famous NextClosure algorithm on the platform of Pentium IV 2.6 GHz, with main memory of 512 MB.

Two types of data are taken into account. For one thing, we used a real application test data, MUSHROOM [8]. Results are shown in Table 3. For another, we designed several random generated test data and focused our attentions on the following three parameters, the number of objects, the number of attributes and the number of relations [9]. Comparison results show that our algorithm performs much better than NextClosure on both dense and sparse contexts, especially on the real application data.

Table 3. Evaluation results for MUSHROOM database.

No. of objects	Table (s)	NextClosure (s)	No. of concepts
1000	10.48	56.67	32,514
2000	22.70	170.80	58,983
3000	34.27	361.81	80,902
4000	49.56	581.55	104,105
5000	79.95	>600.00	136,402
8124	180.06	>600.00	238,710

REFERENCES

1. S.O. Kuznetsov (2001), Machine learning on the basis of formal concept analysis. *Automation and Remote Control*, 62, 10.
2. G. Stumme (2002), *Efficient Data Mining Based on Formal Concept Analysis*. DEXA, Aix-en-Provence, France, pp. 534–546.
3. B. Ganter (1984), Two basic algorithms in concept analysis. *Technical Report 831*, Technische Hochschule, Darmstadt.
4. E.M. Norris (1978), An algorithm for computing the maximal rectangles in a binary relation. *Revue Roumaine de Mathématiques Pures et Appliquées*, 23, 2, pp. 243–250.
5. R. Godin, R. Missaoui and H. Alaoui (1995), Incremental concept formation algorithms based on Galois lattices. *Computation Intelligence*, 11, 2, pp. 246–267.
6. S.O. Kuznetsov and S.A. Obiedkov (2002), Comparing performance of algorithms for generating concept lattices. *Journal of Experimental and Theoretical Artificial Intelligence*, 14, 23.
7. B. Ganter and R. Wille (1999), *Formal Concept Analysis: Mathematical Foundations*. Springer-Verlag, Berlin, Heidelberg.
8. C.L. Blake and C.J. Merz (1998), UCI Repository of machine learning databases. http://www.ics.uci.edu/∼mlearn/MLRepository.html. University of California, Department of Information and Computer Science, Irvine, CA.

ALGORITHM OF MEASUREMENT-BASED ADMISSION CONTROL FOR GPRS

Y.H. Liu[1], B. Li[1] and H.P. Sun[2]
[1]*Department of Computer Science and Technology, Jilin University, Changchun 130012, China*
[2]*Department of Electronics Engineering, Beijing University of Post and Telecommunication, Beijing 100876, China*

Abstract The General Packet Radio Service (GPRS) provides the predictive services for the users by applying the similar algorithm in the QoS management mechanisms used by the ATM. Research shows that a connective Measurement-Based Admission Control (MBAC) algorithm is more adaptable to and efficient for the GPRS network than other admission control algorithms. Based on sufficient analysis of the characteristics of GPRS, a MBAC algorithm is presented which is suitable for GPRS using a worst-case delay calculation model with a priority queue, and equivalent leaky bucket approximation and delay normalization. The most characteristic of this algorithm is that it used both bandwidth and delay at the same time for the control admission. The emulation result shows that this algorithm can achieve a rate of as high as 80% resource utilization while satisfying the requirements of GPRS QoS.

Keywords: GPRS, predictive service, QoS management mechanisms, MBAC algorithms.

1. INTRODUCTION

The GPRS developed on the GSM network aimed for the burst data packet transmission is a transitional scheme providing data and multi-media service [1]. In order to provide the users with an end-to-end packet transmission model, GPRS adds two key components to the original GSM system: a Gateway GPRS Support Node (GGSN) and a Serving GPRS Support Node (SGSN). The QoS management mechanisms of a GPRS network include the description of traffic characteristics, the CAC (Call/Connection Admission/Acceptance Control) [2–8], the shaping/policing and the packet scheduling and so on. The CAC

G. R. Liu et al. (eds.), Computational Methods, 1221–1230.

algorithm is the most important part of QoS, which determines whether or not there are sufficient resources in the network to meet the demands of the connected QoS. Since the MBAC [9–11] obtains the users' resource requirements, and performs real-time measurements of the system, it is more adaptable to a network environment varying frequently.

2. MODELLING ASSUMPTION AND DELAY NORMALIZATION

First, the basic model of the research is presented. This paper only focuses on a single cell in which the GPRS provides the service, instead of considering the handoff among cells. The data service of packet switching within a cell can use eight channels: '0' channel serves as the main channel, and the other channels are the subordinate ones. MS may use several channels when necessary. Every channel operates at 9.05 Kb/s and MS encodes at half of the rate. The bandwidth of the cell is 72.4 Kb/s. The time difference from entering a cell to leaving it is called the end-to-end delay, which can be measured between Um (air interface) and Gi (between GGSN and external PDN).

Using the leaky bucket model [12] to describe each predictive service flow p with class i and priority k, each leaky bucket has two parameters: the token producing rate and the depth of the leaky bucket (r^p, b^p). For every type of predictive service, the cell preserves the bandwidth superimposed volume $\hat{v}_i$, the real queue delay $\hat{D}_k$ and the predefined restrain delay $D_{k.}$ The scheduling principle adopts delayed-requirement based priority distribution scheduling algorithm [13].

2.1 Delay calculation model

A system with many different priorities presumes: (1) the source's peak value rate is infinite; (2) the service principle of each class is FIFO (First in First Out); and (3) the total bandwidth of the system is u. It has been proved that the most worst-case delay of class j is:

$$D_j^* = \frac{\sum_{i=1}^{j} b_i}{u - \sum_{i=1}^{j-1} r_i} \tag{1}$$

Now, use formula (1) as the basic equation to compute the influence caused by admitting new flow to existed predictive service flow. D_k^* is defined as the delay. Regard all flows of the same class as one flow conforming to leaky bucket parameter (v_i, b_i). After admitting a new flow with parameters (r_k^p,

b_k^p), the leaky-bucket parameters of present class can be considered as ($v_i + r_k^p$, $b_i + b_k^p$). According to the above assumptions, it can be summarized as follows:

(1) the effect on the flows of the same priority after admitting a new one is:

$$D_k^{*\prime} = \frac{\sum_{i=1}^{k-1} b_i}{u - \sum_{i=1}^{k-1} v_i} + \frac{b_k + b_k^p}{u - \sum_{i=1}^{k-1} v_i} = D_k^* + \frac{b_k^p}{u - \sum_{i=1}^{k-1} v_i} \tag{2}$$

(2) the effect on the flows of low priority after admitting a new one is:

$$\begin{aligned} D_j^{*\prime} &= \frac{\sum_{i=1}^{k-1} b_i + b_k + b_k^p + \sum_{i=k+1}^{j} b_i}{u - \sum_{i=1}^{k-1} v_i - v_k - r_k^p - \sum_{i=k+1}^{j-1} v_i} \\ &= D_j^* \frac{u - \sum_{i=1}^{j-1} v_i}{u - \sum_{i=1}^{j-1} v_i - r_k^p} + \frac{b_k^p}{u - \sum_{i=1}^{j-1} v_i - r_k^p} \end{aligned} \tag{3}$$

2.2 Delay normalization

The packet length of GPRS is variable, while the packet delay is relative to the length of packets. Since the delays of packets, which have different lengths, are incomparable and they cannot be simply added together. To measure them, the delay normalization is needed. The concrete method is as follows:

$$D_{i,l} = \begin{cases} \dfrac{0.5 \times l}{128}, & i = 1,\ l \le 128, \\ 0.5 + \dfrac{1.5 \times l}{1024 - 128}, & i = 1,\ l > 128, \\ \dfrac{5 \times l}{128}, & i = 2,\ l \le 128, \\ 5 + \dfrac{10 \times l}{1024 - 128}, & i = 2,\ l > 128, \\ \dfrac{50 \times l}{128}, & i = 3,\ l \le 128, \\ 50 + \dfrac{25 \times l}{1024 - 128}, & i = 3,\ l > 128. \end{cases} \tag{4}$$

According to the Table 1, a delay requirement can be obtained from formula (4), where $D_{i,l}$: the packet delay required with i as service class and l as length of packets. The ratio of measurement packet delay to $D_{i,l}$ is defined as the normalized delay.

Table 1. The delay requirement classes.

Service classes			Predictive service			Best-effort transmission
Delay classes			1.0	2.0	3.0	4
Packet sizes	128 octets	Average	0.5	5.0	50.0	No description
		95%	1.5	2.5	150.0	
	1024 octets	Average	2.0	15.0	75.0	
		95%	7.0	75.0	375.0	

Now define the real delay of the measurement as $D'_{i,l}$, the delay normalized as $D^{\Delta}_{i,l}$, then it shows that:

$$D^{\Delta}_{i,l} = \frac{D'_{i,l}}{D_{i,l}} \tag{5}$$

The average normalized delay of every class is:

$$\bar{D}^{\Delta}_{i,l} = \left[\sum_{p=1}^{n} \left(D^{\Delta}_{i,j}\right)_p\right] \Bigg/ n, \quad 1 \le i \le 3 \tag{6}$$

3. ALGORITHM DESCRIPTION

3.1 Basic admission control principles

(1) If the sum of the present bandwidth and the bandwidth required by new flow exceeds the available bandwidth of the system, the request of flow p will be refused.

$$vu > r_k^p + \sum_{i=1}^{3} \hat{v}_i \tag{7}$$

where v is available bandwidth factor. When system load is heavy, the square deviation of the delay grows bigger and the MBAC algorithm will not be effective. So MBAC cannot make full use of the bandwidth. Therefore, an object factor of the bandwidth whose value is decided by the characteristics of connection flow should be set up. When the source rate of each flow compared with the total connection ability of system is very low and the period of burst data is short, the object factor should be bigger. Otherwise, its value should be smaller. Because the retransmission rate is very high and is effected by handoff in GPRS, the value of ‘v’ cannot be too high though the period of burst data is short. In this paper, its value is 0.8.

(2) If the same priority class can destroy the delay restriction after admitting flow p, then the flow p should be refused:

$$D_k > \hat{D}_k + \frac{b_k^p}{u - \sum_{i=1}^{k-1} \hat{v}_i} \tag{8}$$

Or if the low priority class can destroy the delay restriction after admitting flow p, the request of flow p will be refused too:

$$D_j > \hat{D}_j^* \frac{u - \sum_{i=1}^{j-1} \hat{v}_i}{u - \sum_{i=1}^{j-1} \hat{v}_i - b_k^p} + \frac{v_k^p}{u - \sum_{i=1}^{j-1} \hat{v}_i - r_k^p}, \tag{9}$$

After the delay being normalized, the formula (8), and (9) become as follows, respectively:

$$1 > \bar{D}_k^{\Delta} + \frac{b_k^p}{u - \sum_{i=1}^{k-1} \hat{v}_i} \times \frac{1}{D_{k,j}} \tag{10}$$

$$1 > \bar{D}_j^{\Delta *} \frac{u - \sum_{i=1}^{j-1} \hat{v}_i}{u - \sum_{i=1}^{j-1} \hat{v}_i - b_k^p} + \frac{b_k^p}{u - \sum_{i=1}^{j-1} \hat{v}_i - r_k^p} \times \frac{1}{D_{j,l}}, \quad k < j \leq 3 \tag{11}$$

(3) If formula (7), (10) and (11) are all satisfied, this flow can be admitted.

4. TIME WINDOW MEASUREMENT ALGORITHM

4.1 Bandwidth measurement

To estimate the bandwidth to be used, one time window can be divided into 10 sample periods. The average bandwidth used, $\hat{v}_j^s$ (sample period is s), of predictive service of every class j in every sample period should be calculated. The estimated bandwidth should be updated under the following four scenarios: (1) when time window is to its end, maximum $\hat{v}_j^s$ is the estimated bandwidth; (2) when a new flow is admitted, the estimated bandwidth should be updated and restart the time window; (3) when $\hat{v}_j^s$ exceeds the estimated bandwidth, $\hat{v}_j^s$ is the estimated value; and (4) when any flows leaves, the estimated bandwidth should be updated and restart time window. The updating formula of the estimated bandwidth is:

$$\hat{v}' = \begin{cases} MAX(\hat{v}_j^s), & \text{while time window to be end,} \\ \hat{v}_j^s, & \text{if } \hat{v}_j^s > \hat{v} \ (\hat{v} \text{ is the average bandwidth used in } S), \\ \hat{v} + r^p, & \text{while admitting one new flow } P, \\ \hat{v} - \beta r^p, & \text{while one of flows } P \text{ leaves } (\beta \text{ is bandwidth} \\ & \text{decreasing factor).} \end{cases} \tag{12}$$

4.2 Delay measurement

To estimate the delay, the queue delay of each frame is measured. The estimated delay should be updated under the four scenarios: (1) when time window is to its end, the maximum D^{Δ} is the estimated delay; (2) when new flow is admitted, the estimated delay should be updated and restart time window; (3) when D^{Δ} exceeds estimating delay, λ timing D^{Δ} is taken as the estimated delay; and (4) when one of the flows leaves, the time window is restarted immediately. The updating formula of estimating delay is as follows:

$$\bar{D}_j^{\Delta'} = \begin{cases} MAX(D^{\Delta}), & \text{while time windows to be end,} \\ \lambda D^{\Delta}, & \text{if } D^{\Delta} > \bar{D}_j^{\Delta} \ (\lambda \text{ is delay increasing factor),} \\ \text{the right sides of the} & \text{while admitting one new flow } P, \\ \text{(10) and (11),} & \\ \text{no change,} & \text{while one of flows } P \text{ leaves.} \end{cases} \tag{13}$$

There are four variables in the course of time window measurement:

1. *Delay increasing factor* (λ): When the queuing delay of frames exceeds the maximum delay estimated, the bigger maximum delay (the result of increasing factor multiplied by queuing delay of frames) should be used. In this paper, this value is 2.
2. *Sample period (S)*: The sample period controls the sensitivity of measurement to data rate. The shorter the sample period is, the more sensitive to burst data it is. In this paper, there are 54 TDMA circulations in one sample period. The value of S is 996.84 ms.
3. *Time window (T)*: The size of time window controls the adaptability of measurement algorithm. The smaller the size of time window is, the more adaptable the algorithm is, but the bigger the size of time window is, the more stable the algorithm is. And it should be the multiple of sample period. Having admitted a new flow, a time window should be restarted to collect the information of the new flow. In this paper, there are 10 sample periods whose value is 9.9684 s in one time window.
4. *Bandwidth decreasing factor* (β): When any flow leaves the network, the estimated bandwidth used should be decreased. The decreased amount approximately equals to the equivalent bandwidth of that flow. The individual value of β can be determined differently according to different flows. To simplify the algorithm, the definite value of β is set at 1 in this paper.

5. EMULATIONS

5.1 Three data source models

Three models are all supposed to be Poisson arrival process, the interval conforms to the exponential distribution, with average of interval of 15 s (except downstream channels in the WWW Model). The length of packet is described by the following three statistical models. The maximum frame length is 1600 bytes. The packets whose frame length is bigger than it are divided into more frames.

5.2 FUNET model

This model is produced based the statistical data from emails by the Finish University and Research Network (FUNET). The probability distribution function of this model can be considered to conform to Cauchy distribution with parameter (0.8, 1) approximately, M is maximum of packet size (10 KB), $0 < x < M$ (Table 2).

$$f(x) = Cauchy(0.8, 1) = \frac{1}{\pi(1 + (x - 0.8)^2)} \tag{14}$$

5.3 Mobitex model

This model is generated from the statistical data used in Mobitex Wireless Data Network based on the application of management ships in Sweden. The upstream channel: $30 + \text{random}(-15, 15)$ bytes; the downstream channel: $115 + \text{random}(-57, 57)$ bytes.

5.4 WWW model

It is supposed that the available time slot number n of MS varies in the scope 1, 2, 3, 7. The transmission rate of each time slot is 9.05 KB (CS1). The different delay classes and time slots are assigned to three different applications,

Table 2. The distribution of FUNET model.

Packet size (K)	<0.5	<1	<1.5	<2	<3	<5	<10
Percent	10	36	54	67	79	87	91

Table 3. The types and ratio of different groups.

Semanteme	1	2	3	4	5	6	7	8	9	10	11
Service class	1	1	1	1	1	2	2	3	3	3	3
Delay class	1	1	2	3	4	1	1	1	1	2	2
Time slot	1	2	1	1	2	1	2	3	7	3	7
Ratio	1/24	1/24	1/12	1/12	1/12	1/6	1/6	1/12	1/12	1/12	1/12

E-mail (service class 1), Fleet management (service class 2) and www browse (service class 3), then 11 types of data packets are obtained. Table 3 illustrates the concrete description. When data packets are generated, the generation of data source chooses one of 11 types of data packets in a random order according to the ratio indicated in Table 3.

5.5 Emulation results

The results of the several CAC algorithms for GPRS network are compared with each other in experiments using the same data source, the same scheduling mechanism and the same emulating environment. The leaky bucket algorithm is used to describe the resource requirements in all three algorithms, and the time window measurement algorithm is employed in two MBAC algorithms. The time window size is 9.9684 s including 10 sample periods. The MBAC algorithm presented adopts the bandwidth equation and the delay equation at the same time. It can achieve higher resource utilization rate. Some necessary system parameters are recorded in the experiments. According to those records, every CAC's accuracy can be given through the emulation, and a perfect MBAC algorithm is presented right now. Figures 1 and 2 illustrate the emulation comparisons of the four CAC algorithms for GPRS. It can be concluded:

(1) The efficiency of CAC algorithm basing on peak rate (SSAC) is the worst. The average usage rate only runs as high as 40%. Furthermore, the multiplexing among the data sources is not to be considered, so the adaptability to variance of load is worse. Therefore, the traffic passed will soon come to its maximum while loads increasing upon a specific value.

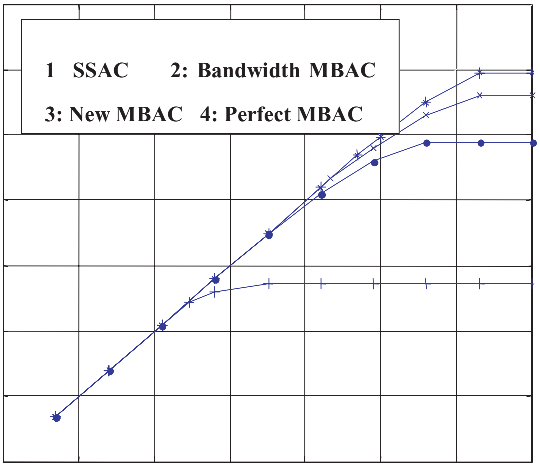


Figure 1. The comparisons of four algorithms under different loads.

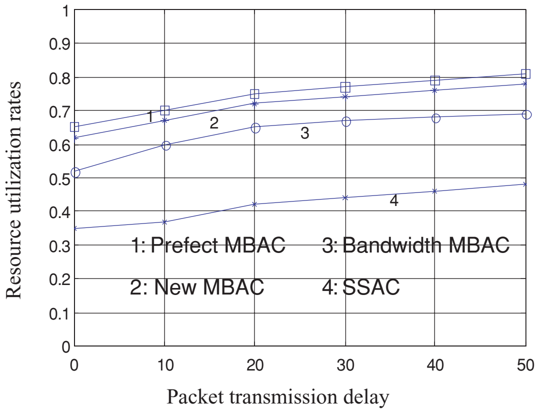


Figure 2. The emulation comparisons of four algorithms.

(2) Compared with the SSAC, the MBAC algorithm only using bandwidth can achieve resource utilization rate of about 70% and it is more adaptable to the variance of load. But the delay increases fast because it is not taken into consideration.

(3) The new algorithm considering both bandwidth and delay at the same time can improve the accuracy of admitting new flow and achieve resource utilization rate of about 80%, which is 5% lower than the perfect MBAC on it’s performance.

6. CONCLUSIONS

The large amount of research shows that MBAC is more adaptable to a network environment varying frequently. Based on sufficient analysis on the characteristics of GPRS, this paper offers a MBAC algorithm adaptable for the GPRS with the CAC equations and measurement procedure using a worst-case delay calculation model of a priority queue as well as the equivalent leaky bucket approximation and the delay normalization. The major characteristics of the algorithm is to apply both bandwidth and delay techniques to achieve the admission control. The emulation results show that the algorithm not only can satisfy the demands of GPRS QoS, but also achieve as high as 80% source utilization rate.

REFERENCES

1. C. Jian and J.G. David (1997), General packet radio service in GSM. *IEEE Communication Magazine*, pp. 122–131.
2. G.P. Harry and M.E. Khaled (1996), Call admission control schemes: a review. *IEEE Communication Magazine*, pp. 82–91.
3. J.Y. Lee and S. Bahk (2001), Simple admission control schemes supporting QoS in wireless multimedia networks. *Electronics Letters*, 37, 11, pp. 712–713.
4. S.J. Wha and G.J. Dong (2001), Call admission control for mobile multimedia communications with traffic asymmetry between uplink and downlink. *IEEE Transactions on Vehicular Technology*, 50, 1, pp. 59–66.
5. C.H. Paik, G.W. Jin and J.H. Ahn (2001), Integrated call control in a CDMA cellular system. *IEEE Transactions on Vehicular Technology*, 50, 1, pp. 97–108.
6. C.W. Leung and W.H. Zhuang (2000), Novel system modeling in call admission control for wireless personal communications. In: *Proc. IEEE GLOBECOM*, pp. 177–181.
7. M. Bozinovski, P. Popovski and L. Gavrilovska (2000), Novel strategy for call admission control in mobile cellular network. In: *Proc. IEEE VTC*, pp. 1597–1602.
8. R.S. Mahmoud, W.H. Ibrahim and K. Parviz (2000), Adaptive allocation of resources and call admission control for wireless ATM using genetic algorithms. *IEEE Journal of Selected Areas in Communications*, 18, 2, pp. 268–282.
9. J. Sugih, B.D. Peter and J.S. Scott (1997), Comparison of measurement-based admission control algorithms for controlled-load services. In: *Proc IEEE INFOCOM'97*.
10. D. Tse and M. Grossglauser (1997), Measurement-based call admission control: analysis and simulation. In: *Proc. IEEE INFOCOM'97*, Kobe, Japan.
11. S. Jamin, P.B. Danzig, S.J. Shenker and L.X. Zhang (1997), A measurement-based admission control algorithm for integrated services packet networks. *IEEE/ACM Transactions on Networking*, 5, 1, pp. 56–70.
12. D. Logothetis and K. Trivedi (1994), The leaky bucket as a policing eevice: transient analysis and dimensioning. In: *Proc. Infocom'94.*
13. G. Kulkarni and N. Gautam (1996), Leaky bucket: sizing and admission control. In: *TR-96-39*. Department of Operations Research, University of North Carolina, Chapel Hill.

SURFACE RECOGNITION OF AUTOMOBILE PANEL BASED ON SECTION CURVE IDENTIFICATION

Ping Hu,[1] Wenbin Hou[2] and Shuhua Guo[3]
[1]*National Key Lab of Automotive Dynamic Simulation in Jilin University, Changchun, China*
[2]*Department of Computer Science and Technology in Jilin University, Changchun, China*
[3]*Institute of Automobile Body and Die-engineering in Jilin University, Changchun, China*

Abstract Automobile panel surface is a kind of complicated 3-D free-from surface. According to the characteristic of the shape of the panel surface, a surface recognition scheme describing the surface in the primary inertia axes coordinate system is proposed. In the coordinate system, the surface equation will keep invariant when the surface is translated or rotated. After extracting the surface's outline projection curve and section curves, the curve's eigenvectors are calculated with an improved invariant moment method. At last, the similarity between two surfaces is figured out based on the similarity between the section curves extracted from the two surfaces separately.

Keywords: surface recognition, invariant moment, intelligent CAD, case-based reasoning.

1. INTRODUCTION

Up to now, since there has been no complete die design theory, it depends on the engineer's experience to design the automobile panel dies. Nowadays, the number of experienced designers is decreasing, while the competition between factories is becoming more and more drastically. To improve the design automatism and make the design more independent from the design specialist, it is necessary to develop the intelligent CAD system. Case-based reasoning (CBR) is one of the major research focuses in the artificial intelligence fields. When working, the CBR system does not reason from the scratch line, but resolve the new problem according to the existing experience and successful design cases. CBR performs much better than rule-based system in the fields where experience plays an important role.

G. R. Liu et al. (eds.), Computational Methods, 1231–1238.

The main idea of case-based intelligent CAD for automobile panel die structures is that: creating the relation between the automobile panel and the whole set of stamping die, then constructing case models of the die assembly structures, finally designing dies according to the relation and the case model [1]. How to describe the case and how to retrieve the useful case that are the two most important technologies to CBR system. Auto recognition of the automobile panel surface is the key of retrieval of case in case-based intelligent computer aided die design.

To compare two automobile panel surfaces, first we describe the surfaces in a local coordinate system, then we extract some section curves and the boundary curves from each surface, finally we calculate the similarity between the corresponding curves on the two surfaces separately and get the similarity between the two surfaces based the curve's similarity.

2. PRIMARY INERTIA AXES COORDINATE SYSTEM

There are two kinds of methods to compare two surfaces. The main idea of the first kind of method is to align the two surfaces before we compare them. ICP (Iterative Closet Point), which is proposed by Schütz and Hügli [1], is a representative algorithm to align two surfaces. It is time exhausting to search the orientation in which two surfaces match best. The second kind of method is comprised of two steps, first extracting the eigenvectors from every surface separately, then getting the similarity between the two surfaces by calculating the Euclidean distance of the two surface's eigenvectors. Chua and Jarvis [2] proposed points with the biggest principal curvature. The above methods show their validity, however, obvious shortcomings of the methods on representative ability are still visualized. Johnson and Hebert [3, 4] presented a new method, in which spin–image is used as eigenvector. He estimate the similarity between the new surface and the model one, according to the similarity between the spin–images. Their experiments have get quite good results. But the spin–image method needs large-scale calculation, which is time exhausted.

The primary inertia axes are the surface's intrinsic characteristic, which keep invariant when the surface is translated and rotated. The automobile panel surface often has a 'flat' shape, between whose three primary inertia moments there is a great difference. We define the three primary inertia axes as x, y and z axis separately according to the corresponding primary inertia moments. That means that we assign the primary inertia axis about which the surface has the biggest moment as y axis, and the primary inertia axis about which the surface has the smallest moment as x axis, the left primary inertia axis as z axis. We also

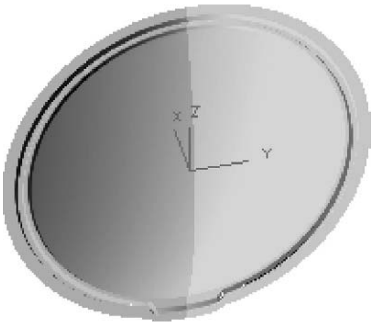


Figure 1. The principal inertia axes coordinate system of a face.

assign the centroid of the surface as the origin point. Thus we get the primary inertia axes coordinate, in which the surface's equation keeps invariant when been translated or rotated. The example of the surface's primary inertia axes coordinate is illustrated in Figure 1.

The angular moment of a 3-D body with respect to the origin of the coordinate system can be expressed as

$$H = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} \cdot \omega \tag{1}$$

Equation (1) shows that the angular moment H might have different direction with the angular velocity ω. There exits a direction in which H is parallel with ω. Under this condition, we can get

$$H = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} \cdot \omega = I\omega \tag{2}$$

From the above equation, we obtain

$$\begin{bmatrix} I_{xx} - I & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} - I & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} - I \end{bmatrix} \cdot \omega = 0 \tag{3}$$

Obviously, the angular velocity ω in Equation (3) is the eigenvector of the inertia moment matrix in equation (1), and ω is one of the primary inertia axes of the 3-D object, I is the object's inertia moment about the axis of rotation. That is to say, we can get a 3-D object's three primary inertia axes though calculating the three eigenvectors of the inertia moment matrix. The three axes and the object's centroid build up the primary inertia axes coordinate system.

3. SCALE INVARIANT MOMENT

There are global and local descriptors to represent the eigenvectors of the curves, for example, Fourier descriptor, auto-regressive model and corner invariant. Invariant moment, which was proposed by Hu [5] is a broadly used method to extract image's eigenvector in image pattern recognition. In this paper, some improvements are made on invariant moment method to get the scale invariant moment, which is used to recognize surfaces' section curves and boundary curve projections. The new method has some excellent features: (1) Each section curve of a surface may consist of many segments, even detached segments. Other methods except for scale invariant moment cannot deal with such condition. (2) Surfaces will keep invariant when they are translated or rotated in the primary inertia axes coordinate system, so section curves need the kind of descriptor, which keeps invariant only when the curves are scaled. Scale invariant moment is designed for such request. (3) Section curves are open curves and boundary curves are close curves. Invariant moment method can deal with open curves and close curves with the same steps, which make programming easier.

There are two curves with similar shape but different size, the (p, q) degree of moment relative to origin of the segment between a and b is

$$M_{ab}^{p,q} = x_a^p y_a^q s_{ab} \approx x_a^p y_a^q |\, ab \,| \tag{4}$$

The (p, q) degree of moment relative to origin of the segment between c and d is

$$M_{cd}^{p,q} = x_c^p y_c^q s_{ab} \approx x_c^p y_c^q |\, ab \,| \tag{5}$$

Supposing that the proportion between the two curves' size in Figure 2 is r, that is to say, $r = |\, oa \,| / |\, oc \,|$, we can get the proportion value between the two

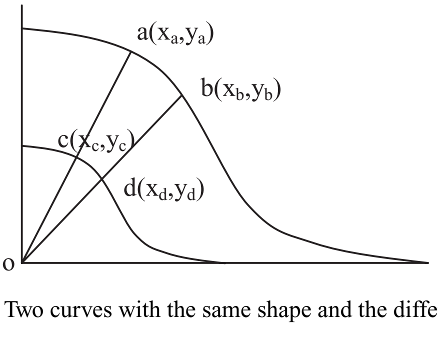


Figure 2. Two curves with the same shape and the different size.

moments relative to the origin

$$\frac{M_{ab}^{p,q}}{M_{cd}^{p,q}} = \left(\frac{x_a}{x_c}\right)^p \cdot \left(\frac{y_a}{y_c}\right)^q \cdot \frac{|\,ab\,|}{|\,cd\,|} = r^{p+q+1}. \tag{6}$$

Obviously, if normalizing the moments relative to the origin with r^{p+q+1}, the scale invariant moment can be written as follows

$$M^{p,q} = \frac{x^p y^q \Delta s}{\sqrt{x^2+y^2}^{p+q+1}}. \tag{7}$$

The scale invariant moment of the whole curve is

$$M^{p,q} = \sum_{i=1}^{n-1} \frac{x_i^p y_i^q \sqrt{(x_{i+1}-x_i)^2 + (y_{i+1}-y_i)^2}}{\sqrt{x_i^2+y_i^2}^{p+q+1}} \tag{8}$$

The above formula gives a moment, which keeps invariant when the curve is scaled and does not need the points locating on the curve with proportional spacing.

In this present paper, the authors calculate seven scale invariant moments for each curve as the eigenvector, the selected moments are $M^{1,2}$, $M^{2,2}$, $M^{1,4}$, $M^{3,2}$, $M^{2,4}$, $M^{4,2}$ and $M^{1,6}$. Generally, Euclid distance between two eigenvectors is used to evaluate the similarity between two curves.

$$D = \sqrt{\sum (f_1(i) - f_2(i))^2} \tag{9}$$

The Euclidean distance between two eigenvectors usually does not keep a linear relation with people's sense about the difference between two curve shapes. This is an open issue in the sense science field. In this paper, the following formula is defined to calculate similarity between curves:

$$S_{curve} = e^{-D^2/scale}. \tag{10}$$

The variable 'scale' in Equation (10) is a coefficient, which is adjusted to make S_{curve} as close to people's sense about the difference between curves as possible.

Figure 3 illustrates a tank shell's section curves and its boundary curve projection. Figure 4 shows a beam shell's section curves and its boundary curve projection. We select four section curves from the tank shell: 'ys1,' 'ys2,' 'ys3' and 'ys4', the boundary curve projection: 'ye.' We select four section curves from the beam shell: 'ls1,' 'ls2,' 'ls3' and 'ls4,' the boundary curve projection: 'le.' The similarity between the selected curves is calculated with Equation (10) (scale = 1) and illustrated in Table 1. It is seen from Table 1 that the result is consistent with the similarity between the curves' shape, which shows that

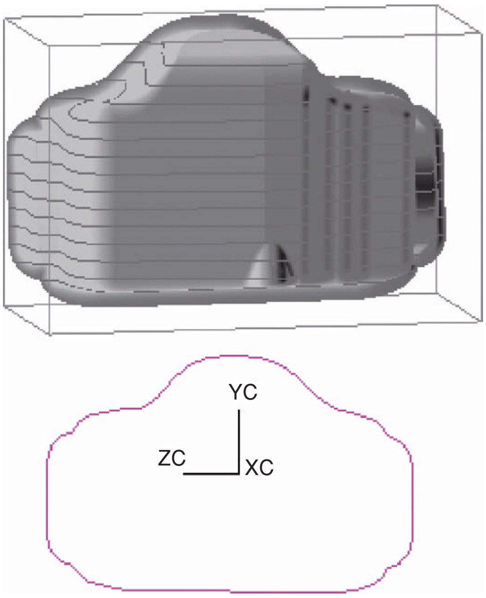


Figure 3. Tank shell’s sections and outline projection.

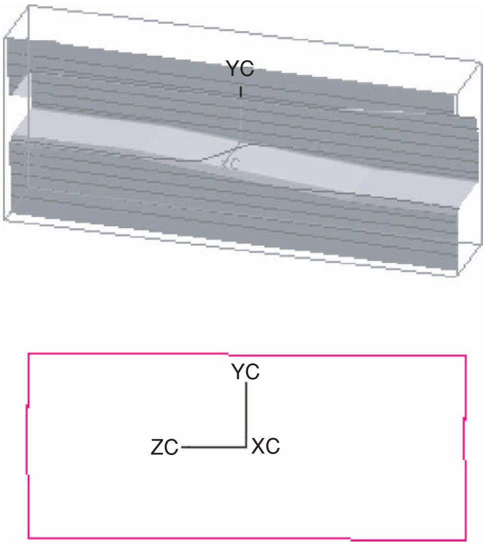


Figure 4. Beam shell’s sections and outline projection.

Table 1. The similarity between the selected curves.

Similarity	Ys1	ys2	ys3	ys4	ye	ls1	ls2	ls3	ls4
ys2	.43	–	–	–	–	–	–	–	–
ys3	.49	.87	–	–	–	–	–	–	–
ys4	.51	.84	.96	–	–	–	–	–	–
Ye	.24	.37	.38	.39	–	–	–	–	–
ls1	.04	.07	.07	.08	.18	–	–	–	–
ls2	.01	.02	.02	.02	.02	.01	–	–	–
ls3	.11	.24	.21	.21	.45	.08	.01	–	–
ls4	.07	.17	.15	.14	.28	.05	.01	.63	–
Le	.24	.37	.37	.39	.58	.18	.02	.45	.28

the scale invariant moment method is a valid method to evaluate the curves' similarity.

In the primary inertia axes coordinate system, a series of planes located alone the y-axis are evenly selected (14 planes are appointed in this paper), which are perpendicular to the y-axis. When these planes intersect with the surface, a set of section curves and the boundary curve projection are obtained by projecting the surface's boundary curve to the y–z plane. Examples of section curves and boundary curve projections are illustrated in Figure 3 and Figure 4.

The similarity between two surfaces is calculated based on the similarity between the corresponding section curves and boundary curve projections belonging to the two surfaces separately.

$$S_{face} = \frac{1}{15}\left(S_{curve}(proj_1 + proj_2) + \sum_{i=1}^{14} S_{curve}(\sec_{1i}, \sec_{2i})\right) \tag{11}$$

One often gives more emphases on some critical section curves. Under such circumstance, we can assign different weights to each section curves and get the new similarity equation:

$$S_{face} = \frac{1}{15}\left(w_{proj} \cdot S_{curve}(proj_1 + proj_2) + \sum_{i=1}^{14} w_i \cdot S_{curve}(\sec_{1i}, \sec_{2i})\right) \tag{12}$$

The similarity between the corresponding section curves belonging to the tank shell and the beam shell separately is shown in Table 2.

Table 2. The similarity between the corresponding section curves.

Section curves	1	2	3	4	5	6	7	8	9	10	11	12	13	14
imilarity	.11	.24	.21	.21	.08	.02	.19	.14	.15	.02	.09	.27	.33	.32

The similarity between the tank's boundary curve projection and the beam's boundary curve projection is 0.58. We can get the similarity between the tank shell and beam shell is $S_{face} = 0.19$.

4. CONCLUSIONS

Free form surface recognition is a critical technology in the case-based intelligent CAD for automobile panel. The primary inertia axes coordinate system, the scale invariant moment and the method to recognize free form surface based on the section curve identification are presented in this paper. The calculation examples have gotten good results. The method proposed here is also useful in other fields where surface recognition is needed.

ACKNOWLEDGEMENTS

This work was all together supported by the Key Project of National Natural Science Foundation of China (No.19832020).

REFERENCES

1. Ch. Schütz and H. Hügli (1995), Recognition of 3-D objects with a closet point matching algorithm, *Proc. Conf. ISPRS*, Zurich, 30, 128–133.
2. C. Chua and R. Jarvis (1996), 3-D free-form surface registration and object recognition. *IJCV*, 17, 77–99.
3. A.E. Johnson and M. Hebert (1999), Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Trans. on pattern analysis and machine intelligence*, 21, 5, 433–449.
4. A. Johnson and M. Hebert (1998), Control of polygonal mesh resolution for 3-D computer vision. *Graphical Models and Image Processing*, 60, 261–285.
5. M.K. Hu (1962), Visual pattern recognition by moment invariants. *IRE Trans. Inf. Theory IT*-8, 179–187.

DYNAMIC CLUSTERING ALGORITHM BASED ON ADAPTIVE RESONANCE THEORY

D.X. Tian[1], Y.H. Liu[1] and J.R. Shi[2]
[1]*College of Computer Science and Technology, Jilin University, Changchun 130012, China*
[2]*Automation Department, Jilin Chemical Institute of Technology, Jilin 132022, China*

Abstract Artificial neural network can be categorized according to the type of learning, that is, supervised learning versus unsupervised learning. Unsupervised learning can find the major features of the origin data without indication. Adaptive resonance theory can classify large various data into groups of patterns. Through analysing the limit of adaptive resonance theory, a dynamic clustering algorithm is provided. The algorithm not only can prevent from discarding irregular data or giving rise to dead neurons but also can cluster unlabelled data when the number of clustering is unknown. In the experiments, the same data are used to train the adaptive resonance theory network and the dynamic clustering algorithm network. The results prove that dynamic clustering algorithm can cluster unlabelled data correctly.

Keywords: artificial neural network, adaptive resonance theory, dynamic clustering, unsupervised learning.

1. INTRODUCTION

Artificial neural networks provide a neurocomputing approach for solving complex problems that might otherwise not have a tractable solution. Applications of neural networks include prediction and forecasting, associative memory, function approximation, clustering, data compression, speech recognition and synthesis, nonlinear system modelling, nonlinear control, pattern classification, feature extraction, combinatorial optimization, solution for matrix algebra problems, blind source separation and solution of differential equations. Artificial neural networks have the ability to learn from their environment and to adapt to it in an interactive manner similar to their biological counterparts. The ability to learn by example and generalize are the principal characteristics of

G. R. Liu et al. (eds.), Computational Methods, 1239–1248.

artificial neural networks. A neural network is trained by presenting several patterns that the network must learn according to a learning rule. The knowledge that has been learned during the training process is stored in the synaptic weights of the neurons.

Neural networks can be classified according to the type of learning, that is, supervised learning and unsupervised learning. Unsupervised learning involves the frequent modification of the network's synaptic weights in response to a set of input samples. The weight modifications are carried out in accordance with a set of learning rules. After repeated applications of these samples to the network, a configuration emerges that is of some significance. Competitive learning is the main method of unsupervised learning. A special of self-organizing neural network is based on competitive learning. In competitive learning networks, the output neurons compete among themselves to determine a winner. There are three basic types of neural networks: the Kohonen self-organizing map (SOM) [1, 2], learning vector quantization (LVQ) [3] and adaptive resonance theory (ART) [4, 5] networks. In conventional competitive learning, when initial conditions are far from the final solutions, some neurons tend to be underutilized, that is, dead neurons. The problem of these dead neurons is one of the serious problems in conventional competitive learning. Therefore, there have been many attempts to solve the dead neuron problem [6–12]. To solve this problem and cluster correctly, this paper proposes a dynamic clustering algorithm, which is based on ART.

2. ASSOCIATIVE LEARNING

An association is any link between a system's input and output such that when a pattern A is presented to the system it will respond with pattern B. When two patterns are linked by an association, the input pattern is often referred to as the stimulus. Likewise, the output pattern is referred to as the response. Two basic rule of associative learning is recognition rule and recall rule.

2.1 Recognition rule

A recognition neuron is presented in Figure 1. The input/output expression for the neuron is

$$a = f(\mathbf{wp} + b) \tag{1}$$

The neuron will be active whenever the product between the weight vector and the input is greater than or equal to $-b$. For two vectors of constant length, the product will be largest when they point in the same direction. Based on

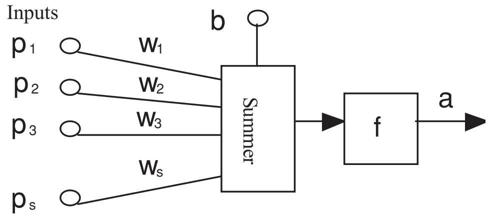


Figure 1. Recognition neuron.

these arguments, the neuron will be active when **p** is close to **w**. By setting the bias b appropriately, we can select how close the input vector must be to the weight vector in order to activate the neuron. The larger the value of b, the more patterns there will be that can activate the neuron, thus making it the less discriminatory. Based on the Hebbian postulate the leaning rule is:

$$w_i(n) = w_i(n-1) + \alpha a(n) p_i(n) \tag{2}$$

To get the benefits of weight decay, while limiting the forgetting problem, a decay term that is proportional to $a(n)$ is added:

$$w_i(n) = w_i(n-1) + \alpha a(n) p_i(n) - \gamma a(n) w_i(n-1) \tag{3}$$

by setting γ equal to α and when the neuron is active $a(n) = 1$, Equation 3 can then be written

$$w_i(n) = (1-\alpha) w_i(n-1) + \alpha p_i(n) \tag{4}$$

in vector form:

$$\mathbf{w}(n) = (1-\alpha)\mathbf{w}(n-1) + \alpha \mathbf{p}(n) \tag{5}$$

When the neuron is active, the weight vector is moved towards the input vector along a line between the old weight vector and the input vector. The distance the weight vector moves depends on the value of α.

2.2 Recall rule

The recall neuron is shown in Figure 2. The input–output expression for this neuron is

$$\mathbf{a} = f(\mathbf{w}p) \tag{6}$$

where the input p is scalar and the output **a** is vector. It can perform recall by associating a stimulus with a vector response. In the recognition rule, forgetting is limited by making the weight decay term of the Hebbian rule proportional

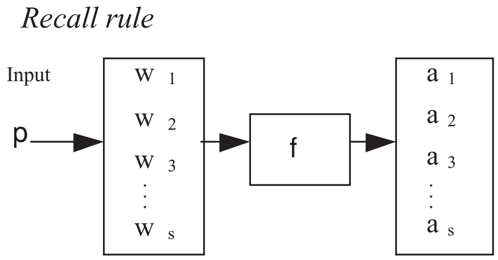


Figure 2. Recall neuron.

to the output of the neuron. Conversely, the recall rule makes the weight decay term proportional to the input of the neuron:

$$w_i(n) = w_i(n-1) + \alpha a_i(n) p(n) - \gamma p(n) w_i(n-1) \tag{7}$$

If the decay rate γ is set equal to the learning rate α

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \alpha(\mathbf{a}(n) - \mathbf{w}(n-1))p(n) \tag{8}$$

In recall neuron, learning occurs whenever p is nonzero. When learning occurs, weight vector **w** moves towards the output vector.

3. COMPETITIVE LEARNING

Competitive learning comprises one of the main classes of unsupervised artificial neural networks, where only one neuron or a small group of neurons, called winning neurons, are activated according to the degree of proximity of their weight vectors to the current input vector. This type of algorithm is used in tasks of pattern recognition and classification, such as clustering and vector quantization. In these applications, the weight vectors are called the prototypes of the set input patterns. The basic structure of competitive neuron is shown in Figure 3. The basic competitive algorithm is Winner-Take-All (WTA). In WTA type learning, after an input sample is presented, only one neuron (the winner)

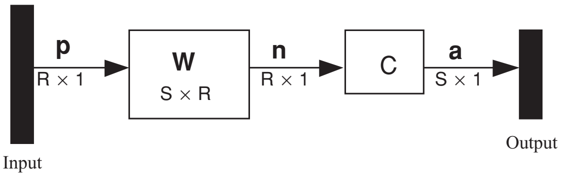


Figure 3. Competitive neuron.

in the competitive network will remain active (or switch on) after a number of iterations (or a single iteration, depending on which type of network is used) and its corresponding long-term memory will be updated.

$$a_i = \begin{cases} 1, & i = i^* \\ 0, & i \neq i^* \end{cases} \tag{9}$$

where i^* is the winner neuron.

The modifications of WTA are frequency sensitive competitive learning (FSCL) and rival penalized competitive learning (RPCL) [10, 11, 13]. FSCL's idea is 'the winner neuron learns, the losers are frozen':

$$a_i = \begin{cases} 1, & \text{if } d(\mathbf{x}, \mathbf{w}_i) = \min(d(\mathbf{x}, \mathbf{w}_k)) \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

where $d(\,)$ is the distance function. RPCL is 'the winner and the first loser learn, the remaining neurons are frozen':

$$a_i = \begin{cases} 1, & \text{if } d(\mathbf{x}, \mathbf{w}_i) = \min(d(\mathbf{x}, \mathbf{w}_k)) \\ -\beta & \text{if } d(\mathbf{x}, \mathbf{w}_j) = \min(d(\mathbf{x}, \mathbf{w}_k)) \text{ and } k \neq i \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

4. DYNAMIC CLUSTERING ALGORITHM

4.1 Adaptive resonance theory

Adaptive resonance theory (ART), developed by Carpenter and Grossberg, overcomes the stability/plasticity dilemma by accepting and adapting the stored prototype of a category only when the input is sufficiently similar to it. When an input pattern is not sufficiently similar to any existing prototype, a new category is formed with the input pattern as the prototype using a previously uncommitted output unit. There are three basic types of unsupervised ART networks: ART1 (binary-valued input vectors), ART2 (continuous-valued input vectors), and fuzzy ART (both binary and continuous-valued input vectors).

The basic ART architecture is shown in Figure 4. ART consist of three parts: Layer 2 to Layer 1 expectations, the orienting subsystem and gain control. The equation of operation of Layer 1:

$$\varepsilon \frac{d\mathbf{n}^1(t)}{dt} = -\mathbf{n}^1(t) + ({}^{+}\mathbf{b}^1 - \mathbf{n}^1(t))(\mathbf{p} + \mathbf{W}^{21}\mathbf{a}^2(t)) - (\mathbf{n}^1(t) + {}^{-}\mathbf{b}^1)({}^{-}\mathbf{W}^1\mathbf{a}^2(t)) \tag{12}$$

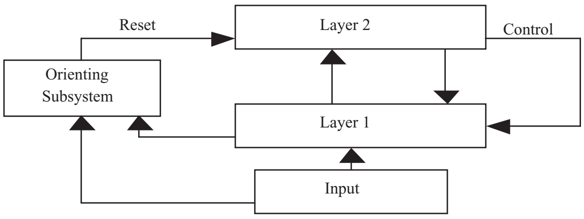


Figure 4. ART architecture.

The learning rule for weight matrix is

$$\frac{d\mathbf{w}_j^{21}(t)}{dt} = \mathbf{a}_j^2(t)(-\mathbf{w}_j^{21}(t) + \mathbf{a}^1(t)) \tag{13}$$

The equation of operation of Layer 2 is

$$\begin{aligned} \varepsilon\frac{d\mathbf{n}^2(t)}{dt} = &- \mathbf{n}^2(t) + ({}^{+}\mathbf{b}^2 - \mathbf{n}^2(t))({}^{+}\mathbf{W}^2\mathbf{f}^2(\mathbf{n}^2(t)) + \mathbf{W}^{12}\mathbf{a}^1) \\ &- (\mathbf{n}^2(t) + {}^{-}\mathbf{b}^2)({}^{-}\mathbf{W}^2\mathbf{f}^2(\mathbf{n}^2(t))) \end{aligned} \tag{14}$$

The learning rule for weight matrix is

$$\frac{d({}_i\mathbf{w}^{12}(t))}{dt} = \mathbf{a}_i^2(t)(k({}^{+}\mathbf{b} - {}_i\mathbf{w}^{12}(t))({}^{+}\mathbf{W}\mathbf{a}^1(t)) - ({}_i\mathbf{w}^{12}(t) + {}^{-}\mathbf{b})({}^{-}\mathbf{W}\mathbf{a}^1(t))) \tag{15}$$

One of the key elements of the ART architecture is the Orienting Subsystem. Its purpose is to determine if there is a sufficient match between the L2-L1 expectation and the input pattern. When there is not enough of a match, the Orienting Subsystem should send a reset signal to Layer 2. The reset signal will cause a long-lasting inhibition of the previous winning neuron, and thus allow another neuron to win the competition. The equation of operation of the Orienting Subsystem is

$$\varepsilon\frac{dn^0(t)}{dt} = -n^0(t) + ({}^{+}b^0 - n^0(t))({}^{+}\mathbf{W}^0\mathbf{p}) - (n^0(t) + {}^{-}b^0)({}^{-}\mathbf{W}^0\mathbf{a}^1) \tag{16}$$

4.2 Dynamic clustering algorithm

In ART learning, after an input sample is presented, only one neuron (the winner) will remain active after a number of iterations and its corresponding long-term memory will be updated. It has been known for a long time that some long-term memory of the output neurons simply cannot learn. These neurons are usually referred to as dead neurons. It is found that a sensitivity or conscience parameter λ (which can be implemented as a winner frequency counter) can be included to prevent such dead neurons from occurring. The sensitivity parameter counts the number of times that a certain neuron wins. If that node wins more frequently than other neurons, it will reduce its chances from winning in future iterations. Thereby, giving other less frequently won neurons opportunities to learn.

Unfortunately, dead neurons may still arise if the number of output neurons is more than the number of clusters or the neurons are unfairly initialized. In these cases, the dead neurons will usually position themselves near the boundaries of the desired clusters or somewhere around their initial location. And some times the sensitivity parameter leads to wrong results because it destroys the normal learning. To solve these problems, this paper proposes a dynamic clustering algorithm, which add neuron when the input sample is not sufficiently similar to any existing neuron and delete neuron whose wins time below a threshold value θ. The dynamic clustering algorithm is

Step 1. Initialize the weight matrices $\mathbf{w}_{ij}^{21} = 1$, $\mathbf{w}_{ij}^{12} = \frac{k}{k+S1-1}$;
Step 2. Present an input pattern to the network $a_i^1 = p_i$;
Step 3. Compute the input to Layer 2, $n_j = ({}_j\mathbf{w}^{12})^T \mathbf{a}^1$;
Step 4. Activate the neuron in Layer 2 with the Winner-Take-All rule,

$$a_i^2 = \begin{cases} 1, & \text{if } (({}_i\mathbf{w}^{12})^T \mathbf{a}^1 = \max(n_j)) \\ 0, & \text{else} \end{cases},$$

if no matching neuron, go to Step 10;
Step 5. Compute the L2-L1 expectation, $\mathbf{w}_j^{21} = \mathbf{W}^{21}\mathbf{a}^2$;
Step 6. Adjust the Layer 1 output to include the L2-L1 expectation, $\mathbf{a}^1 = \mathbf{p} \cap \mathbf{w}_j^{21}$;
Step 7. The Orienting Subsystem determines the degree of match between the expectation and the input sample,

$$a^0 = \begin{cases} 1, & \text{if } \sum_{i=1}^{S1} a_i^1 \Big/ \sum_{i=1}^{S1} p_i < \rho \\ 0, & \text{else} \end{cases};$$

Step 8. If $a^0 = 1$, then set $a_j^2 = 0$, inhibit it until an adequate match occurs, go to Step 2, else continue;
Step 9. The win times of neuron j is added and update column j of $\mathbf{W}^{21}$ and row j of $\mathbf{W}^{12}$, $\mathbf{w}_j^{21} = \mathbf{a}^1$,

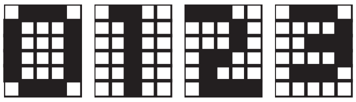
$${}_j\mathbf{w}^{12} = \frac{(1-\alpha)_j\mathbf{w}^{12}\ old + \alpha\mathbf{p}}{\sum_{i=1}^{S1} (1-\alpha)_j w_i^{12}\ old + \alpha p_i},$$

go to Step 11;

Step 10. Add a new neuron, $\mathbf{W}^{21}$ add one column and $\mathbf{W}^{12}$ add one row, $\mathbf{w}_p^{21} = \mathbf{p}$, ${}_p\mathbf{w}^{12} = 1/\sum_{i=1}^{S1} p_i^1$;

Step 11. Restore all inhibited neurons in Layer 2, and return to Step 2 with a new input sample.

5. EXPERIMENTS



The patterns we want to train are shown to the below. They represent the digits $\{0, 1, 2, 3\}$ displayed in a 6×5 grid. We need to convert these digits to vectors, each white square will be represented by a '0', and each dark square will be represented by a '1'. Then, to create the input vectors, we will scan each 6×5 grid one column at a time. For example, the digit '0' will be $p_0 = \{011101000110001100011000101110\}$.

First, train an ART1 network and present the vectors in the order 0-1-2-3. Use the learning law parameters $\xi = 2$, vigilance $\rho = 0.6$ and category parameter $S^2 = 3$. We begin by initializing the weight matrices. The initial $\mathbf{W}^{21}$ matrix is an 30×3 matrix of 1's. The initial $\mathbf{W}^{12}$ matrix is normalized, therefore it is an 3×30 matrix, with each element equal to $\xi/(\xi + S^1 - 1) = 2/(2 + 30 - 1) = 0.0645$. The training results are shown in Figure 5.

Second, use the same sample and present in the same order to dynamic clustering network. The vigilance is set 0.6 and the training results is shown in Figure 6.

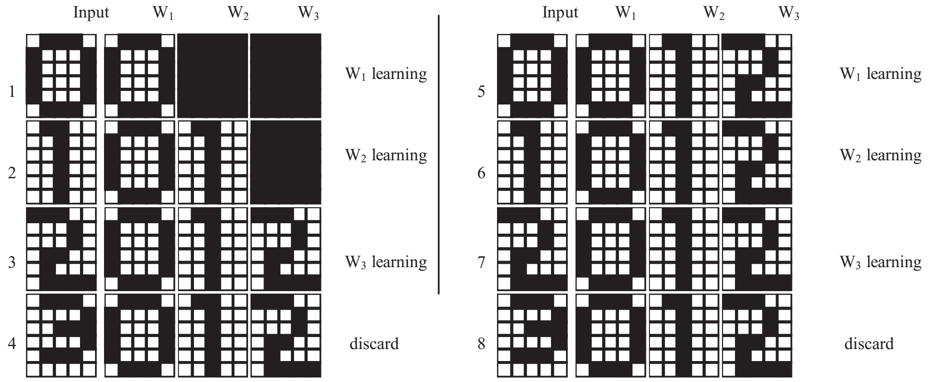


Figure 5. The results of ART.

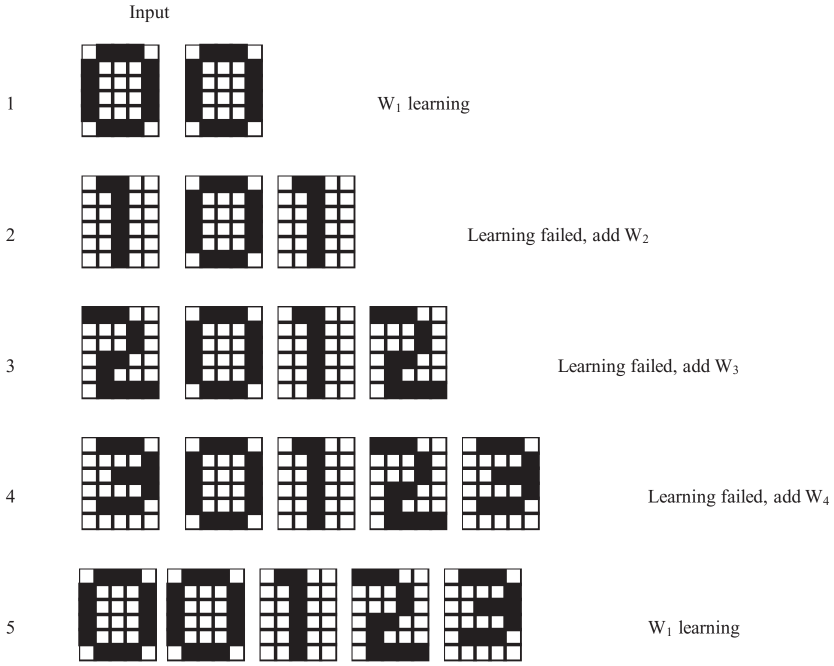


Figure 6. The results of dynamic clustering.

We can find that in ART the digit ‘3’ is covered by digit ‘2’, but in dynamic clustering all the digits are learned.

6. CONCLUSIONS

In this paper, a dynamic clustering algorithm based on adaptive resonance theory is proposed. The algorithm uses unsupervised learning and updates the weight vectors by associative learning rule. To overcome the problem of dead neurons, the dynamic clustering algorithm add new neurons when the input sample is not sufficiently similar to any existing neuron and delete neuron whose wins time below a threshold value θ. The algorithm makes the clustering of unlabelled data more correctly. Finally, the algorithm is applied to train digits and compare it with ART network. The results prove that dynamic clustering algorithm can correctly clustering and prevent from giving rise to dead neurons.

REFERENCES

1. H. Ressom, D. Wang and P. Natarajan (2003), Adaptive double self-organizing maps for clustering gene expression profiles. *Neural Network*, 16, 633–640.

2. B. Hammer, A. Micheli, A. Sperduti and M. Strickert (2004), A general framework for unsupervised processing of structured data. *Neurocomputing*, 57, 3–35.
3. K.L. Wu and M.S. Yang (2003), A fuzzy-soft learning vector quantization. *Neuroncomputing,* 55, 681–697.
4. W.K. Fung and Y.H. Liu (2001), A game-theoretic adaptive categorization mechanism for ART-type networks. *Lecture Notes in Computer Science*, 2130, 170–176.
5. W.K. Fung and Y.H. Liu (2003), Adaptive categorization of ART networks in robot behavior learning using game-theoretic formulation. *Neural Networks*, 16, 10, 1403–1420.
6. R. Kamimura (2003), Competitive learning by information maximization: Eliminating dead neurons in competitive learning. *Lecture Notes in Computer Science,* 2714, 99–106.
7. L. Andrew (2000), Analyses on the generalized lotto-type competitive learning. *Lecture Notes in Computer Science,* 1983, 9–16.
8. M. Stephen, S. Jonathan and N. Ulrich (2002), A self-organising network that grows when required. *Neural Networks*, 15, 1041–1058.
9. I. Hirotaka and N. Hiroyuki (2003), Efficiency of self-generating neural networks applied to pattern recognition. *Mathematical and Computer Modelling*, 38, 1225–1232.
10. T.M. Nair, C.L. Zheng and J.L. Fink (2003), Rival penalized competitive learning (RPCL): A topology-determining algorithm for analyzing gene expression data. *Computational Biology and Chemistry,* 27, 565–574.
11. Z.Y. Liu, K.C. Chiu and L. Xu (2003), Local PCA for line detection and thinning. *Lecture Notes in Computer Science*, 2683, 21–34.
12. M.A. Garcia-Bernal, J. Munoz-perez, J.A. Gomez-Ruiz and I. Ladron de Guevara-lopez (2003), A competitive neural network based on dipoles. *Lecture Notes in Computer Science*, 2686, 398–405.
13. G. Acciani, E. Chiarantoni, G. Fornarelli and S. Vergura (2003), A feature extraction unsupervised neural network for an environment data set. *Neural Networks*, 16, 427–436.

ONTOLOGY LEARNING USING WORDNET LEXICON

H. Hu[1,2], X.Y. Du[1], D.Y. Liu[2], and J.H. Ouyang[2]
[1]*School of Information, Renmin University of China, Beijing, China 100872*
[2]*Open Symbol Computation and Knowledge Engineering Laboratory of State Education Committee, College of Computer Science and Technology, Jilin University, Changchun, China 130012*

Abstract Current Semantic Web community has popularized ontology research. However, ontology building by hand has proven to be a very hard and error-prone task and become the bottleneck of ontology acquiring process. WordNet, an electronic lexical database, is considered to be the most important resource available to researchers in computational linguistics. The paper proposes an ontology learning approach, which uses WordNet lexicon resources to build a standard OWL ontology model. The approach will help the automation of ontology building and be very useful in ontology-based applications.

Keywords: ontology learning, WordNet, OWL.

1. INTRODUCTION

Ontologies in fact turn out to be the backbone technology for the Semantic Web; Tim Berners-Lee [1] has portrayed Semantic Web as a layered architecture where ontology layer lies in the middle of the other layers. Research in ontologies to date has mainly addressed the basic principles, such as knowledge representation formalisms, devoting only limited attention to more practical issues such as techniques and tools aimed at ontology's actual construction and acquisition. We propose an ontology learning approach in this paper, which uses WordNet lexicon resources to build a standard OWL ontology. The approach will help the automation of ontology building and will be very useful in ontology-based applications.

The paper is organized as follows. We first give some preliminary knowledge about ontologies, OWL and WordNet system in Section 2, then describe how

G. R. Liu et al. (eds.), Computational Methods, 1249–1253.

we can use the ontology mapping approach to get OWL ontologies with the help of WordNet lexicon in Section 3. Section 4 discusses our overall framework for ontology learning. After an analysis of current related works on the subject, we conclude our research work and give some ideas for future works in the last section.

2. PRELIMINARY

2.1 Ontology approach and OWL language

Tom Gruber [2] has defined ontology as ‘a specification of a conceptualization’. Ontologies provide a deeper level of meaning by providing equivalence relations between concepts; they can standardize meaning, description, representation of involved concepts, terms and attributes; capture the semantics involved via domain characteristics, resulting in semantic metadata and ‘ontological commitment’ which forms basis for knowledge sharing and reuse. Ontologies can provide a domain theory using an expressive language for capturing the domain. One of the properties of ontologies is that all relevant knowledge has been made explicit; this constitutes in the necessity of specifying many relationships that are otherwise left implicit and are only made explicit in the applications developed for working with the ontology.

A number of ontology definition languages have been developed over the past years. Among them, the Web Ontology Language (OWL) [3] is the newly emerging standard proposed and supported by W3C for defining ontologies in *Semantic Web*. It is based on description logic, a subset of first-order logic that provides sound and decidable reasoning support. The OWL Web Ontology Language is designed for use by applications that need to process the content of information instead of just presenting information to humans. OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF, and RDF Schema (RDF-S) by providing additional vocabulary along with a formal semantics.

2.2 WordNet system

WordNet is an on-line lexical database, which was developed at the Cognitive Science Laboratory at Princeton University under the direction of George Miller [4]. The design of WordNet is inspired by current psycholinguistic theories of human lexical memory. WordNet is considered to be the most important resource available to researchers in computational linguistics, text analysis and many related areas. Its design is inspired by current psycholinguistic and computational theories of human lexical memory. English nouns, verbs, adjectives and adverbs are organized into synonym sets, each representing one underlying lexicalized concept. Different relations link the synonym sets including: antonymy, hypernymy, hyponymy, holonymy, meronymy, synonymy, troponymy etc. A web interface to WordNet is available at: http://www.cogsci.princeton.edu/cgi-bin/webwn.

3. A MAPPING FROM WORDNET TO OWL

WordNet system uniquely identifies a word sense in two ways: with a set of terms called synset and a textual definition called gloss. For example, for the third sense of ‘transport’, the synset list would consist of the words ‘transportation’, ‘shipping’, and ‘transport’. The gloss textual definition of that third sense would be ‘the commercial enterprise of transporting goods and materials’. WordNet codes other types of semantic relations as well, such as kind-of, part-of and several types of similarity relations. Based on the analysis above, we give an illustration of the meta-model of WordNet system in OWL format.

```
<owl:Class rdf:ID="Synset">
</owl:Class>
<owl:Property rdf:ID="synsetRelation">
    <rdfs:domain rdf:resource="#Synset"/>
    <rdfs:range rdf:resource="#Synset"/>
</owl:Property>
<owl:Property rdf:ID="antonym">
    <rdf:type rdf:resource="&owl; SymmetricProperty"/>
    <rdfs:subPropertyOf rdf:resource="#synsetRelation"/>
</owl:Property> ......
<owl:FunctionalProperty rdf:ID="typeOfSynset">
    <rdfs:domain rdf:resource="#Synset"/>
    <rdfs:range rdf:resource="#TypeOfSynset"/>
</owl:FunctionalProperty>
```

Synset, typeOfSynset and word of WordNet are defined as OWL concepts; antonymy, hypernymy, hyponymy, holonymy, meronymy etc. are defined as OWL properties between synsets. These properties have different characteristics: antonymy is symmetric; other properties such as hypernymy, hyponymy, holonymy and meronymy are transitive. These characteristics encoded in OWL can support ontology reasoning tasks. The initialized elements of the mapping process are words, for every word in the input word set, we lookup WordNet lexicon through calling API and get the semantic resources including its synsets, antonymy, hypernymy, hyponymy, holonymy, meronymy etc., then An OWL ontology definition will be generated based on the meta-model above. The overall architecture will be presented in the next section.

4. AN ONTOLOGY LEARNING FRAMEWORK

The input for the framework is domain or application free texts corpus; the framework uses a language analyzer to extract terminology from the corpus, the linguistic knowledge such as the grammar, morphological rules and some syntactic and semantic templates are used to do the natural language processing in this process. Words or lexical tokens will be generated after the natural language processing. Next, we use the WordNet lexical knowledge bases to retrieve semantic concepts and relations of the terms. Based on the mapping described in the last section, the words or lexical tokens are mapped into OWL ontologies with the help of WordNet lexicon. It is a domain/application independent framework and can learn lexical and ontological knowledge for general and specific domains.

5. RELATED WORKS

Alessandro Lenci *et al.* [5] have researched in formalizing the EuroWordNet Synsets and Top Ontology in RDF and writing semantic frames in RDF/S as basis for interlingua representations. However, RDFS has unclear semantics, no clean separation between: Instances, Ontologies and meta-ontologies (e. g. RDFS language itself); moreover, RDFS has no inference model which is of crucial importance for automatic tasks. Our mapping are targeted to OWL, which has clear semantics bringing by description logic systems; OWL can distinguishes between Instances and ontologies etc.; OWL also enjoys a well-founded inference model from some particular description logics (SHOQ(D) [6]. There are many other research works intended to extend WordNet or to achieving a formal specification of WordNet. The focus of paper [7] is to guide and ease the representation, retrieval and sharing of general knowledge; the focus of paper [8] is the extension and axiomatization of conceptual relations in WordNet. Neither of them uses mapping approaches as illustrated in this paper, and their resulting ontologies are not target at OWL.

6. CONCLUSION AND FUTURE WORKS

In this paper, we propose an ontology learning approach and framework based on the mapping from WordNet lexicon to OWL ontologies. Synset, typeOfSynset and word of WordNet are defined as OWL concepts; antonymy, hypernymy, hyponymy, holonymy, meronymy etc. are defined as OWL properties between synsets. This approach will help the automation of ontology building and be very useful in ontology-based applications. We plan to study Chinese ontology learning with the support of Chinese Concept Dictionary (Liu *et. al.*, 2002) in our future works.

REFERENCES

1. T.B. Lee, J. Hendler and O. Lasilla (2001), The Semantic Web. *The Scientific American*, 284, 5, pp. 34–43.
2. T.R. Gruber (1995), Toward principles for the design of ontologies used for knowledge sharing, *International Journal of Human and Computer Studies*, 43, 5/6, pp. 907–928.
3. W3C (2003), "OWL Web Ontology Language Overview", http://Web.w3.org/TR/owl-features/.
4. G. Miller (1990), WordNet: An on-line Lexical data-base. *International Journal of Lexicography*, 3, 4, pp. 235–312.
5. N. Calzolari, A. Zampolli and A. Lenci (2002), Towards a Standard for a Multilingual Lexical Entry: The EAGLES/ISLE Initiative, LNCS 2276, Springer, pp. 264–279.
6. I. Horrocks and U. Sattler (2001), Ontology Reasoning in the SHOQ(D) Description Logic, *Proceedings of the seventeenth International Joint Cenference on Artificial Intelligence*, pp. 199–204.
7. P. Martin (2003), Correction and extension of wordNet 1.7, 11th International Conference on Conceptual Structure (ICCS), LNAI 2746, pp. 160–173.
8. G. Aldo, N. Roberto and V. Paola (2003), The Onto WordNet project. Extension and axiomatization of conceptual relation in WordNet, LNCS 2888, pp. 820–838.

GENETIC PROGRAMMING FOR MAXIMUM-LIKELIHOOD PHYLOGENY INFERENCE

H.Y. Lv, C.G. Zhou and J.B. Zhou
College of Computer Science and Technology, Jilin University, Changchun, P. R. China 130012

Abstract Phylogeny reconstruction is a difficult computational problem, because the number of possible solutions increases with the number of included taxa. For this reason, phylogenetic inference methods commonly use clustering algorithms or heuristic search strategies to minimize the amount of time spent evaluating non-optimal trees. Even heuristic searches can be painfully slow, especially when computationally intensive optimality criteria such as maximum likelihood are used. I describe here a genetic programming to heuristic searching that can tremendously reduce the time required for maximum-likelihood phylogenetic inference, especially for data sets involving large numbers of taxa, and we confirm its availability by experiments.

Keywords: genetic programming, phylogenetic tree, phylogeny, maximum likelihood.

1. INTRODUCTION

Phylogenetics [1] is a method widely used by biologists to investigate evolutionary pathways followed by organisms currently or previously inhabiting the Earth. Given a data set that contains a number of different species, each with a number of (phenotypic or genetic) attribute values, phylogenetics software constructs phylogenies, which are representations of the possible evolutionary relationships among the given species. A phylogeny is a tree structure: The root of a tree can be viewed as the common ancestor, the leaves are the species, and subtrees are subsets of species that share a common ancestor. Each branching of a into subtrees represents a divergence in one or more attribute values of the species within the two subtrees.

G. R. Liu et al. (eds.), Computational Methods, 1255–1259.

A typical phylogenetics approach uses a deterministic hill climbing methodology to find phylogenies for a given data set, saving one or more ‘best’ phylogenies as the result of the process, where ‘best’ is defined by the specific metric used (in this work, we are using parsimony). The phylogeny-building approach adds each species into the phylogeny in sequence, searching for the best place to add the new species. The search process is deterministic, but different phylogenies may be found by running the search with different random orderings of the species in the data set. For example, a Phylip search for a most parsimonious phylogeny consists of specifying the number of random orderings of the data set (called ‘jumbles’) to search with. Each of these searches proceeds independently, and each will end when the search has added all species into the phylogeny and thereby reached a local optimum. In contrast, an EA [2, 3] search is able to explore multiple phylogenies in parallel, and by recombining these, has the potential to jump from one region of the search space to another.

In this paper, I describe here a genetic programming to heuristic searching that can tremendously reduce the time required for maximum-likelihood phylogenetic inference, especially for data sets involving large numbers of taxa, and we confirm its availability by experiments.

2. ALGORITHM

2.1 Initialization

Here, we give n sequences, and code them from 1 to n as a set. We select two sequences with minimum distance to build a subtree, and delete them from the set, then insert the subtree to set as an element. Repeat the procedure until no element in the set, we build a tree with all sequence. Based the tree, random arrange the leaves, we get some trees. These trees are initial population.

2.2 Fitness

The fitness lies on original value of individual and maximum original value of population. Equation (1) is the original value:

$$f = \sum (D_{ij} - d_{ij}) \times (D_{ij} - d_{ij})/(D_{ij} \times D_{ij}) \quad (i \neq j) \tag{1}$$

D_{ij} is the distance of leaf node i and j, d_{ij} is the path length of i and j. In order to calculate d_{ij}, we use fast minimum-evolution method [4]. If d_{ij} is negative, we evaluate it as 0, because the length of branch could not be zero.

$$Nf = (1.5 - f/\mathrm{Max}) \times 100 \tag{2}$$

In Equation (2), Nf is the fitness value and Max is the maximum original value of the generation.

2.3 Selection

In a generation, one son individual is copy from father generation, which has max fitness, and the others decided by roulette.

2.4 Crossover

We select two trees (individual) from father generation. From tree1, we select a subtree N1 from tree1, and get subtree N2 from tree2, then cross the two subtrees, and gained two new trees. In crossover, we use two methods to select cross node. We get a random value M. If $M \geq 0.5$, we use method 1, otherwise we use method 2.

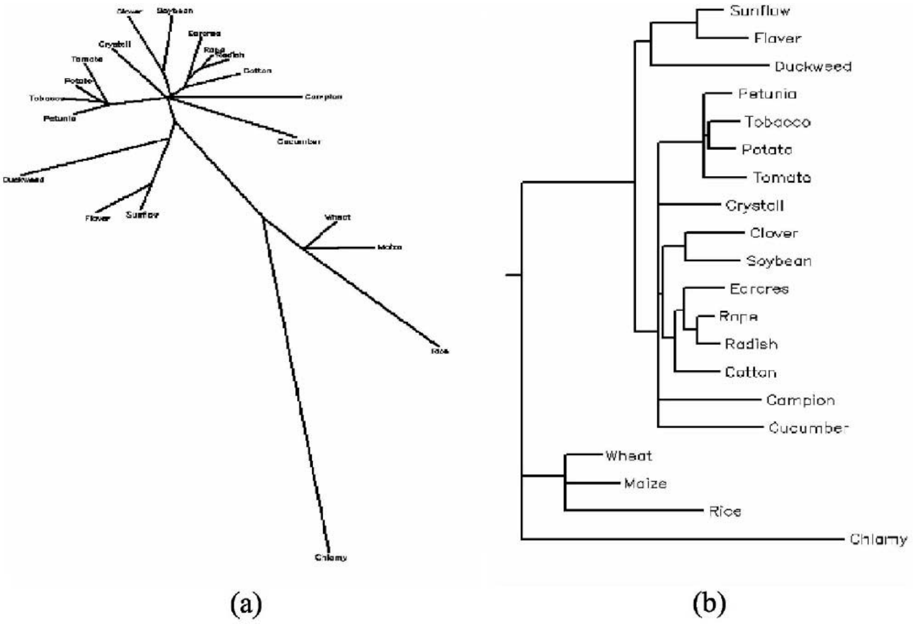
Method 1 is random select node to cross method 2 is depth equality method. In method 2, every node has a value. At the same floor, the values of node are same. We regard the sum value of one floor is W, and the number of floors is L. Firstly, we get a random value between 0 and $W \times L$, then from leaves, add up the value of those node. When the first node which cumulate value is bigger than or equal to the random value, the node is cross node.

For new individual, in order to avoid node repeat or lose, we need adjust them with insert or delete. During delete, we delete the leaf node and his father node, the other subtree (or leaf node) which has the same father node with deleted leaf node, need to connect with his grandfather node. For insert, we find the min distance subtree with the inserted leaf, and use the new subtree replace originally leaf node.

3. RESULT AND DISCUSSION

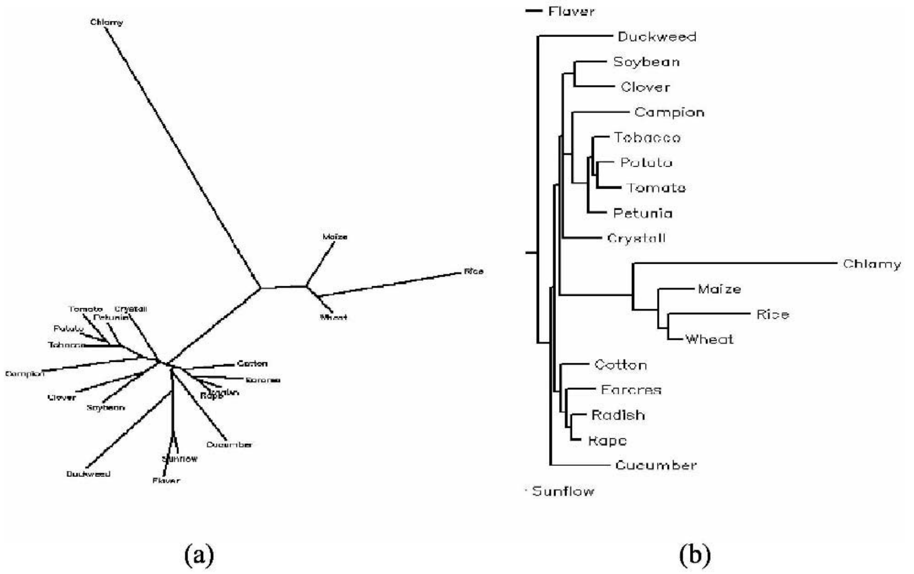
We select 20 amino acid sequences from http://pir.georgetown.edu/pirwww. Sequence alignments run by http://www.ebi.ac.uk/clustalw, and obtain phylogenetic tree by genetic programming.

Population size is 100, and generation is 1000, crossover probability is 75%. We build phylogenetic tree for the populations with maximum-likelihood genetic programming, Fitch–Margoliash and least-squares distance methods. The result is seen from Figures 1 and 2. In every figure, (a) and (b) are all pictures of the tree.



(a) (b)

Figure 1. Maximum-likelihood genetic programming.



(a) (b)

Figure 2. Fitch–Margoliash and least-squares distance methods.

4. CONCLUSIONS

From the above introduction, we can clear to see from it, the process of building phylogenetic tree is to select the best tree from population, and this came from characteristic of GP. Compared with other method, building tree by maximum-likelihood genetic programming can reduce time of the procedure, and the veracity is improved too. And we should research new operators to improve the result of experiment and to avoid becoming trapped in local optima.

REFERENCES

1. R.D.M. Page and E.C. Holmes (1998), *Molecular Evolution: A Phylogenetic Approach*. Blackwell Science Ltd, Oxford.
2. O.L. Paul (1998), A genetic algorithm for maximum-likelihood phylogeny inference using nucleotide sequence data. *Molecular Biology and Evolution*, l, 15, pp. 277–283.
3. C.B. Congdon (2002), Gaphyl: An evolutionary algorithms approach for the study of natural evolution. In: *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO2002)*, MorganKaufmann, San Francisco, CA, pp. 1057–1064.
4. A. Rzhetsky and M. Nei (1993), Theoretical foundation of the minimum-evolution method of phylogenetic inference. *Molecular Biology and Evolution*, 10, pp. 1073–1095.

MINING DOMINANCE ASSOCIATION RULES IN PREFERENCE-ORDERED DATA

Y.B. Liu[1,2], D.Y. Liu[1,2] and Y. Gao[1,2]
[1]*College of Computer Science and Technology, Jilin University, Changchun 130012, China*
[2]*Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China*

Abstract In general, there is a preference order on the domain of an attribute in preference-ordered data, but most data mining approaches ignore it. Such attribute is a criterion. In fact, such rules are useful for prediction: if the mathematics score of a student goes up then his/her physics score will go up. Such rules are called dominance association rules here. Based on criteria, a preference-ordered data table can be transformed into a tri-value data table, in which dominance association rules can be mined. Dominance association rules uncover the correlation between criteria and reflect when values of a set of criteria change, how values of another set of criteria change. One use of dominance association rules is to predict the unknown values of criteria in an object by comparing with other objects.

Keywords: association rule, criterion, dominance association rule

1. INTRODUCTION

In general, there is a preference order on the domain of an attribute (especially of a quantitative attribute) in preference-ordered data, but most data mining approaches ignore it. Such attribute is a criterion. *Association rule* mining designed for basket data [1] mines association rules and further research was carried out aiming at quantitative attributes and categorical attributes [2, 3]. Although association rule mining in preference-ordered data [4] was proposed, the rule form couldn't reflect the correlation between criteria. In fact, such rules are needed for prediction: if the mathematics score of a student goes up then his/her physics score will go up. Such rules are called dominance association rules here, which can be mined from a preference-ordered data table.

G. R. Liu et al. (eds.), Computational Methods, 1261–1266.

They uncover the correlation between criteria and reflect when values of a set of criteria change, how values of another set of criteria change. One use of dominance association rules is to predict the unknown values of criteria in an object by comparing with other objects. We describe the prediction method with an example.

2. PREFERENCE-ORDERED DATA TABLE

Each object in a data table is often described by attributes and attributes' values. The attribute on whose domain there is a preference order is called a criterion, otherwise called a regular attribute. The domains of criteria are called preference scales (increasing or deceasing). For each criterion c, V_c is a preference scale of c. The preference scale V_c of each criterion c induces a preference relation $\geq_c$. For values c_1, $c_2 \in V_c$, $c_1 \geq_c c_2$ means $c_1 >_c c_2$ or $c_1 =_c c_2$, but can't be the both.

Definition 1. A preference-ordered data table $PDT = (U, C)$, where U is a non-empty finite set of objects, C is a non-empty finite set of criteria and for each $c \in C$: $U \to V_c$, $c(x)$ is the value of criterion c in object $x \in U$.

Definition 2. Given a preference-ordered data table $PDT = (U, C), c \in C$, $P \subseteq C$ and $x, y \in U$, a binary dominating relation on U is defined as: $x \gg_c y$ if and only if $c(x) \gg_c c(y); x \gg_P y$ if and only if $x \gg_c y$ for all $c \in P$. A binary dominated relation on U is defined as: $x \ll_c y$ if and only if $c(y) \geq_c c(x)$; $x \ll_P y$ if and only if $x \ll_c y$ for all $c \in P$.

3. MINING DOMINANCE ASSOCIATION RULES

Definition 3. $\phi \to \psi$ is a dominance association rule, where $\phi = (c_1,\gg) \wedge \ldots \wedge (c_s,\gg) \wedge (c_{s+1},\ll) \wedge \ldots \wedge (c_m,\ll)$ and $\psi = (c_{m+1},\gg) \wedge \ldots \wedge (c_t,\gg) \wedge (c_{t+1},\ll) \wedge \ldots \wedge (c_n,\ll)$. $\forall c_i \in \{c_1, \ldots, c_n\}$ is a criterion and $c_i \neq c_j$ for $i \neq j$. $(c_i,\gg)$ or $(c_i,\ll)$ is a dominance item, the compound of k dominance items is denoted *k-dominance itemset*.

The approach proposed mines dominance association rules as the following steps.

Step 1. Transform algorithm in Figure 1 transforms the preference-ordered data table $PDT = (U, C)$ to a tri-value data table $PDT^+ = (U^+, C)$ where $U^+ = (U \times U) - \{(x, x) \mid x \in U\}$. Each row, a record of PDT^+, corresponds to a pair of objects in PDT. A record label i is added to the ith row/record so that *AprioriTid* algorithm can be used to discovery frequent dominance

```
Algorithm Transform
    Input: PDT=(U,C)
    Output: PDT=(U,C)
Begin
For each record x in U do
    For each record y in U (y≠x)do
      For each criteria c in C do
          Insert record r in U
              If $c(x) >_r c(y)$ then $c(r)$ =2
              If $c(x) =_r c(y)$ then $c(r)$ =1
              If $c(x) =_r c(x)$ then $c(r)$ =0
      End for
    End for
End for
End
```

Figure 1. Algorithm *Transform*.

itemsets. Table 1 is a preference-ordered table. Table 2 is the tri-value data table corresponding to Table 1.

Step 2. Dominate_AprioriTid algorithm in Figure 2 discoveries all frequent dominance itemsets satisfying *minsup* defined by user. For a tri-value data table $PDT^+ = (U^+, C)$, $c_i \in C$ and a pair $(x, y) \in U^+$, (x, y) supports $(c_i, \gg)$ if and only if $c_i((x, y)) = 2$ or $c_i((x, y)) = 1$; (x, y) supports $(c_i, \ll)$ if and only if $c_i((x, y)) = 0$ or $c_i((x, y)) = 1$; (x, y) supports a dominance itemset ϕ if and only if (x, y) supports each dominance item in ϕ. The ratio between amount of pairs supporting ϕ in U^+ and $|U^+|$ is the support of ϕ, denoted *sup*(ϕ). If *sup*(ϕ) is not less than *minsup*, then ϕ is a frequent dominance itemset. After discovering the frequent dominance itemsets of length 1, we

Table 1. A preference-ordered data table.

Object	a_1	a_2	a_3
X_1	5000	Medium	Medium
X_2	7500	Low	High
X_3	10,000	High	Low
X_4	3000	Low	High

Table 2. A tri-value data table.

Object pairs	a_1	a_2	a_3	Object pairs	a_1	a_2	a_3
1 (X_1,X_2)	0	2	0	7 (X_3,X_1)	2	2	0
2 (X_1,X_3)	0	0	2	8 (X_3,X_2)	2	2	0
3 (X_1,X_4)	2	2	0	9 (X_3,X_4)	2	2	0
4 (X_2,X_1)	2	0	2	10 (X_4,X_1)	0	0	2
5 (X_2,X_3)	0	0	2	11 (X_4,X_2)	0	1	1
6 (X_2,X_4)	2	1	1	12 (X_4,X_3)	0	0	2

```
Algorithm Dominate_AprioriTid
    Input: $PDT^+$=($U^+$,$C$), minsup
    Output:  DARS(Dominance Association rules)
    Begin
    For each record $x$ in $U^+$ do
        For each criteria $c$ in $C$ do
            If $c(r)$ =2 or $c(r)$ =1
            Then add lablel of $r$ to (c, »).tidllist
            If $c(x)$ =1 or $c(r)$ =0
            Then add lablel of $r$ to ($c_2$,«).tidllist
        End for
    End  for,
    $L_1$={frequent 1-dominance iternset}
    Apply AprioriTid ($L_1$,minsup)
    End
```

Figure 2. Algorithm *Dominate_AprioriTid*.

can iteratively generate the frequent dominance itemsets of length $(k + 1)$ from the dominance itemsets of length k. *AprioriTid* algorithm can be applied here.

Step 3. Generating dominance association rules satisfying *minconf* from the frequent dominance itemsets. For a frequent dominance itemset ϖ, we can generate a dominance association rule $\phi \rightarrow \psi, \phi \cup \psi = \varpi$ and $\phi \cap \psi = \emptyset$. Its support is $sup(\phi \rightarrow \psi) = sup(\phi \wedge \psi)$ and its confidence is $conf(\phi \rightarrow \psi) = sup(\phi \wedge \psi)/sup(\phi)$.

The time complexity of mining dominance association rules is $O(N^2)$. Given $minsup = 5/12$ and $minconf = 80\%$, part of dominance association rules mined from Table 2 are listed in Table 3.

4. DOMINANCE ASSOCIATION RULES FOR PREDICTION

Definition 4. c_i and c_j are positive correlated criteria, if and only if the confidences of $(c_i, \gg) \rightarrow (c_j, \gg)$ or $(c_i, \ll) \rightarrow (c_j, \ll)$ is 1; c_i and c_j are negative

Table 3. Dominance association rules.

Dominance association rule	Confidence
$(a_1,\gg) \rightarrow (a_2,\gg)$	5/6
$(a_1,\gg) \rightarrow (a_3,\ll)$	5/6
$(a_2,\gg) \rightarrow (a_3,\ll)$	1
$(a_3,\ll) \rightarrow (a_2,\gg)$	1
$(a_1,\gg) \wedge (a_2, \gg) \rightarrow (a_3, \ll)$	1
$(a_1,\gg) \wedge (a_3, \ll) \rightarrow (a_2, \gg)$	1

correlated criteria, if and only if the confidences of $(c_i, \gg) \rightarrow (c_j, \ll)$ or $(c_i, \ll) \rightarrow (c_j, \gg)$ is 1.

Positive correlation of c_i and c_j means that for an object x, increasing the value of c_i in x will not decrease the value of c_j in x and decreasing the value of c_i in x will not increase the value on c_j in x. Negative correlation of c_i and c_j means that increasing the value of c_i in x will not increase the value of c_j in x and decreasing the value of c_i in x will not decrease the value of c_j in x.

For simple to describe how to predict the unknown value of one criterion in an object, a preference-ordered decision table is considered. A preference-ordered decision table is a decision table in which the condition attributes and the decision attributes are all criteria. Assume Table 1 is a preference-ordered decision table, in which a_1 and a_2 are condition criteria and a_3 is the decision criterion. And for an object y, the value of a_2 in y is *low*, but the value of a_3 in y is unknown. Because the value of a_2 in X_1 is *medium*, the value of a_3 in X_1 is *medium*, $a_3(y)$ should be greater than or equal to *medium* based on $(a_2, \gg) \rightarrow (a_3, \ll)$. At the same time because the value of a_2 in X_3 is *high*, the value of a_3 in X_3 is *low*, $a_3(y)$ should be greater than or equal to *low* based on $(a_2, \gg) \rightarrow (a_3, \ll)$. So $a_3(y)$ should be greater than or equal to *max*({*low*, *medium*}), i.e., greater than or equal to *medium*. The predict algorithm is omitted here.

5. CONCLUSIONS

A domain of an attribute is often preference-ordered in real world. By transforming the preference data table into a tri-value data table, Dominance association rules can be found and used for predicting unknown value of objects' (decision) criterion. Dominance association rules also can be used to measure which condition criteria is more significant to decision criteria in a preference-ordered decision table. So in future it can be a new measurement and can be combined with other measurements.

ACKNOWLEDGEMENTS

This paper is supported by NSFC of China (project numbers: 60373098, 60173006), '863' Project of China (project number: 2003AA118020) and Project of Jilin Province (20020303-2).

REFERENCES

1. R. Agrawal, T. Imielinski and A. Swami (1993), Association rule mining between sets of items in large databases. In: *Proceedings of 1993 ACM SIGMOD International Conference on Management of Data*, Washington, DC, pp. 207–216.

2. R. Srikant and R. Agrawal (1996), Mining quantitative association rules in large relational tables. In: *SIGNOD-96*.
3. R.J. Miller and Y. Yang (1997), Association rules over interval data. In: *Proceedings of the 1997 ACM SIGMOD International Conference on Management of Data*, Tucson, Arizona, United States, pp. 452–461.
4. S. Greco, B. Matarazzo, R.J. Slowinski and Stefanowski (2002), Association rule mining in preference-ordered data. In: *Proceedings of the 13th International Symposium on Methodologies for Intelligent System*, pp. 442–450.

MINING ORDINAL PATTERNS FOR DATA CLEANING

Y.B. Liu[1,2] and D.Y. Liu[1,2]
[1]College of Computer Science and Technology, Jilin University, Changchun 130012, China
[2]Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China

Abstract It is well recognized that sequential pattern mining plays an essential role in many scientific and business domains. In this paper, a new extension of sequential pattern, ordinal pattern, is proposed. An ordinal pattern is an ordinal sequence of attributes, whose values commonly occur in ascending order over data set. After each record in data set is transformed into an ordinal sequence of attributes according to their ordinal values, ordinal patterns can be mined by means of mining sequential patterns. One use of ordinal patterns is to identify possible error records in data cleaning, in which the values of attributes are inconsistent with the ordinal patterns which most of the data conform to. Experiments verify the high efficiency of the method presented.

Keywords: sequential pattern, ordinal sequence, ordinal pattern, data cleaning.

1. INTRODUCTION

It is well recognized that sequential pattern mining plays an essential role in many scientific and business domains [1, 2], such as customer shopping behaviour prediction, medical treatment, etc. In this paper, ordinal values (for example, numerical value) of attributes are focused on. A new extension of sequential pattern, ordinal pattern, is proposed and applied in detecting possible error records in data cleaning. After each record in data set is transformed into an ordinal sequence of attributes according to their ordinal values, ordinal patterns can be mined by means of mining sequential patterns, for example, *GSP* algorithm is used the to discovery ordinal patterns. An ordinal pattern is an ordinal sequence of attributes, whose ordinal values commonly occur in ascending order over data set. Ordinal patterns can identify possible error

G. R. Liu et al. (eds.), Computational Methods, 1267–1272.

records efficiently in data cleaning, which deals with detecting and removing errors and inconsistencies from data in order to improve the quality of data set. The error records detected by our method are inconsistent with the ordinal patterns which most of the data conform to. Until now few similar methods directly tackle this problem automatically [3, 4]. The approach of the paper gives a useful, practical and automatic approach, which can deal with multiple attributes.

2. MINING ORDINAL PATTERNS FOR DATA CLEANING

Assume data set R is composed of records and each record r is described by attributes $\{a_1, a_2, \ldots, a_m\}$ and their values, $a_i(r)$ $(1 \leq I \leq m)$. Usually a_i has an ordinal domain, in which any two values c_1 and c_2 meet $c_1 < c_2$, or $c_1 = c_2$, or $c_1 > c_2$. Sometimes values of different attributes can be compared. For example, the attributes of the glass, the weight percent of Na and Mg, can be compared to know which value is greater. In the following discussion, assume (1) $\forall a_i \in \{a_1, a_2, \ldots, a_m\}$ has an ordinal domain; (2) $\forall a_i, a_j \in \{a_1, a_2, \ldots, a_m\}(a_i \neq a_i)$, v_1 from the domain of a_i, and v_2 from the domain of a_j can be compared.

Definition 1. A record r corresponds to an *ordinal sequence* $os(r) = \{s_1, s_2, \ldots, s_l\}$ $(1 \leq l \leq m)$, meeting (1) s_i is a set of attributes; (2) $\forall a_s$ and $\forall a_t$ in s_i, $a_s(r) = a_t(r)$ is true; (3) $\forall a_s$ in s_i and $\forall a_t$ in s_j $(1 \leq i < j \leq l)$, $a_s(r) < a_t(r)$ is true. l is the length of $os(r)$. If there exist integers $i_1 \leq i_2$ such that $a_s \subseteq s_{i1}, a_t \subseteq s_{i2}$ in $\{s_1, s_2, \ldots, s_n\}$, then a_s *precedes* a_t, or else a_s *succeeds* a_t.

In Table 1, any two values of attributes of a, b, c and d can be compared, but those of a and f are not comparable. So what we focus on is only $\{a, b, c, d\}$ in R, not $\{a, b, c, d, e\}$. While sorting record r_2 based on ascending order of values of a, b, c and d, the result is $a(r_2) < d(r_2) < b(r_2) = c(r_2)$. Actually the ordinal sequence $\{\{a\}, \{d\}, \{b, c\}\}$ is enough to express the ordinal relationship

Table 1. A data set R.

	a	b	c	d	e
r_1	4	2	1	5	Green
r_2	3	3	1	2	Blue
r_3	2	4	2	1	Red
r_4	3	4	1	2	Green

```
Algorithm FOP
Input: data set R, minsup
Output ordinal patterns
For each record $r$ in data set
    Normalize or convert $r$
    Sort attributes based on their values.
    Generate ordinal sequence $os(r)$.
End for
Generate ordinal patterns by GSP.
```

Figure 1. Algorithm *FOP*.

between the values of a, b, c and d in r_2. By the same method, each record in R can be transformed into ordinal sequences.

Definition 2. An ordinal sequence $\{a_1,a_2,\ldots,a_n\}$ is a *subsequence* of another $\{b_1,b_2,\ldots,b_m\}$ if there exist integers $i_1 \leq i_2 \leq \cdots \leq i_n$ such that $a_1 \subseteq b_{i1}$, $a_2 \subseteq b_{i2},\ldots,a_n \subseteq b_{in}$.

Record r in R is said to support an ordinal sequence s if s is a subsequence of $os(r)$. The ratio between those records supporting s and $|R|$ is the support of s, denoted $sup(s)$. Given a minimum support threshold *minsup*, the problem of mining ordinal patterns is to discovery common ordinal sequences whose $sup(s)$s are not less than *minsup.* After each record in R is transformed into an ordinal sequence, ordinal patterns can be mined by means of mining sequential patterns. The algorithm *FOP* in Figure 1 discovering ordinal patterns first normalizes the data if necessary and then sorts every record. Only one scan of the data set is required. A temporary file with the results of sort is generated. *FOP* uses the algorithms *GSP* for mining sequential patterns in [1]. Assume m is the number of attributes, and then the complexity of sorting m attributes is m^2. Therefore the complexity of this algorithm is O(N*m^2), where N is the number of records in data. The results of this algorithm are written to another temporary files for use in data cleaning.

Now the problem is given a set of ordinal patterns, how can we judge whether r is error or not in data cleaning? Further, what should we take to measure the inconsistency between an ordinal pattern$p$ and an ordinal sequence $os(r)$?

Definition 3. Given an ordinal pattern p and an ordinal sequence $os(r)$, if a_i precedes a_j in $p(a_i \neq a_j)$, but a_i succeeds a_j in $os(r)$, then $\langle a_j, a_i\rangle$ is an *anti-pair* between p and $os(r)$. *Let AP* $(p,os(r))$ denote all anti-pairs between p and $os(r)$ and $|p|$ is the length of p, the *inconsistency* between p and $os(r)$ can be defined as:

$$inconsistency = |AP(p, os(r))|/C^2_{|\mathrm{P}|}$$

Table 2. Ordinal sequences.

	Sequences
$os(r_1)$	c, b, a, d
$os(r_2)$	$c, d, \{a, b\}$
$Os(r_3)$	$d, \{a, c\}, b$
$os(r_4)$	c, d, a, b

```
Algorithm Error
Input: an ordinal sequence $os(r)$,
       an ordinal pattern $p$.
Ouput: inconsistency
begin
for attribute $a$ and $b$ in $p$ do
  if $a$ proceeds $b$ in $p$, but $a$ succeeds $b$ in $os(r)$
   then anti-pairs=anti-pairs+1
   end if
end for
inconsistency = anti-pairs/ $C_{|p|}^2$
end
```

Figure 2. Algorithm *Error*.

```
Algorithm Error-Detect
Input: a data set $R$, an ordinal patterns set $P$,
       a threshold $\tau$
Ouput: a set of error records $E$
begin
for $os(r)$ of each record $r$ in $R$ do
sum=0 count=0
   for each ordinal patterns $p$ in $P$
      sum=sum+ Error($os(r)$,$p$), count=count+1
   end for
   if sum/count > $\tau$ then $E$=$E$+r  end if
   end for
end
```

Figure 3. Algorithm *Error-Detect.*

Given an ordinal pattern $p = \{(a), (b), (c), (d)\}$ and an ordinal sequence $os(r) = \{(d), (a), (b), (c)\}$, there are three anti-pairs ($\langle d,a\rangle$, $\langle b,a\rangle$, $\langle c,a\rangle$) between p and $os(r)$. Because $3/C_4^2 = 3/6 = 1/2$, $os(r)$ is 50% possibility inconsistent with p. So r may be an error record with 50% possibility based on p (Table 2).

Figure 2 shows the algorithm *Error* that computes the inconsistency between an ordinal pattern p and $os(r)$. Figure 3 shows the algorithm *Error-Detect* that computes the average inconsistency based on of a set of ordinal patterns and judging whether a record r is an error or not. If the average is not lower than a threshold τ chosen by user, it maybe an error record. The precedent work [4] focuses on using only two attributes. But our work takes into account ordinal patterns involving multiple attributes. The time complexity of this step is $O(N^{*} C_m^2)$.

3. EXPERIMENTS

Experiments are designed to prove the performance of *Error-Detect*. The Glass Identification Database including 214 instances from UCI data is used. The nine attributes used are RI (refractive index), the weight percent of Na, Mg, Al, Si, K, Ca, Ba and Fe in corresponding oxide. We generate x error records, which does consistent to the data in the database, and add them to the database. In fact, the records in the database are a little inconsistent. Even if we run *FOP* and *Error-Detect* only on itself, there are a few records identified as error records. Assume *Error-Detect* detects *error* records from x records and *source* records from the database. Define the accuracy of *Error-Detect* as: $Accuracy = error/(x + source) \times 100\%$. The parameters, *minsup*, τ and x, affect the result. The results show that threshold τ plays a more important role in identifying error records. And the results are independent of the data distribution. Let $\lambda = x/(x + 214) \times 100\%$, when $minsup \leq 1 - \lambda$, the accuracy gets high values, or else, the accuracy is low. In Tables 3 and 4, the bold parts

Table 3. Accuracy when $x = 40$.

$\lambda = 15.74\%$	*Accuracy*		
minsup	$\tau = 0.1$	$\tau = 0.15$	$\tau = 0.2$
70%	93%	100%	70%
75%	93%	100%	70%
80%	93%	100%	70%
85%	93%	70%	70%
90%	70%	70%	70%

Table 4. Accuracy when $x = 20$.

$\lambda = 8.55\%$	*Accuracy*		
minsup	$\tau = 0.1$	$\tau = 0.15$	$\tau = 0.2$
75%	87%	100%	70%
80%	87%	100%	70%
85%	87%	70%	100%
90%	100%	70%	70%
95%	70%	70%	70%

are those tests, in which *minsup* $\leq 1 - \lambda$. So if we can know λ according to the related information then the performance can get better. But when *minsup* is 1, no error records can be detected. Is it right? Yes. Because the ordinal patterns found with *minsup* $= 1$ are hold by each record, *Error-Detect* can detect none error record. The results of other experiments are omitted.

4. CONCLUSIONS

By extension to ordinal patterns, sequential pattern mining proves to be useful in identifying ordinal patterns that uncover error records in data set for data cleaning. An ordinal pattern reflects a kind of ordinal relationships between attributes that commonly occur over the data. One use of ordinal patterns is to identify possible error records, whose attributes do not conform to those uncovered orderings. The experiments show that our method can accurately identity most of the error records in data set.

ACKNOWLEDGEMENTS

This research has been supported by NNSF under grants no. 60173006.

REFERENCES

1. R. Srikant and R. Agrawal, Mining sequential patterns: Generalizations and performance improvements. In: *EDBT'96*.
2. R. Agrawal and R. Srikant, Mining sequential patterns. In: *ICDE'95*, 3–14, Taipei, Taiwan.
3. H. Galhardas, D. Florescu, D. Shasha and E. Simon (1999), An extensible framework for data cleaning, Institute National de Recherche en Informatique ét en Automatique.
4. A. Marcus, J.I. Maletic and Lin (2001), Ordinal association rules for error identification in data sets. In: *CIKM01*, pp. 589–591.

USER ASSOCIATION MINING BASED ON CONCEPT LATTICE

H. Qi, D.Y. Liu, L. Zhao and M. Lu
College of Computer Science and Technology, Jilin University, P. R. China

Abstract The problem of mining user association from rating table plays an essential role in rule-based recommender system. Using the closure of the Galois connection, we define two user association bases: the exact base (i.e., for all rules with a 100% confidence) and the approximate base (i.e., with confidence <100%) from which all valid user association rules with support and confidence can deduced. These user association bases are characterized using frequent closed itemsets and their reductions within concept lattice. Algorithm for extracting these two bases is presented and experimental evaluated on real-life databases. The results show that the proposed user association bases can considerably reduce number of rules in user association and do not loss any information.

Keywords: user association, concept lattice, user association base, recommender system.

1. INTRODUCTION

With the development of data mining, association rule-based recommend technique, which makes recommendation by using both item association and user association mined from the rating table, becomes a new interesting research topic. Fu et al. [1] developed a recommender system SurfLen, which uses Apriori algorithm to make web page recommendation for users online. Sarwar et al. [2] proposed a Top-N algorithm which first selected the users that are the most similar to the current user, and then mined association rules on these users. Lin et al. [3] proposed an adaptive-support algorithm ASARM which use the current user to limit the association rule's form and increase the efficiency of recommendation.

Recent research show that concept lattice is an efficient tool for association rules mining, since the set of all closed frequent itemsets can be orders of magnitude smaller than the set of all frequent itemsets and any information

G. R. Liu et al. (eds.), Computational Methods, 1273–1277.

do not be lost. Bastide et al. [4] define two bases for association rules which union is a generating set for all valid association rules with support and confidence. Zaki et al. [5] present a new framework for associations based on the concept of closed frequent itemsets. Stumme et al. [6] mine association rules by constructing Iceberg Concept Lattices.

Using the idea of concept lattice based association rules mining we define two user association bases from which all valid user association rules with support and confidence can deduced. These user association bases are characterized using frequent closed itemsets and their reductions within concept lattice.

2. SEMANTIC FOR ASSOCIATION RULE BASED ON CONCEPT LATTICE

In FCA [7] a *formal context* is a triple of $K = (O,I,R)$, where O is a set of object, I is a set of item and $R \subseteq \times I$ is a binary relation. The inclusion $(o, i) \in R$ is read as 'object o and item i has relation'. For $X \subseteq O$, we define $\varphi\ (X) = \{i \in I | \forall o \in X, (o,i) \in R\}$; and for $Y \subseteq I$, we define dually $\psi\ (Y) = \{o \in \mathrm{O} \mid \forall I \in Y, (o,i) \in R\}$. Then $\gamma = \varphi o \psi$ is a *closure operator* on I. A *formal concept* is a pair (X,Y), with $X \subseteq O$, $Y \subseteq I$, $\varphi\ (X) = Y$, $\psi\ (Y) = X$. X is called *extent* and Y is called *intent* of the concept. The set of all concepts of a formal context K together with the partial order $(X_2, Y_2) \leq (X_1,Y_1) \Leftrightarrow Y_1 \subseteq Y_2$ (which is equivalent to $X_1 \supseteq X_2$) is a complete lattice, called *concept lattice* of K.

A set of items $l \subseteq I$ is called an *itemset*. The *support* of an itemset l is the percentage of objects in O containing l: *sup* $(l) = |\psi(l)|/|O|$. l is a *frequent itemset* if *sup*$(l) \geq$ *minsupport*. A frequent itemset $l \subseteq I$ is a *frequent closed itemset* iff $\gamma(l) = l$. An itemset $g \subseteq I$ is a *reduction* of a closed itemset l iff $\gamma(g) = l$ and $\neg\exists\ \mathrm{g}' \subset \mathrm{g}$ such that $\gamma(g') = l$.

An *association rule r* is an implication between two frequent itemsets l_1, $l_2 \subseteq I$ of the form $l_1 \rightarrow (l_2 \backslash\ l_1)$ where $l_1 \subset \mathrm{l}_2$. The *support* and *confidence* of r are defined as: *sup*$(r) =$ *sup* (l_2), *conf*$(r) =$ *sup* (l_2) / *sup* (l_1).

3. USER ASSOCIATION MINING BASED ON CONCEPT LATTICE

Recommender system first collects all users' rating information about kinds of recourse to form the rating table. The rating table is usually expressed by a table, in which the value in the ith row and the jth column is the rating that the

jth user remarks on the ith resource. If the value is 1 it means that the jth user likes the ith resource, otherwise it means no rating about the ith resource.

In recommender system based on user association, given the current user U_c, all association rules of the form $\wedge U_i \rightarrow U_c$ (*sup*,*conf*) (called as a *user association* about U_c, where U_i are users in the system, $U_i \neq U_c$, and *sup* and *conf* are support and confidence of the rule) first be found, then the common interest of the users in the condition of the user association is suggested to the current user U_c.

The exact user associations are rules of the form r: $\wedge U_i \Rightarrow U_c$ and $conf(r) = 100\%$. Therefore $sup\ (\cup U_i \cup U_c) = sup\ (\cup U_i)$, that is U_i and U_c appear in the intent of the same concept. Hence, we only need to consider the concept containing the current user to mine exact user association. Here we use G_+ to denote the set of nodes containing the current user.

Definition 1. *Exact User Association Base* is: $EB = \{r : g \Rightarrow U_c | U_c \in f \wedge g \in G_f \wedge\ U_c \notin g\}$, where f is a frequent closed itemset containing U_c, G_f is the reduction set of f.

Proposition 1. All valid exact user associations, their supports and their confidences (that are equals to 100%) can be deduced from the rules of the exact user association base and their supports.

The approximate user associations are rules of the form r: $\wedge U_i \rightarrow U_c$ and $conf(r) < 100\%$. Therefore the node containing U_i is the ancestor of the node containing U_c. Since the ancestors containing U_c have been considered in exact base we only need to consider relationship between nodes that are in G_+ and have ancestors not in G_+ (denoted by G_b).

Definition 2. *Approximate User Association Base* is: $AB = \{r : g \rightarrow U_c | U_c \in f \wedge g \in G \wedge\ \gamma(g) \subset f\}$, where f is a frequent closed itemset containing U_c, G is the reduction set of all frequent closed itemset.

Proposition 2. All valid approximate user associations, their supports and their confidences can be deduced from the rules of the approximate user association base and their supports.

Algorithm: UAM

Algorithm: UAM (User Association Mining)

Input: Concept lattice G constructed from rating table D, current user U_c, support threshold *min_sup* and confidence threshold *min_conf*.

Output: Exact user association base *EB* and approximate user association base *AB*.

```
1. construct $G_+(U_c, min\_sup)$ and $G_b(U_c, min\_sup)$
2. For $H \in G_+$ do
```

```
    $G_f \leftarrow$ int_red $(H.f)$
    For $g \in G_f$ and $U_c \notin g$ do
        $EB \leftarrow EB \cup \{\grave{r} : g \Rightarrow U_c(H.c/|O|, 1)\}$
    End for
    If $H \in G_b$ then
        push ($stack$, $H$)
        While not empty ($stack$) do
            $N \leftarrow$ pop ($stack$)
            For each parent $P$ of $N$ do
                If $P \notin G_+$ and not mark($P$) and $P.f \neq \emptyset$ and $H.c/P.c \geq min\_conf$ then
                    $G_f \leftarrow$ int_red $(P.f)$
                    For $g \in G_f$ do
                        $AB \leftarrow AB \cup \{r : g \rightarrow U_c\ (H.c/|O|, H.c/P.c)\}$
                    End for
                    push ($stack$, $P$)
                    mark ($P$)
                End if
            End for
        End while
    End if
End for
```

4. EXPERIMENTAL RESULT AND CONCLUSION

We use the MovieLens dataset of GroupLens Research Project as test-bed of our approaches, which contains ratings from 943 users for 1682 movies. The number of rules for different support threshold and confidence threshold are presented in Table 1 in which '*Ers*' is exact rules and '*Ars*' is approximate rules.

Table 1. Number of exact and approximate user associations extracted.

					$min\_sup = 10\%$					
	$min\_sup =$ 10%		$min\_sup =$ 20%		$min\_conf =$ 30%		$min\_conf =$ 70%		$min\_conf =$ 90%	
U_c	*EB*	*Ers*	*EB*	*Ers*	*AB*	*Ars*	*AB*	*Ars*	*AB*	*Ars*
6	138	262	0	0	810	14,593	657	11,349	210	1758
49	366	4456	10	54	834	6217	798	6203	462	4016
55	439	3199	5	14	692	8517	525	7869	220	3592
97	344	3696	1	6	857	7337	782	7126	343	3670
182	58	1139	0	0	743	12,842	454	10,298	153	2044

The experiments on real-life databases show that the proposed user association bases can considerably reduce number of rules in user association and are easily usable from the point of view of the users.

ACKNOWLEDGEMENTS

This research benefits from the support of NNSF (National Natural Sciences Foundation of China under grants No 60173006) and Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University.

REFERENCES

1. X. Fu, J. Budzik and K.J. Hammond (2000), Mining navigation history for recommendation. In: 2000 *International Conference on Intelligent User Interfaces*, New Orleans, pp. 106–112.
2. B. Sarwar, G. Karypis, J. Konstan and J. Riedl (2001), Item-based collaborative filtering recommendation algorithms. In: *10th International World Wide Web Conference*. Hong Kong, China, pp. 285–295.
3. W. Lin, S.A. Alvarez and C. Ruiz (2001), Efficient adaptive-support association rule mining for Recommender Systems. *Data Mining and Knowledge Discovery*, 6, 1, pp. 83–105.
4. Y. Bastide, N. Pasquier, R. Taouil, G. Stumme and L. Lakhal (2000), Mining minimal non-redundant association rules using frequent closed itemsets. In: *1st International Conference on Computational Logic*, London, pp. 972–986.
5. M.J. Zaki (2000), Generating non-redundant association rules. In: *6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Boston, pp. 34–43.
6. G. Stumme, R. Taouil, Y. Bastide, N. Pasquier and L. Lakhal (2001), Intelligent structuring and reducing of association rules with formal concept analysis. *The Joint German/Austrian Conference on AI*. Vienna, Austria, pp. 335–350.
7. B. Ganter and R. Wille (1999), *Formal Concept Analysis: Mathematical Foundations*. Springer-Verlag, Heidelberg, Germany.

A PLM-ORIENTED WORKFLOW MODEL

Wan-jun Yu[1,2], Da-you Liu[2] and Quan Liu[2]
[1]*Northeast China Institute of Electric Power Engineering, Jilin 132012*
[2]*College of Computer Science & Technology, Jilin University, Changchun, 130021 China*

Abstract Workflow management systems (WFMSs) are used to coordinate and streamline business processes. The feedback iteration occurred frequently in very large process instances, thus bring on the high cost. In this paper, the issue that process modelling based on product lifecycle management (PLM) is studied and a parallel model – PLM-Based Parallel Workflow Model is proposed. By using the model, the complexity of process definition is reduced. In addition, the adhoc access control policies can be created within each phase of lifecycle, so the access control requirement of least privilege can be satisfied. Finally, we give two algorithms used to parse role.

Keywords: product lifecycle management, workflow, parallel model.

1. INTRODUCTION

Since the concepts of product lifecycle management (PLM) were proposed firstly by Eean [1] and Levirt [2], the intension and extension of PLM are changing and developing at all times. Up to now, it really contains completely the whole process of product. The product data management (PDM) is a typical application of PLM. As a driver, workflow is one of the most important functional components of PDM system. Based on the WfMC reference model [3] and the requirement of workflow management in PDM [4], we proposed a lifecycle oriented parallel workflow model (LOPWM). LOPWM can decompose the process definition with complex structure into relative simple subprocess of lifecycle phase that can execute parallelly, and have the capability of creating adhoc access control and parsing roles to tasks participants. Figure 1 shows the comparison between Sequential and Parallel.

G. R. Liu et al. (eds.), Computational Methods, 1279–1284.

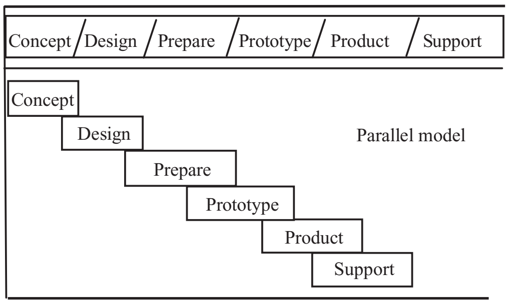


Figure 1. The comparison of sequential and parallel.

2. TEAM-ROLE MANAGEMENT

Definition 1. An object is called team managed if it can be associated with a team-role object. Only lifecycle object is team managed in the paper.

RBAC96 is both important and popular access control model [5], but RBAC96 do not define the user group and authorization about data object. We extend the RBAC_0 of RBAC96 to eRBAC and we use the eRBAC as the team-role management model.

Definition 2. Team-role $= [U, R, G, P, O, U, G, U - G, RoP]$, where

(1) User set U; Role set R; Group set G; Permission set P; Object set O;

(2) $U : U \to R$, the mapping for user to role. $r = U(u)$ is a role that u plays. The number of roles that one user can play is limited to one, but one role can be played by different users, hence U is of many to one relationship.

(3) $G : G \to R$, the mapping for group to role. $r = G(g)$ is a role that g plays, i.e., every member in g plays role r. similar to U, G is also a many to one relationship.

(4) $U - G \subseteq U \times G$, the membership of users and group. A group can have many members, and one user can be the member of more than one groups. Hence it is a many to many relationship.

(5) $RoP \subseteq R \times O \times P$. $(r, \tau_o, \mathrm{p}) \in RoP$, means that role r is granted access on object o of type τ_o with privilege p. The connection between permission and user is made by role. According to Definition 2, u has the privileges that the privileges of his roles plus the privileges of his groups. The privilege set of u is $\{\mathrm{p} | r = U(u) \wedge (r, \tau_o, \mathrm{p}) \in RoP\} \cup \{\mathrm{p} | (u, g) \in U - G \wedge r = G(g) \wedge (r, \tau_o, \mathrm{p}) \in RoP\}$.

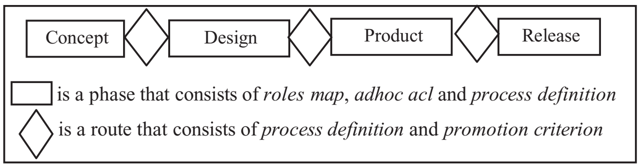


Figure 2. An example of lifecycle model.

3. PRODUCT LIFECYCLE MANAGEMENT

3.1 Lifecycle management

Definition 3. An object is called lifecycle managed if the object has different lifecycle states and different access control policies in different state, i.e., it can associate with a lifecycle object.

The lifecycle managed objects in this paper is document, part and engineering change. Lifecycle object manages the creating, modifying and auditing and release of part, administrates the reviewing of document, and controls the change requisition, change order and change activity.

3.2 The element of lifecycle model and relationship

A lifecycle is a series of phase states that lifecycle managed object has to pass through. One phase consists of a state and a route, shown in Figure 2.

Definition 4. $\text{lifecycle} = [\text{phase}_1, \text{phase}_2, \ldots, \text{phase}_n]$, the phase_i is defined as:

$$\text{phase}_i = < \text{state, route} >, \text{if} i \neq n.$$
$$\text{phase}_n = \text{state}_n, \text{if} i = n.$$

State: is a lifecycle phase of lifecycle managed object. The state should associate with a workflow process if object in the state will be processed to achieve some goals. Otherwise there is not.

Route: is an estimate step inserted between phases that checkouts the object and make a decision whether the object should be promoted to the next phase. A state and a route can associate with a workflow process respectively.

3.3 The special requirement about lifecyclemanagement

Besides the workflow associated with the state or route, the role mapping, the adhoc access control policies and the promotion criterion also should be defined.

Algorithm 1 ***The 2-level roles parse algorithm***

```
INPUT:  1) A lifecycle = [$phase_1$, $phase_2$, ... , $phase_n$]

            $i \neq n$, $phase_i = <state_i, route_i>$
            $i = n$, $phase_n = state_n$
        2) A team-Role=$<U, R, G, P, O, \mathcal{U}, \mathcal{G}, U\text{-}G, RoP>$,where $|R|=m$.
OUTPUT: roleMap[][]
    1. for $i$ from 1 to $n$
           for $j$ from 1 to $m$
               set roleMap[$i$][$j$] as null
           endfor
       endfor
    2. for each $state_i$
        for j from 1 to m
           if ($\Sigma^i(r_j)=r' \in R$ ) , then roleMap[$i$][$j$]= $\{u|u\in U \wedge \mathcal{U}(u)=r'\}\cup$
                      $\{g|g\in G \wedge \mathcal{G}(g)=r'\}$;
              else if($\Sigma^i(r_j) \in U \cup G$), then roleMap[$i$][$j$]=$\Sigma^i(r_j)$;
                 else roleMap[$i$][$j$]= $\{u|u\in U \wedge \mathcal{U}(u)=r_j\}\cup\{g|g\in G \wedge \mathcal{G}(g)=r_j\}$
         endfor
       endfor
```

Figure 3. 2-Level role parse algorithm.

Role mapping: Σ^i: $R \to 2^S$. If $\Sigma^i(r) \subseteq U \cup G$, then we get the participants of workflow activity directly. If $\Sigma^I(r) \subseteq R$, then we parse the participants of workflow activities by team-role object. According to definition 2, we know $S = U \cup G \cup R$. The 2-level role parse algorithm is shown in Figure 3.

Adhoc access control: for $r \in R$, suppose o′ is the object processed by workflow in $state_i$ and the adhoc privileges of r is $P_i(r)$, then $P_i(r) \cap (\{\mathrm{p}\ |r = U(u) \wedge (\tau_o, r, \mathrm{p}) \in RoP\} \cup \{\mathrm{p}|(u,g) \in U - G \wedge r = G(g) \wedge (\tau_o, r, \mathrm{p}) \in RoP\})$ is the permissions for r access on o′.

4. PROCESS MANAGEMENT

Process management includes workflow modelling, workflow engine and worklist-handler. We suppose that object managed by workflow must be managed by lifecycle, but it is not true conversely.

Definition 5. An object is called workflow managed if it can be processed by a workflow instance.

4.1 Process modelling

The model element of process definition is defined as follows:
processDefinition =<procType, procDepAttr, procNorAttr> where:

```
Algorithm 2  The 3-level roles parse algorithm
 INPUT:   1) assignedActivity$_k$ .Participants
          2) state$_i$
          3) teamRole = $[U, R, G, P, O, \mathcal{U}, \mathcal{G}, U\text{-}G, RoP]$
OUTPUT:  roleMap[]
    1. for $j$ from 1 to $m$
         set  roleMap[$j$] as null
       endfor
    2. for $j$ from 1 to $m$
         if $\Pi^k(r_j) \in U \cup G$ then roleMap[$j$]= $\Pi^k(r_j)$
           else if $\Pi^k(r_j)=r'' \in R \wedge \Sigma^i(r'')=r' \in R$ then roleMap[$j$]=
             $\{u|u \in U \wedge \mathcal{U}(u)=r'\} \cup \{g|g \in G \wedge \mathcal{G}(g)=r'\}$
               else if $\Sigma^i(r'') \in U \cup G$ then roleMap[$j$]=$\Sigma^i(r'')$
                   else roleMap[$j$]=$\{u|u \in U \wedge \mathcal{U}(u)=r_j\} \cup \{g|g \in G \wedge \mathcal{G}(g)=r_j\}$
       endfor
```

Figure 4. 3-Level role parse algorithm.

1. procType: is the goal of process, e.g., business planning, change management, customer services, human resource and quality control, etc.
2. procDepAttr: is the dependent attributes, e.g., goal constraints, input data, output data, pre- and post- activity conditions, start conditions, execution conditions, transition conditions, deadline and except, etc.
3. procNorAttr: is normal attributes. e.g., the process ID and the goal description.
4. *ActNorAttr*: is the normal activity's attributes that can be used to identify the activity, e.g., the name of activity and the task description.

Lets Π^k: $R \to S$ is the mapping from R to S, $S = U \cup G \cup R$. (there is a hypothesis that limit one activity can associate with one team-role only). We give the 3-level role parse algorithm for the former shown in Figure 4.

4.2 Workflow engine

The workflow engine provides environments of workflow enactment, interprets the process definition, controls the creating, activating, suspending and terminating of process instance, take charge of the states transition of process and activities, navigates the activities running, maintains the relevant data and control data, and invoke the external applications.

4.3 Worklist-handler

The task in worklist corresponds to the AssignedActivity. The user's worklist is managed by worklist-handler which transmits messages of state transition between users and workflow systems and monitor interacting between users and workflow engine. The manner that task is completed by participants is: selecting a workitem, reassigning a workitem, notifying the completed of task, updating the deadline of selected workitem and invoking external facilities or applications.

5. RELATED WORK AND CONCLUSION

Up to now, researches about the PLM more focus on the concepts and actions [6]. It is importance for enterprise informatization to research the model of PLM deeply and integrate PLM with workflow. It is also one of the trends for enterprise informatization. At present, there are several commercial software systems in the marketplace [7].

LOPWM is a flexible model to define the task participants for the complex business process, we can softly define process by assigning participants to roles not to user or groups, then use the roles parsing algorithm to parse the final participants to complete the tasks.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China under Grant No. 60173006 and the National High-Tech Research and Development Plan of China under Grant No.2003AA118020.

REFERENCES

1. J. Dean (1950), Pricing policies for new products. *Harvard Business Review*, 28, 6, pp. 45–53.
2. T. Levirt (1965), Exploit the product life cycle. *Harvard Business Review*, 43, 6. pp. 81–94.
3. WfMC (January 1995), TC00-1003. *Workflow Management Reference Model*. Workflow Management Coalition, Hampshire, UK.
4. J. Schöttner and Qi Guoning (2000), *Produktdatermanagement in der Fertigurgsindustrie: Prinzip-Korzepte-Strategien*. China Machine Industry Press.
5. R. Sandhu, E. Coyne, H. Feinstein et al. (1996), Role-based access control models. *IEEE Computers*, 29, 2, pp. 38–47.
6. Huang Shuang-xi and Fan Yu-shun (2004), Overview of product management. *Computer Integrated Manufacturing System-CIMS*, 1, 10, pp. 1–9.
7. M. Burkett, J. Kemmeter, K. O'Marah (2002), Product lifecycle management: what's real now. *ARM Research*.

A LANGUAGE FOR SPECIFYING CONSTRAINTS IN WFMSs*

Wan-Jun Yu[1], Da-You Liu[2] and Li Jiang[2]
[1]*Northeast China Institute of Electric Power Engineering*
[2]*College of Computer Science and Technology, Jilin University, Changchun 130012, China*

Abstract Although there are some approaches to formally describe the authorization constraints, such as role-based access control models, they are not adequate to model such WFMSs' constraints. To address this issue, this paper analyzes the existing methods and proposes a workflow authorization constraints model which including an authorization specification language named RWAL (role-based workflow authorization language) used to specify both static and dynamic authorization constraints. Its basic elements, syntax and semantics are discussed. It is the expanded version of Bertino's and more intuitionistic. The expression power of the language is better than the Bertino's in that it can not only represent both static authorization constraints and dynamic authorization constraints but also can capture the historical information of authorizations.

Keywords: workflow, workflow management systems (WFMSs), authorization, constraints, security policy.

1. INTRODUCTION

The security management of large system is very hard. To simplify the complexity of security administration, it is common practice to allocate a role to each activity in the process and then assign one or more users to each role, i.e., the authorization is for the roles not for the users [1]. Security policies are often expressed as constraints (or rules) upon users and roles, e.g., *the separation of duties* [2]. Unfortunately, the existing role-based authorization models are not adequate to model WFMSs' constraints.

---

*supported by the National Natural Science Foundation of China under Grant No. 60173006 and the National High-Tech Research and Development Plan of China under Grant No. 2003AA118020.

G. R. Liu et al. (eds.), Computational Methods, 1285–1290.

In the fields of role-based authorization, Sandhu et al. proposed several types of authorization constraints [3]. Jonscher et al. modeled the constraints resulting from the separation of duties using the activation conflict relation and the association conflict relation [4]. However, these attempts are not sufficient to model all the constraints required in WFMSs because they failed to capture the history of events. The significant work in this orientation owes to Nyanchama and Osborn [5] and Ahn and Sandhu [6]. But this research did not attempt to specify access control in terms of activities or tasks, so it is inadequate for WFMSs.

As regards workflow security authorization, Atluri et al. have proposed an authorization model suitable for workflows, called *Workflow Authorization Model* (WAM) [7]. WAM ensures that the tasks are executed only by authorized users or processes (subjects). One significant work in this direction is due to Bertino and Ferrari [8]. But this language is insufficient for supporting the role hierarchy and the number of rules is very large resulting in the higher cost for checking the constraints consistency.

2. THE RCL2000 AND THE AUTHORIZATION CONSTRAINTS OF WFMSs

The RCL2000, proposed by Ahn and Sandhu, used to represent the constraints formally, consists of two parts. One is the basic elements and system functions on which RCL2000 is based, the other is the additional elements and system functions. Another formal language proposed by Bertino uses several rules to represent an authorization constraint, so it can express formally authorization constraints as clauses of a logic program. Restricted by the size of this paper futher details are omitted.

3. THE BASIC REQUIREMENTS OF RWAL

3.1 The Authorization Model of Workflow System

The constraints on role and user assignments to tasks in a workflow are categorized into three main categories by Bertino according to the time at which they can be evaluated. (1) Static constraints (SC) can be evaluated without executing the workflow. (2) Dynamic constraints (DC) can be evaluated only during the execution of the workflow. (3) Hybrid constraints (HC). Constraints whose satisfiability can be partially verified without executing the workflow.

Definition 1 (The WFMSs Authorization Model). Symbol U, R, T, C and *UA* denotes the users set, roles set, workflow tasks set, authorization constraints set and user/role relationships respectively. RH denotes roles hierarchies, $\leq_r$ for briefly.
role: T → R, for $\forall$ $t_i \in$ T, role $(t_i) \in$ R represents the role executing the t_i.
user: T → U, for $\forall$ $t_i \in$ T, user $(t_i) \in$ U represents the user executing the task t_i, and (user(t_i), role(t_i)) $\in UA$.
roles: U → 2^R, for $\forall$ $u_i \in$ U, roles $(u_i) = \{r \in R \mid (\exists r' \geq_r r) \wedge (u_i, r') \in UA\}$ is the set of roles that the user u_i belongs to.
users: R → 2^U, for $\forall$ $r_i \in$ R, user $(r_i) = \{u \mid (\exists r' \geq_r r_i) \wedge (u, r') \in UA\}$ represents the set of users that belong to the role r_i.
tasks: U → 2^T, for $\forall$ $u_i \in$ U, tasks (u_i) represents all tasks executed by the user u_i.

3.2 K-level Role Function

RH is a partial order in RBAC models. The lowest roles in RH correspond to the bottom staves in an organization. The higher the role is in RH, the more senior he is in the organization, so we define the roles level in RH as follow: the level of lowest role r in RH is defined 0, denoted as *level* $(r) = 0$, and *level* $(r') = 1$ where r' is r's direct superior role, and the like. The highest level of RH is defined as *level (RH)*.

Definition 2 (*k*-level role function). For $\forall r_i \in$ R, we define the *k-level* role function of r_i as RL^k: R → 2^R:

- When k > 0, if $\exists r \in R$, such that *level* $(r_i) + k \leq$ *level* (RH), then $RL^k(r_i) = \{r \mid (r >_r r_i) \wedge (level(r) - level(r_i) = k)\}$, otherwise $RL^k(r_i) = \phi$ or say that RL^k (r_i) is not exist.
- When k < 0, if $\exists$ r $\in$ R, such that *level* $(r_i) + k \geq 0$, then $RL^k(r_i) = \{r \mid (r_i >_r r) \wedge (level(r) + k = level(r_i))\}$, otherwise $RL^k(r_i) = \phi$ or say that RL^k (r_i) is not exist.
- When $k = 0$, R^k $(r_i) = \{r_i\}$.

In the following, we always assume R^k $(r_i) \neq \phi$, i.e., $R^k(r_i)$ is existing.

3.3 Workflow Specification and Conflicting Tasks

Definition 3 (The Specification of Workflow). A workflow W is defined formally as $W = [TRS_1, TRS_2, \ldots, TRS_n]$. Where $TRS_i = (t_i, PR_i$, role (t_i), user $(t_i))$, $t_i \in$ T is a workflow task, $RS_i \subseteq$ R is candidate roles set authorized to execute the task t_i. Role $(t_i) \in RS_i$ is the actual role executing the task t_i. User

(t_i) is the accrual user executing the task t_i.

From the definition 1 we have role (t_i) $\in$ roles (user (t_i)) and user (t_i) $\in$ users (role (t_i)).

Definition 4 (The Tasks Partial relation *TH*). We call $TH \subseteq T \times T$ as a partial order relation on tasks set T, denoted as $\leq_t$.

The meaning of conflicting role (CR) and conflicting user (CU) is as the same as in RCL2000. Conflicting task (CT) represents that two or more tasks would not be executed by the same role or user.

Definition 5 (The Conflicting Tasks in Workflow System). The conflicting tasks of a workflow *W* is the set CT (*W*) $= \{CT_1, CT_2, CT_3, \ldots, CT_m\}$, where $CT_i \subseteq T$, for $\forall\, t_i, t_j \in CT_i$, we have role ($t_i$) $\neq$ role (t_j) and user (t_i) $\neq$ user (t_j).

4. CONSTRAINTS SPECIFICATION LANGUAGE-RWAL

4.1 The Essential Elements of RWAL

Definition 6 (Essential Elements of RWAL). The essential elements of RWAL consist of constant, variable symbols, function symbols and predicate symbols.

(1) Constant: U: users set, R: roles set, T: tasks set, C: constraints set and N: natural number set.
(2) Variable symbols: Arbitrary element of U, R, T, C and N denoted as V_U, V_R, V_T, V_C and V_N.
(3) Function symbols: The set of function symbols is {role, user, roles, users, tasks, OE(X), AO(X), RL^k}. The symbols listed above are called *term*.
(4) Predicate symbols: The set of predicate symbols consists of four sets:

- A set of binary comparison predicates. Its elements are:
  (1) Equation (x_1, x_2): if Equation (x_1, x_2) is true, then x_1 equal to x_2 (x_1, $x_2 \in$ N);
  (2) Greater (x_1, x_2);
  (3) Less (x_1, x_2). The semantics of last two is similar to the first.
- A set of membership predicates. Its elements are: Belong_To (x, X): if Belong_To (x, X) is true, then the individual x is an element of X.
- A set of capturing the effect of a workflow execution. Its elements are:
  (1) User_Success_Done (u_i, t_j): User_Success_Done (u_i, t_j) is true if user u_i executed task t_j successfully. Where $u_i \in$ U, $t_j \in$ T;

(2) Role_Success_Done (r_i, t_j);
(3) User_Execut_Abort (u_i, t_j);
(4) Role_Execut_Abort (r_i, t_j). The semantics of last three is similar to first one.

- A set of expressing authorization predicates.

(1) User_Must_Execute (u_i, t_j): User_Must_Execute (u_i, t_j) is true if user u_i is authorized to execute task t_j. Where $u_i \in U$, $t_j \in T$.
(2) Role_Must_Execute (r_i, t_j);
(3) User_Cannot_Do (u_i, t_j):
(4) Role_Cannot_Do (r_i, t_j): Role_Cannot_Do (r_i, t_j). The semantics of last three is similar to the first one.

4.2 The Rules of RWAL

Definition 7 (rules of RWAL). rules of RWAL is an expression of the form: $H \leftarrow L_1, \dots, L_m$ ($m \geq 0$), where H called head of the rule is an atom, $L_1, \dots, L_m$ called the rule *body* are literal. The rule $H \leftarrow$ represents the fact.

4.2.1 The Consistency of The Rules Set

Definition 8 (Consistency of Rules Set). The Rule-Set (W) is called consistence if Rule-Set (*W*) satisfies the following two conditions at the same time.
(1) for $\forall$rule $(i, j) \in$ Rule-Set (W), if $t_p \in$ rule(i, j).head, $t_q \in$ rule(i, j). body, then $t_p \leq_t t_q$.
(2) for $\forall$rule $(i, j) \in$ Rule-Set (W) and $\forall$rule$(m, n) \in$ Rule-Set (W), if $t_p \in$ rule (i, j). head, then $t_p \notin$ rule (m, n).body and if $t_p \in$ rule (m, n).head, then $t_p \notin$ rule (i, j).body.

4.2.2 The Authorized Roles and Users and Prevented Roles and Users for a Task t_i

The Authorized roles and users and the prevented roles and users for every task t_i can be evaluated after the C-Rule-Set (W) is obtained. For $\forall t_i \in T$, the authorized roles set permitted-users (t_i), authorized users set permitted-users (t_i), prevented roles set denied-roles (t_i) and prevented users set denied-users (t_i) of t_i are defined respectively as follow:

permitted-roles (t_i) = $\cup$ {r | Role_Must_Execute (r, t_i) $\in$ C-Rule-Set (W)}
permitted-users (t_i) = $\cup$ {u | User_Must_Execute (u, t_i) $\in$ C-Rule-Set (W)}
denied-roles (t_i) = $\cup$ {r | Role_Cannot_Do (r, t_i) $\in$ C-Rule-Set (W)}
denied-users (t_i) = $\cup$ {u | User_Cannot_Do (u, t_i) $\in$ C-Rule-Set (W)}

5. CONCLUSIONS

This paper proposed an authorization specification language which intended to specify both static and dynamic authorization constraints. The definition of consistency of rules set and checking algorithm are also given. It is the expanded version of Bertino's and is more intuitive. It can not only represent both static authorization constraints and dynamic authorization constraints but also can capture the history information of authorization.

REFERENCES

1. E. Bertino, F. Ferrari and V. Auluri (1997), A flexible model supporting the specification and enforcement of role-based authorizations in workflow management systems. In: *Proceedings of the Second ACM Workshop on Role-Based Access Control*, Fairfax, VA, pp. 1–12.
2. R. Sandhu (1991), Separation of duties in computerized information systems. *Database Security IV: Status and Prospects*, North Holland, pp. 179–189.
3. R. Sandhu, E.J. Coyne, H.L. Feinstein and C.E. Youman (1996), Role-based access control models. *IEEE Computer* 29, 2, pp. 38–47.
4. D. Jonscher, J. Moffet and K. Dittrich, Complex subjects or the striving for complexity is ruling our world. In *Database Security VII: Status and Prospects*. Elsevier North-Holland, Inc., Amsterdam, The Netherlands, pp. 19–37.
5. M. Nyanchama and S. Osborn, (1993), Role-based security, object oriented databases and separation of duty. *SIGMOD Rec.* 22, 4, pp. 45–51.
6. G.-J. Ahn and R. Sandhu (2000), Role-based authorization constraints specification. *ACM Transactions on Information and System Security*, 3, 4, pp. 207–226.
7. V. Atluri and WK Huang (1996), An authorization model for workflows. In: *Proceedings of the 5th European Symposium on Research in Computer Security, in Lecture Notes in Computer Science No. 1146. Springer-Verlag*, pp. 44–64.
8. Elisa Bertino and Elena Ferrari. (1999), The specification and Enforcement of Authorization constraints in workflow management systems. *ACM Transaction on Information and System Security*, 2, 1, pp. 65–104.

ONTOLOGY BASED WORKFLOW MODEL

Zhao-hui Zhang, Da-you Liu, Wei-jiang Liu and Bo Yang
College of Computer Science and Technology, Key Laboratory of Symbolic Computing and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, China, 130012

Abstract This paper is focused on semantic spatial information in workflow. The spatial relations is described by Description Logic, the ontologies is then constructed using the RDF Schemas and OWL Capabilities. These theories and methods are enclosed in a middleware, and the middleware was embedded in workflow system of the government in Jilin province.

Keywords: workflow, semantic web, ontology, description logic, middleware

1. INTRODUCTION

Workflow technology continues to be subjected to on-going development in its traditional application areas of business process modeling and business process coordination, and now in emergent areas of component frameworks and inter-workflow, business-to-business interaction. Addressing this broad and rather ambitious reach, a large number of workflow products, mainly workflow management systems, are commercially available [0]. The research of spatial information process quickly developments in the past 10 years. Since many information are interrelated with its location (or spatial properties), spatial information manage are also necessary for workflow systems, especially in some special domains such as cadastre management. To embed spatial information management in applications, nowadays there are many enterprise-scale spatio-temporal (or simply spatial) middleware products. They have different roles in future componentized OpenGIS architectures. In workflow systems or other applications, users need more semantic information. But the OpenGIS technology based productions can only handle the syntax spatial information, not the

G. R. Liu et al. (eds.), Computational Methods, 1291–1295.

semantic one. As we all known, Ontologies and semantic web are the proper technologies for managing semantic information. This paper is focused on semantic spatial information in workflow. The spatial relations was described by Description Logic, the ontologies was then constructed using the RDF Schemas and OWL Capabilities. These theories and methods are enclosed in a middleware, and the middleware was embedded in workflow system of the government in Jilin province.

2. SEMANTIC WEB ONTOLOGIES WITH DESCRIPTION LOGIC

The current web aims to expand the quantity but the semantic web, known as a trust web, means an expansion of quality. The semantic web appears correct to be called an evolution of the Web than calling it a reformation of the web. The semantic web will enable intelligent services such as information brokers, search agents, information filters etc. Such intelligent services on the knowledgeable web should surpass the currently available versions of these services, which are limited in their functionality, and only work as stand-alone services that do not interoperate. Two things constituting the semantic Web are ontology, which represents the semantic constitution, and markup language, which represents well-defined information. Humans and machines should communicate with each other in order to realize the semantic web to process and interpret information. The languages for representing information are XML for representing information structures and RDF, DAML, and OWL for representing the information meaning, have been developed and standardized in various ways using W3C [0,0,0].

The main idea of this paper is to handle the spatial relationships that occur when web ontologies are constructed. The existing methods for representing the relationships do not have the capability to manage the spatial relationships. These are five steps of building a web ontology.

Step 1: Collect all the concepts related to spatial regions.

Step 2: Classify the collected concepts.

Step 3: Define the relationship between concepts using the existing representation method. Most concepts are possible to define use of ‘subClassOf relationship.

(ex) Enyu subClassOf Yushu.

Step 4: Define the relationship by RCC and direction relations.

(ex) *Enyu* EC *Gongpeng, Gongpeng* SW *Yumin*.

Step 5: Build the web ontology on the nations using the OWL language.

3. IMPLEMENT AND APPLY TO WORKFLOW MANAGEMENT SYSTEM

The above theories and methods are enclosed in a middleware: Spatial Semantic Web Ontologies Middleware (SSWOM for short). Middleware is a class of software technologies designed to help manage the complexity and heterogeneity inherent in distributed systems. It is defined as a layer of software above the operating system but below the application program that provides a common programming abstraction across a distributed system. In doing so, it provides a higher-level building block for programmers than Application Programming Interfaces (APIs) such as sockets that are provided by the operating system. This significantly reduces the burden on application programmers by relieving them of this kind of tedious and error-prone programming.

SSWOM consists of four parts:

(1) Ontology Editor: By applying the above methods, Ontology Editor generates new ontology which is available to be browsed by the end user.
(2) Ontology Base: The ontology is described by OWL, the format of OWL document is like the sample of step 5.
(3) Mediator: Mediators look for spatial information and translate it into a format understandable by the end user.
(4) Ontology Server: The ontology server has a central role in SSWOM because it provides the control of other components.

Figure 1 is the architecture of SSWOM. The lowest layer is Spatial Database Interface which directly interacts with GIS platforms or Spatial Databases. The upper component is the Application Interface which provide API or

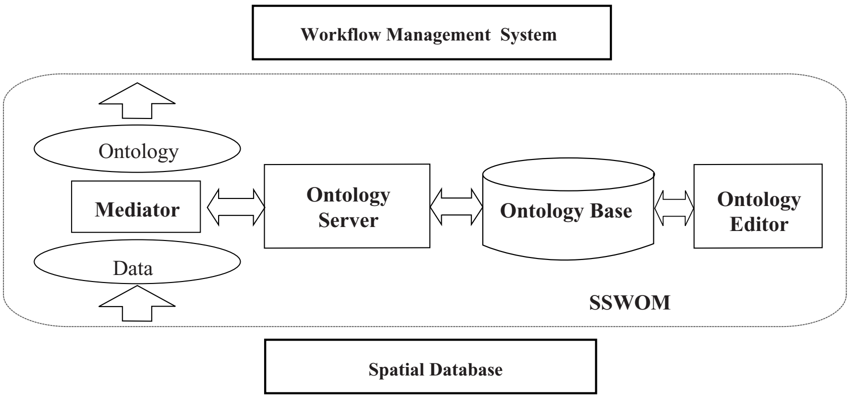


Figure 1. Spatial semantic web ontologies middleware.

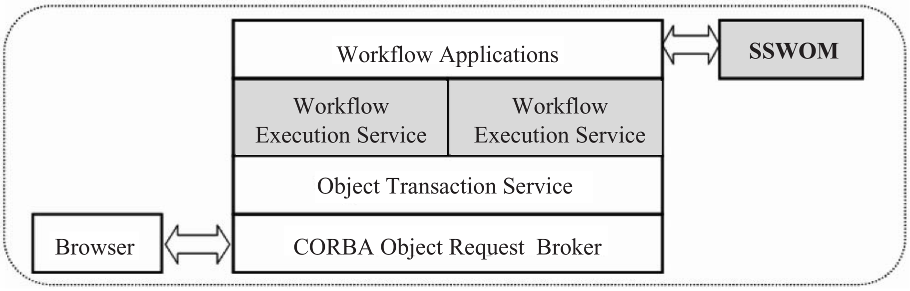


Figure 2. Distributed workflow management system structure.

CORBA based interaction with workflow systems. SSWOM was embedded in workflow management system of the government in Jilin province of China. The government always deal with many location relative affairs like that cadastre management, irrigation works and city planning. SSWOM gives the semantic spatial information interpretation to the workflow, thus the workflow management system is more efficient.

The whole distributed workflow management system structure is shown in Figure 2. The browser is a Java capable browser. The most important components of the system are the two transactional services, the workflow *repository* service and the workflow *execution* service. These two facilities make use CORBA Object Transaction Service. The SSWOM is directly interacts with the Workflow Applications layer. Graphical user interface to these applications has been provided by making use of Java applets which can be loaded and run by any Java capable Web browser.

4. CONCLUSION

We put forward a new method which using semantic web ontologies for workflows. This method successfully solves the problem of handling semantic spatial information in workflow management system. While other works based on *ALCRP(D)* only contains partial RCC topological relations[0], both full RCC relations and direction are considered in this paper. A future research will aim to define more complex spatial relationships such as distance.

REFERENCES

1. W.M.P. van der Aalst and K.M. van Hee (2002), *Workflow Management: Models, Methods, and Systems*, MIT Press, Cambridge, MA.

2. V. Haarslev, R. Moller and M. Wessel (2000), Visual spatial query languages: a semantics using description logic. *Diagrammatic Representation and Reasoning*.
3. A.G. Cohn and S.M. Hazarika (2001), Qualitative spatial representation and reasoning: an overview. *Fundamental Informatics*, 46 (1–2), pp. 1–29.
4. Cohn, Z. Cui and D. Randell (1992), A spatial logic based on regions and connection. In: *Proceedings Third International Conference on Principles of Knowledge Representation and Reasoning* (KR ’92).
5. V. Haarslev, C. Lutz and R. Moller (1999), A description logic with concrete domains and a role-forming predicate operator. *Journal of Logic and Computation*, 9(S), pp. 351–384.
6. Hyunjang Kong, Representing the Spatial Relations in the Semantic Web Ontologies. T.D. Gedeon and L.C.C. Fung (eds.). AI 2003, LNAI 2903, pp. 77–87.
7. P. F. Patel-Schneider, P. Hayes and I. Horrocks (2003), OWL Web Ontology Language Semantics and Abstract Syntax, W3C Working Draft 31 March, 2003, http://www.w3.org/TR/2003/WD-owl-semantics-20030331.
8. Yongjuan Zou (April 23, 2003), Integrating concrete domains into description logic: a survey.
9. Ankolekar, F. Huch and K. Sycara (2002), Concurrent Semantics for the Web Services Specification Language Daml-S. In: *Proceedings of the Coordination 2002*.

CONSTRAINED MULTI-SAMPLE TEXTURE SYNTHESIS

W.H. Li, Y. Zhang, Y. Meng, Z.J. Tan and Y.J. Pang
Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, College of Computer Science and Technology, Jilin University, Changchun, China 130012

Abstract It is a significant thing that generating specified texture picture in specified position. This paper proposed a novel texture synthesis method based upon multi-sample of texture. particle swarm optimization was applied to carry out single-sample texture synthesis, and then it was applied to constrained multi-sample texture synthesis in extension. The users can generate new pictures according to their own demands. Experiment shows that the algorithm is satisfied in the aspect of speed and quality.

Keywords: texture synthesis, particle swarm optimization, constrained multi-sample texture synthesis.

1. INTRODUCTION

Texture synthesis is of great importance to computer vision and graphics. It has remained an active research topic for years and many algorithms have been proposed. In the study of texture synthesis, most synthesis algorithm is only suitable for single-sample texture synthesis; however, it is considered very important that generating specified texture in specified position. There are many methods for controlling the growth of the texture. The simplest way is using the existent photo or simple design texture as object picture, in the guidance of the objective photo to implement the final synthesis result. Based on the schemes proposed in previous work by the authors [1–5], this paper presents constrained multi-sample texture synthesis. The users can generate new pictures according to their own demands.

G. R. Liu et al. (eds.), Computational Methods, 1297–1301.

2. PARTICLE SWARM OPTIMIZATION-BASED TEXTURE SYNTHESIS

Kennedy and Eberhart [6] proposed a particle swarm optimization (PSO) algorithm. This algorithm tends to find a best solution for the swarms by optimize each particle's solution according to the swarm's solution.

Texture synthesis should *look like* the sample texture and keep the randomness of textures. The PSO algorithm will either give us the best location or an approximate best location. Now we apply the PSO algorithm to texture synthesis [7]. Randomly set a number of positions in input image I and treat these points as virtual particles. These particles can be thought as virtual points of the image and each particle determine a patch by setting the left upper corner of the patch as the position of this particle. When the particles travel through the image we compare the patches they determine with the synthesized area and find the best-match patch.

Each particle will travel through I according to its location and velocity function. At each location we need to calculate the fitness of the current location. The fitness is then the distance between the boundary zone A determined by particle and boundary zone B, which is the patch just synthesized. The formula for calculating fitness between two patches A and B is as follows:

$$d(A, B) = \left[\frac{1}{M} \sum_{i}^{M} (p_A^i - p_B^i)^2 \right]^{1/2} \tag{1}$$

where M is the number of pixels in the boundary zone. p_A^i and p_B^i present the values (greyscale or colour) of the ith pixel in the boundary zone A and B.

Texture synthesis is performed in 2D space, so we need to adjust the original formula to 2D space. We define n particles on the sample texture. The location of a particle i will be recorded as $Present_i$. The best fitness location of particle i is marked as $Lbest_i$, with best fitness $LBest_i$. The global fitness is the best fitness among all n particles so $GBest = min(LBest1, LBest2, \ldots LBestn)$. The location of the particle, which has best global fitness, is recorded as *Gbest*. Each particle will be given an initial velocity vector V_i. We define another two vectors L_i and G_i.

$$L_i = Lbest_i - Present_i, G_i = Gbest - Present_i. \tag{2}$$

Now we update the particles using the formula 2.

The PSO algorithm is an iteration process. We will terminate the iteration when the program find an approximate best location. In the former case it gives an appropriate location. A function *dmin* is used to determine the best location. If the difference between a patch and the synthesized area is less than *dmin* then we get the best location and stop the iteration.

$$dmin = \varepsilon \left[\frac{1}{M} \sum_{i=1}^{M} (p_B^i)^2 \right]^{1/2} . \tag{3}$$

Where M is the number of pixels in the boundary area and p_B^i represents the value of the ith pixel in the boundary zone of the just synthesized patch. ε is the error tolerance and it is set to 0.2 in our algorithm because as Liang [2], this error threshold is most suitable for keeping the randomness while ensuring synthesis quality and avoiding repetition.

3. CONSTRAINED TEXTURE SYNTHESIS

A normal texture synthesis technique synthesizes textures in an image or a surface. The constrained texture method synthesizes textures on a specific area, which can be used in many ways, e.g., photo repair, designing and image filling. Ashikhmin [1] proposed a constrained texture synthesis method using a single texture sample. However sometimes we need two or more textures or pictures as input. In previous work by the authors [5], they did multi-sample constrained texture synthesis based on Ashikhmin [1]. However, his method synthesizes textures one pixel at a time and thus is very computationally expensive. We have extended swarm intelligence-based texture synthesis method to be used in multi-sample constrained texture synthesis.

We take $n + 1$ images as input: an target image D and several texture samples $S_1, S_2, \ldots, S_n$. Then we synthesize texture on different areas from different sample textures. Basically, we first synthesize n temporary textures of suitable size from the sample textures $S_1, S_2, \ldots, S_n$. Then we render one of the temporary textures to a specific area according to the user's need.

Figure 1 shows two example of constrained texture synthesis. For these applications, we take three images as input: one is a target image D and the others are two texture samples S_1 and S_2. D has a clearly defined background and foreground so that they are filled by textures synthesized from S_1 and S_2. We first generate two temporary texture S'_1 and S'_2 from the two sample textures. They both have the same size as D. Then we take any point from D and set it as background. Convert it to the YIQ system [8] and call the Y value of this point. For each pixel, we check whether the Y value of this pixel is within some tolerance. If it is, then we render S'_1 texture to this pixel, otherwise we render S'_2 texture to this pixel. We keep doing this until the entire image has been filled.

In previous work by the authors [4] the texture-by-numbers was processed by employing the method of image analogies, and the algorithm demonstrated

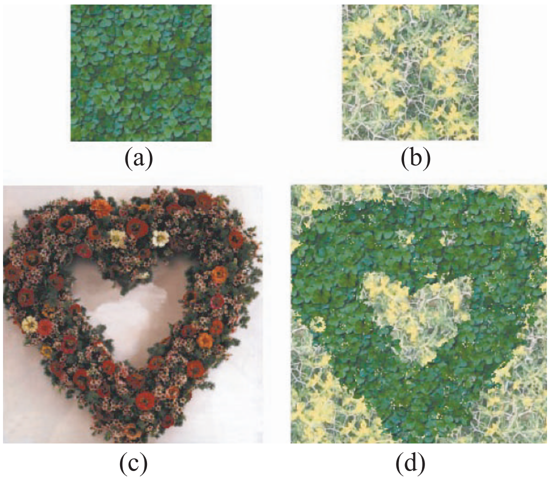


Figure 1. Constrained texture synthesis. (a), (b) are two sample textures; (c) is the target picture and (d) is the synthesized result.

some effect. However, the main idea is still utilizing the texture synthesis algorithm based on pixel. We improved the above algorithm, adopted the swarm intelligence-based texture synthesis to fulfil the texture-by-numbers. For an input sample texture p, we numbered the each part of the texture with different colours manually and generated a tab graph T_1, then we analyse the colour distribution of the tab graph T_1 and design the tab graph T_2 as a object graph to carry through constrained synthesis. Through the swarm intelligence-based texture synthesis, the texture, which has the same colour as the tag graph in the input sample texture p, was synthesized to the texture with the same size as tag graph T_2, and was stored in a temporary picture. The number of the temporary pictures is the same as the number of the colours in tag graph T_1. Then according to the constrained multi-sample texture synthesis algorithm, the pixel corresponded to colour position in tag graph T_2 was filled with corresponded colour, thus to realize the texture-by-numbers. Finally, we use feathering to attain the effect of the seamless mosaics between various textures. The result was shown in Figure 2.

4. CONCLUSION

This paper proposed a constrained texture synthesis algorithm based upon swarm intelligence which can generate new texture for multi-sample texture, thus to extend the application area of texture. We can design various new graphs according to the user's demand; we can also edit the existing photo to attain the

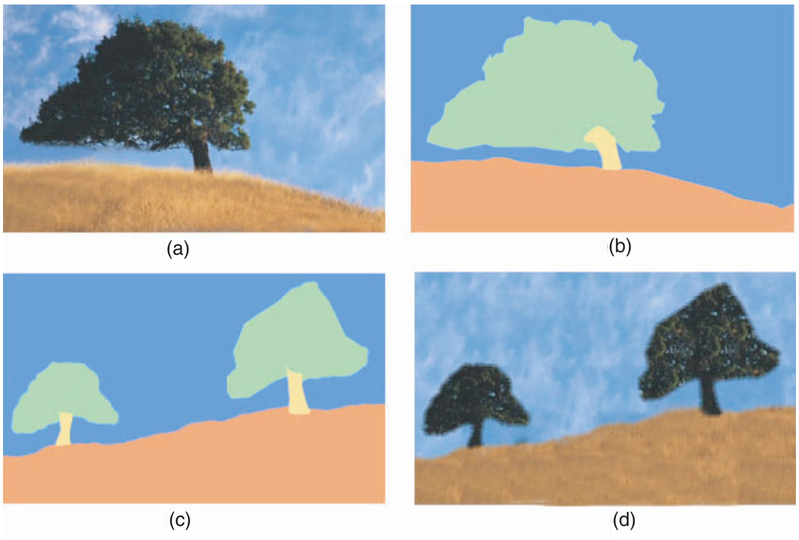


Figure 2. Texture-by-numbers. The source image (a) was painted by hand to annotate (b). The target image (c) was created in a paint program and refined with our interactive editor; the result is shown in (d).

seamless mosaics effect between textures. The speed of the algorithm is fast and the implement is simple.

REFERENCES

1. M. Ashikhmin (2001), Synthesizing natural textures. In: *2001 ACM Symposium on Interactive 3D Graphics*, ISBN 1-58113-292-1, pp. 217–226.
2. L. Liang, C. Liu, Y. Xu, et al. (2001), Real-Time texture synthesis using patch-based sampling. *Microsoft Research, Tech Rep: MSRTR-2001-40*.
3. A.A. Efors and W.T. Freeman (2001), Image quilting for texture synthesis and transfer. In: *Proceedings of ACM SIGGRAPH,* ACM Press, Los Angeles, pp. 341–347.
4. A. Herzmann, et al. (2001), Image analogies. In: *Proceedings of SIGGRAPH 2001*, ACM Press, New York, pp. 327–339.
5. Xu X.G. et al. (2002), Relativity principle-based multi-sample texture synthesis. *Progress in Natural Science* (Chinese edition). pp. 665–668.
6. J. Kennedy and R.C. Eberhart (1995), Particle swarm optimization. In: *Proceedings of IEEE International Conference on Neural Networks*, Vol. IV, IEEE service center, Piscataway, NJ, pp. 1942–1948.
7. Y. Zhang and Y. Meng, et al. (2004), Particle swarm optimization-based texture synthesis and texture transfer. In: *Proceedings of the Third International Conference on Machine Learning and Cybernetics*, pp. 4037–4042.
8. Y.D. Jia (2000), *Machine Vision* (in Chinese). Science Press, Beijing, p. 154.

A GENERAL INCREMENTAL HIERARCHICAL CLUSTERING METHOD

L.L. He[1], H.T. Bai[2], J.G. Sun[1] and C.Z. Jin[1]
[1] *College of Computer Science and Technology, Jilin University of P.R. China, QianWei Road 10, Changchun, Jilin, 130012*
[2] *New Technology Center, ShenZhen Telecom, P.R. China, ShenZhen, 518048*

Abstract Data mining, i.e., clustering analysis, is a challenging task due to the huge amounts of data. In this paper, we propose a general incremental hierarchical clustering method dealing with incremental data sets in data warehouse environment for data mining to reduce the cost further. As an example, we put forward ICHAMELEON, the improvement of CHAMELEON, which is a hierarchical clustering method, and demonstrate that ICHAMELEON is highly efficient in terms of time complexity. Experimental results on very large data sets are presented which show the efficiency of ICHAMELEON compared with CHAMELEON.

Keywords: hierarchical clustering method, ICHAMELEON, incremental method.

1. INTRODUCTION

Clustering in data mining is a discovery process that groups similar data into one set [1], which can be used in enterprises' marketing, such as identifying customers, adopting market tactics, raising the percentage of market shares, etc. Clustering on several million customers' data would consume several hours long, which is intolerable for On-Line Analysis Mining (OLAM). We discover that the data stored in data warehouse are incremental upgraded regularly and timely. It is hopeful to improve the clustering speed by a large margin if an algorithm can combine the historical clustering result with 'incremental' data items. A general incremental hierarchical clustering method dealing with incremental data items is proposed according to this situation. As an example, we put forward ICHAMELEON, the improvement of CHAMELEON, and demonstrate

G. R. Liu et al. (eds.), Computational Methods, 1303–1307.

that ICHAMELEON is highly efficient in terms of time complexity. Experimental results on very large data sets show the efficiency of ICHAMELEON compared with CHAMELEON.

2. GENERAL INCREMENTAL HIERARCHICAL CLUSTERING METHOD

The hierarchical method, such as AGNES [2], CURE [3] and CHAMELEON [4], is a very important kind of clustering algorithms. A hierarchical method makes up data items on a tree of clustering and executes agglomerative or divisive hierarchical clustering. Therefore, if data items are represented by the points of n-space, and the weighted value on arc linking two points is the similarity measurement; the clustering process is the process of merging connected subgraphs.

General incremental hierarchical clustering method includes two processes: first, incremental data set which may add new data items to and/or delete data items from the original data set, is combined together with historical clustering result which produces sub-clusters represented by connected sub-graphs; second, each two connected sub-graphs are merged continuously until suitable clusters are gotten.

Added data items as a whole form some sub-clusters. Different hierarchical algorithm has its choice, for example, CHAMELEON may use a graph partitioning algorithm to cluster k-nearest neighbour graph which is composed of the added and original data items, into a large number of relatively small sub-clusters; while AGNES may merge several times from every added data items as a cluster.

Deleted data items affect the historical clustering result. Original clusters may be preserved or divided according to the connectivity of graphs. Examples are shown in Figures 1 and 2. We may adopt depth or width priority searching algorithm to judge the connectivity of graphs.

Based on the general incremental hierarchical clustering method above, we improve the CHAMELEON algorithm with incremental feature. An

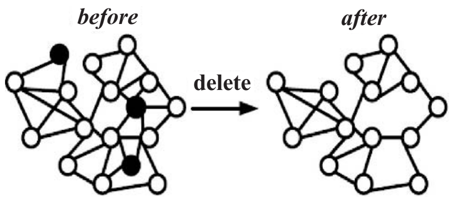


Figure 1. Example of undivided Cluster.

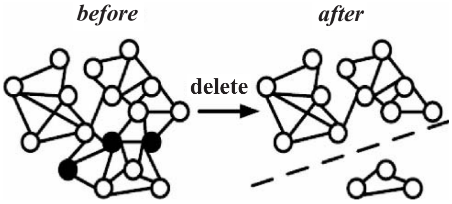


Figure 2. Example of divided Cluster.

incremental CHAMELEON algorithm—ICHAMELEON is presented in the following section.

3. ICHAMELEON ALGORITHM

3.1 ICHAMELEON Algorithm

```
//input: k: the number of items' nearest neighbours; DataSet: InPointSet; $\alpha$:the weight of RC;
//HisResult: HistoryCluster; T: the threshold of $RI(C_i, C_j)^* RC(C_i, C_j)^\alpha$
//output: the incremental clustering result
ICHAMELEON(int k, double $\alpha$, DataSet InPointSet, HisResult:HistoryCluster)
Begin

Step 1. InDataNearestGraph = ConstructSparseGraph(InPonitSet, k);
            // construct k-nearest neighbour graph.
Step 2. InDataSubClusters = PartitionGraph(InDataNearestGraph);
            // initial sub-clusters are obtained by partitioning k-nearest neighbour graph.
Step 3. DelSubCluster = GraphSearch (InPonitSet);
            //judge the connectivity of graph.
Step 4. //the number of sub-clusters is InM
        MaxRIRC < –0;
        while (1) {
            for (i=1, i<=InM–1; i++)
            for(j=i+1;j<=InM, j++) {
                Compute RI and RC;
                if MaxRIRC< $RI(C_i, C_j)^* RC(R_i, C_j)^\alpha$
                    MaxRIRC= $RI(C_i, C_j)^* RC(C_i, C_j)^\alpha$ ;
            }
            if MaxRIRC>T { MergeCluster($C_i$, $C_j$); InM–; MaxRIRC<–0;}
            else break; //final clusters are obtained }
End;
```

3.2 Performance Analysis

Let n_1, n_2 and n_3 respectively be the number of added items, original data items and deleted items, $n = n_1 + n_2$. Let m_1 be the number of sub-clusters produced by added items.

It has been shown that for n_1 added data items, the average cost of inserting based on *k-d* trees [5], as well as the expected *k*-nearest neighbour search time is $O(\log n)$, leading to an overall complexity of Step I is $O(n_1 \log n)$.

To simplify the analysis, we assume that (i) each sub-cluster consisting of added items has the same number of nodes n_1/m_1; (ii) during each successive merging step, ICHAMELEON selects only a single pair of clusters to merge, and (iii) each cluster includes the same number of data items after division of historical clusters.

The amount of time required of the Step II of ICHAMELEON depends on the amount of time required by hMETIS [6]. Given a graph $G = (V, E)$, hMETIS requires $O(|V| + |E|)$ time to compute a bisection. Thus, the overall computational complexity of Step II is $O(n_1 \log(n_1/m_1))$ which is bounded by $O(n_1 \log n_1)$.

Let n_f be the number of historical clusters. The time complexity of the Step III of ICHAMELEON is $O(n_3 \log(n_2/n_f))$ [7]. Let m_f be the number of intermediate sub-clusters after Step III. The time complexity of Step IV is $O(m_f^2 \log m_f + nm_f)$.

Above all, when $n_2 >> n_1$ and $n_2 >> n_3$, the overall time complexity of ICHAMELEON is $O(m_f^2 \log m_f + nm_f + n_1 \log n)$.

3.3 Experimental Results

ICHAMELEON and CHAMELEON are tested by data from a company's data warehouse. Beginning with 100,000 initial data items, increments are 10%, 20% ... 70% the number of initial items. Testing environment includes DELL OPTIPLEX GX260, Intel Pentium4 2.4 GHz CPUs and 1G PC266A DDR memory. The real execution time of the two is presented in figure 3, which shows the ICHAMELEON algorithm reduces the time of clustering largely.

4. CONCLUDING REMARKS

We propose a general incremental hierarchical clustering method in this paper by taking ICHAMELEON, the incremental version of CHAMELEON, as an example. ICHAMELEON keeps the dynamic clustering characteristic of

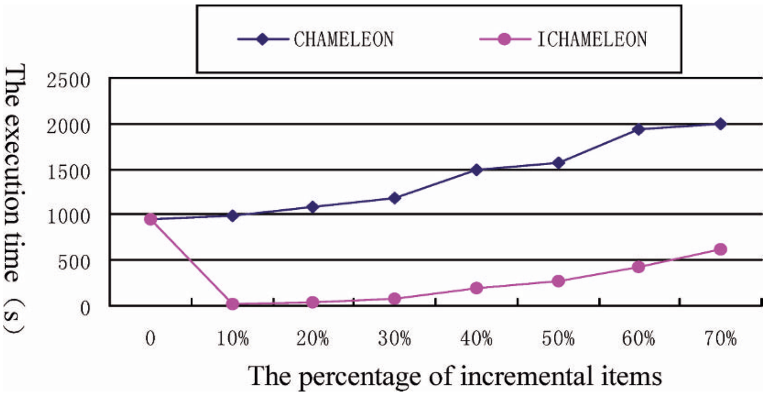


Figure 3. Comparison of two algorithms.

CHAMELEON; moreover, it utilizes the historical clustering result to expedite the clustering process. Experimental results on very large data sets show that ICHAMELEON is especially suitable for incremental clustering analysis in data warehouse environment.

5. ACKNOWLEDGEMENTS

This research benefits from the support of Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University.

REFERENCES

1. J.W. Han and M. Kamber (2001), *Data Mining Concepts and Techniques*. Higher Education Press, Beijing.
2. L. Kaufman and P.J. Rousseeuw (1990), *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons Press, New York.
3. S. Guha, R. Rastogi and K. Shim (1998), Cure: an efficient clustering algorithm for large databases. *International Conference on Management of Data*, pp. 73–84.
4. G. Karypis, E.-H. Han and V. Kumar (1999), CHAMELEON: a hierarchical clustering algorithm using dynamic modeling. *IEEE Computer*, pp. 68–75.
5. H. Samet (1990), *The Design and Analysis of Spatial Data Structures*. Addison-Wesley Press.
6. G. Karypis and V. Kumar (1998), *hMETIS 1.5: a hypergraph partitioning package*. Technical report, Available at http://www.cs.umn.edu/∼ metis.
7. H.T. Bai, J.G. Sun, Y. Jiao and C.Q. Xu (2002), Implementation and comparison of network optimization algorithms. *Journal of Jilin University* (Information Science edition), 40, 4, pp. 59–68.

ACTIVATING IRREGULAR DIMENSIONS IN OLAP*

Z.H. Li, J.G. Sun, H.H. Yu
College of Computer Science and Technology, Jilin University, Changchun, P.R. China 130012
Key Laboratory of Symbolic Computation and Knowledge Engineer of Ministry of Education, Changchun, P.R. China 130012

Abstract OLAP (on-line analytical processing) systems support decision-making process by providing dynamic analytical operations on high volumes of data. Usually, the operations require dimensions to be regular, however, in real-world applications, many complex dimensions fail to meet the requirement. In this paper, we first propose a new conceptual model, compared with traditional multidimensional models, this model extends the surjection between the domains of two levels to the partial mapping. Afterwards, we present a transforming algorithm to prove that the model offers a practical way for handling irregular dimensions in OLAP systems.

Keywords: irregular dimension, model, partial mapping, OLAP.

1. INTRODUCTION

In OLAP systems, drill-down and roll-up are the two basic analysing operations. In previous work by the authors [1–3], dimensions are formally defined on the basis of the surjection between the domain of a child level and the domain of a parent level. Therefore, the two operations usually require dimension hierarchies to be onto and covering. However, in real-world applications, many irregular dimensions fail to meet this requirement, and they could be non-onto, non-covering or non-balanced [4].

Some models for handling irregular dimensions have appeared in recent work by the authors [5,6]. Pedersen et al. [4] propose a model to mainly deal

---

*Supported by National Natural Science Foundation of China (No. 60073039, 60273080)

G. R. Liu et al. (eds.), Computational Methods, 1309–1313.

with many-to-many relationships between facts and dimensions, and he also concerns the modeling of irregular dimensions. However, the partial order in the model is not defined based on the mappings between dimension levels. Jensen and Kligys [5] present a new approach to the modeling of partial containment relationships among elements of spatial dimensions; while we mainly concern the modeling of relationships among levels in irregular dimensions. Therefore, they should be modeled separately. Jagadish and Lakshmanan [6] propose the SQL(H) model that allows the designer to use multiple tables to represent a single level. Thus, it would be very complicated to execute an aggregation query, for we couldn't know in advance which table should be used to join with the fact table.

In this paper, we propose a new conceptual model to handle irregular dimensions by defining the partial mapping between the domains of two dimension levels, and we impose closeness on the domain of a dimension to ensure correct aggregation paths. Therefore, the model offers a practical way for implementing OLAP operations on irregular dimensions. The rest of the paper is organized as follows. In Section 2, we formally define the conceptual model. In Section 3, based on the conceptual model, we present an algorithm to transform a dimension into a relational table.

2. CONCEPTUAL MODEL

In this section, we define a new conceptual model by extending the surjection between the domains of two dimension levels in previous work [1–3] to the partial mapping. Then, to handle irregular dimensions, we define two partial orders $\prec$ and $\leq$ on the basis of the partial mapping. And, to ensure correct aggregation paths in a dimension, we require the domain of a dimension to be closed w.r.t. the partial order $\leq$.

Definition 1. Given $D = \{l_1, l_2, \ldots, l_k\}$, where l_i denotes a set and $dom(l_i) = A_i$ ($1 \leq i \leq k$). Let $\prec$ be a bi-relationship on D, for any l_i, $l_j \in D$ ($1 \leq i, j \leq k$), $l_i \prec l_j$ if $\prec$ satisfies the following conditions:

1. For subsets $P \subseteq dom(l_i)$, $Q \subseteq dom(l_j)$, there is an surjection σ: $P \rightarrow Q$.
2. If there is another surjection τ: $R \rightarrow S$, then $R \subseteq P$, $S \subseteq Q$.

In fact, function σ is the partial mapping from $dom(l_i)$ to $dom(l_j)$. We call $\prec$ the aggregation relationship (AR) on D, and there are four cases for the possible relationships between P, Q and $dom(l_i)$, $dom(l_j)$:

1. If $P \subset dom(l_i)$ and $Q \subset dom(l_j)$, then l_i,l_j satisfy the relationship $\prec$ in a 'weak–weak' way, written $l_i(w \prec w)\, l_j$.

2. If $P = \mathrm{dom}(l_i)$ and $Q \subset \mathrm{dom}(l_j)$, then l_i,l_j satisfy the relationship $\prec$ in a 'strong–weak' way, written $l_i(s \prec w)\, l_j$.
3. If $P \subset \mathrm{dom}(l_i)$ and $Q = \mathrm{dom}(l_j)$, then l_i,l_j satisfy the relationship $\prec$ in a 'weak–strong' way, written $l_i(w \prec s)\, l_j$.
4. If $P = \mathrm{dom}(l_i)$ and $Q = \mathrm{dom}(l_j)$, then l_i,l_j satisfy the relationship $\prec$ in a 'strong–strong' way, written $l_i(s \prec s)\, l_j$.

Let $\prec'$ be the reflexive and transitive closure of $\prec$, and substitute $\prec'$ for $\prec$ as the AR on D. Based on $\prec'$, we define a partial order on $\underset{1\le i\le k}{\cup} \mathrm{dom}(l_i)$ as follows.

Given $D = \{l_1,l_2,\ldots,l_k\}$ and the AR $\prec'$ on D, let $E = \underset{1\le i\le k}{\cup} \mathrm{dom}(l_i)$ and $\le$ a bi-relationship on E, we call $\le$ the element aggregation relationship (EAR) on E, if it satisfies the following conditions:

1. Given $l_i, l_j \in D(1 \le i, j \le k)$, such that $l_i \prec' l_j$, let σ_{ij} be the corresponding mapping from P to Q , where $P \subseteq \mathrm{dom}(l_i)$, $Q \subseteq \mathrm{dom}(l_j)$. For any $a \in \mathrm{dom}(l_i)$, $b \in \mathrm{dom}(l_j)$, if the pair $(a, b) \in \sigma_{ij}$, then we say that a, b satisfy the relationship $\le$, written $a \le b$.
2. Given $l_i,l_j \in D(1 \le i, j \le k)$, for any $a \in \mathrm{dom}(l_i)$, $b \in \mathrm{dom}(l_j)$, if $a \le b$, then $l_i \prec' l_j$.

In fact, $\le\, = \underset{1\le i,j\le k}{\cup} \sigma_{ij}$. Given an element $a \in E$, if no other element $b \in E$ satisfies $a \le b$ or $b \le a$, then a is called an isolated element in E. For any element $a \in E$, if there is always another element $b \in E$ satisfying $a \le b$ or $b \le a$, then we say that $(E, \le)$ is **closed**, or E is closed w.r.t. $\le$. To model a dimension, two special levels are also needed, i.e., the common top-level *All* and the minimal level denoted by *Atomic* here. Then a dimension can be defined as follows.

Definition 2. Given $D = \{l_1,\ldots l_k,l_{k+1}\}$, where $\mathrm{dom}(l_i) = A_i$ $(1 \le i \le k + 1)$ and l_{k+1} *All*. Let $E = \underset{1\le i\le k+1}{\cup} \mathrm{dom}(l_i)$, $\prec'$ the AR on D, $\le$ the EAR on E, if there is a unique level *Atomic* in D and $(E, \le)$ is closed, then $d = (D,\prec')$ is called a dimension, where l_i is a level attribute, d is called the dimension attribute, and let $\mathrm{dom}(d) = E$.

In fact, $(D, \prec')$ is the schema of dimension d. In a dimension, $\prec'$ describes the aggregation path between two levels, and $\le$ shows the aggregation path between two elements. And, the closeness is used to ensure correct aggregation paths in dimension hierarchies. A dimension hierarchy is defined as follows.

Definition 3. Given a dimension $d = (D,\prec')$, where $D = \{l_1,\ldots l_k,l_{k+1}\}$, let $\underline{H} = \{l'_1,\ldots,l'_m\}$ a totally ordered subset of D, if $All \in \underline{H}$ and $Atomic \in \underline{H}$, then $h = (\underline{H},\prec')$ is called a hierarchy, where h is the hierarchy attribute, and let $\mathrm{dom}(h) = \mathrm{dom}(l'_1) \cup \mathrm{dom}(l'_2) \cup \ldots \cup \mathrm{dom}(l'_m)$.

Given a dimension d, and two levels $l_i, l_j \in D$, suppose all mappings on D are onto and covering except the mappings related with l_i and l_j, and only the mapping between l_i and l_j is concerned here, then dimension d and the four cases in definition 1 could have the following relationships:

1. If the first case is true, then d is non-covering and non-onto.
2. Else if the second case is true, then d is non-onto if $l_i = Atomic$, otherwise, it is non-covering and non-onto.
3. Else if the third case is true, then d is non-covering because of the closeness.
4. Else if the fourth case is true, then d is regular or at most non-balanced.

A data cube is constructed with a set of dimensions and a set of facts, and it consists a collection of cells. Then a data cube is defined as follows.

Definition 4. A data cube is defined as C = (D, M, f), where:

1. $D = \{d_1, \ldots, d_n\}$ is a set of dimensions, and $d_i = (D_i, \prec_i')$ $(1 \leq i \leq n)$,
2. $M = \{m_1, \ldots, m_k\}$ is a set of facts, and $m_j = (M_j, agr_j)$ $(1 \leq j \leq k)$ is a fact, M_j is a numeric set, agr_j is a function: $2^{M_j} \rightarrow M_j$ and let $\mathrm{dom}(m_j) = M_j$.
3. let $\mathrm{dom}(D) = \mathrm{dom}(d_1) \times \ldots \times \mathrm{dom}(d_n)$, $\mathrm{dom}(M) = \mathrm{dom}(m_1) \times \ldots \times \mathrm{dom}(m_k)$, then $\mathrm{dom}(C) = \mathrm{dom}(D) \times \mathrm{dom}(M)$.
4. f is a mapping: $P \rightarrow \mathrm{dom}(M)$, where $P \subseteq \mathrm{dom}(D)$ and $P \neq \phi$.

In the definition, (D, M) describes the schema of cube C, written schema(C) = (D, M), and f defines the domain of cube C. Obviously, the cube model allows members of multiple levels in the same cube, thus it can handle various complex hierarchies, e.g., non-onto hierarchies and non-covering hierarchies.

3. TRANSFORMING ALGORITHM

In this section, we devise an algorithm to transform a dimension into a relational table on the basis of the data model. Formal descriptions of the algorithm are given below.

1. Schema transformation:
   Given a dimension $d = (D, \prec')$, let R = (Tname, A, f, g) be the structure of the transformed table, where Tname = d is the table name, A = {Name, Parent, Level, Type} is the schema of the table, f is a functional dependency in A: Parent → Name and g defines the domains of Level and Type: dom(Level) = D, dom(Type) = {leaf, middle, root}.
2. Domain transformation:

Given a dimension d, the EAR $\leq$ on dom(d), and R = (Tname, A, f, g), let TR denote the domain of R, according to the schema A in R, TR = {(n, p, v, t)|n $\in$ dom(d), p $\in$ dom(d) and n $\leq$ p, especially if n = '*All*' then p = '*All*',

$$\mathrm{v} = \begin{cases} '\mathrm{leaf}', & \text{if } \mathrm{n} \in \mathit{Atomic} \\ '\mathrm{root}', & \text{else if } \mathrm{n} =' \mathit{All}', t = l_{\mathrm{i}} \text{ if } l_{\mathrm{i}} \in \mathrm{D}(1 \leq i \leq \mathrm{k}+1) \text{ and } \mathrm{n} \in \\ & \mathrm{dom}(\mathrm{l_i})\}. \\ '\mathrm{middle}', & \text{else} \end{cases}$$

4. CONCLUSIONS

In this paper, we focus on irregular dimension modeling, and propose a new conceptual data model, compared with traditional multidimensional models, this model extends the surjection between the domains of two levels to the partial mapping. And we impose closeness on the domain of a dimension to ensure correct aggregation paths in irregular dimensions. We also present a transforming algorithm to prove that the model can provide an effective way to handle irregular dimensions in building data cubes. As future work, we plan to define dimension operations and cube operations, and model many-to-many relationships between facts and dimensions.

REFERENCES

1. T.B. Nguyen, A.M. Tjoa and R. Wagner (2000), An Object Oriented Multidimensional Data Model for OLAP. In: *1st International Conference on Web-Age Information Management*, Heidelberg, Berlin, pp. 69–82.
2. Q. Mangisengi and A.M. Tjoa (1999), Multidimensional Modeling Approaches for OLAP Based on Extended Relational Concepts. In: *9th International Database Conference on Heterogeneous and Internet Databases*, Hong Kong, China, pp. 329–345.
3. A. Datta. and H. Thomas (1999), The cube data model: a conceptual model and algebra for on-line analytical processing in data warehouses. *Decision Support Systems*, 27, 3, pp. 289–301.
4. T.B. Pedersen, C.S. Jensen and C.E. Dyreson (2001), A foundation for capturing and querying complex multidimensional data. *Information System*, 26, 5, pp. 383–423.
5. C.S. Jensen and A. Kligys. (2002), Multidimensional Data Modeling For Location-Based Services. In: *10th ACM International Symposium on Advances in GIS*, USA, pp. 55–61.
6. H.V. Jagadish and L.V.S. Lakshmanan (1999), What can Hierarchies Do for Data Warehouses? In: *25th International Conference on Very Large Data Bases*, USA, pp. 530–541.

A NEW COMPUTATIONAL METHOD OF INTERSECTION FOR RAY TRACING

Quan Yong, Li Wenhui and Yang Qingyun
Department of Computer Science, Jilin University, Qianwei Road, Changchun 130012, P. R. China

Abstract In this paper a new point-based computational method of intersection for ray tracing is presented. With global illumination we gain a higher rendering speed of ray tracing compared with previous methods. During pre-process, an important attribute is added to each point. During rendering, an intersection algorithm different from previous ones has been demonstrated to yield satisfactory results.

Keywords: points, neighbour radius, ray tracing, global illumination.

1. INTRODUCTION

Recently a significant trend in 3D modelling has been the shift towards using points as representing primitives. Point-based rendering [1] which is actually an old idea revisited [2] is receiving a growing amount of attention. A direct ray tracing algorithm for point-sets has been developed by Schaufler and Wann Jensen [3]. They define the intersection of a ray and the point set as follows: at each point a disc is constructed using the point normal. A cylinder around the ray is intersected with the discs. The intersection is computed as a weighted average of discs whose centres are inside the cylinder. Point set surfaces (PSS) [1] are a smooth manifold surface approximation from a set of sample points. The surface definition is based on a projection operation that constructs local polynomial approximations and respects a minimum feature size. In Ref. [4], Adamson and Alexa study techniques for ray tracing PSSs. Due to the lengthy projection procedure ray tracing PSSs leads to slower speed. A surface approximation technique that is based on an iterative ray-surface intersection algorithm is proposed in Ref. [5]. The improved approach is comparable in speed to Schaufler and Jensen's approach.

G. R. Liu et al. (eds.), Computational Methods, 1315–1320.

2. PRE-PROCESS OF POINT DATA

The pre-process here mainly refers to the process of computing point attributes. In this paper we add an important attribute that we call ‘three-dimensional neighbour radii’ to sample point. Thus point attributes comprise position, normal, material and neighbour radii. Our emphasis is the computation of neighbour radii.

First we would like to explain the concept of ‘neighbour radius’. To decide if a ray intersects a point is to decide if the shortest distance between the ray and the point is within the scope of tolerance. Suppose coordinates of current point is P(x,y,z), coordinates of starting point and another point of the ray respectively are P1($x1$,$y1$,$z1$), P2($x2$,$y2$,$z2$). Then the square of distance between point P and ray P1P2 is:

$$\begin{aligned} d^2 = {} & (x - x1)^2 + (y - y1)^2 + (z - z1)^2 - [(x2 - x1)(x - x1) \\ & + (y2 - y1)(y - y1) + (z2 - z1)(z - z1)]/[(x2 - x1)^2 \\ & + (y2 - y1)^2 + (z2 - z1)^2] \end{aligned} \tag{1}$$

When $d^2 < \varepsilon$, we assume that the ray intersects the point. Moreover, we still need to decide if the point is within valid part of the ray. Here, ε is a small positive value given by us. Each sample point on object surface corresponds to a this kind of positive value. We call the small positive value ε ‘nighbour radius’ of point. Commonly, we look on viewpoint as starting point of the ray and we choose one point on projection screen (different from display screen) as another point of the same ray. Thus intensities of points on object surface can be displayed on projection screen by ray tracing.

Actually, in order to accelerate the speed of intersection algorithm and make boundary processing of image easier, we extend the original meaning of neighbour radius and afterwards improve the intersection method. After extending the meaning of neighbour radius, each point corresponds to not one small positive value, but three small positive values. One is for x direction, one is for y direction and the rest one is for z direction. This is what we called ‘three-dimensional neighbour radii’.

Next, we will show how to compute neighbour radius. We may have three methods for computation. The first one is to let neighbour radius of each point be a proper positive value. We can usually get it according to experience. The second method is to compute neighbour radius of current point according to density of points spatially around it. The third method is to use ‘three-dimensional neighbour radii’. For regular object, there is little difference between images rendered respectively using the three methods. This is because density of points on regular surface doesn’t vary much. For irregular object, sometimes density of points varies much. This may lead to quite different rendering effects between the three methods. So, we are inclined to use the second method or the third method. Consider that effect of boundary processing of the third method is better than the second one. This paper adopts the third method.

The main idea of the second method is: we make ‘neighbour radius’ smaller where there are more points and make it larger where there are fewer points. Suppose current point is P0. We manage to get position vectors of n points which are the nearest to P0. After that, we can calculate position vector of point that is the farthest from P0 among those n points. Suppose the farthest point among those n points is P1. Then the distance between P0 and P1 is regarded as the neighbour radius of point P0.

The main idea of computing neighbour radius in this paper is: we make ‘neighbour radii’ smaller where there are points whose parallel projections are dense on the planes $x = k$, $y = k$ and $z = k$ (k is an arbitrary real number) and make them larger where there are points whose parallel projections are sparse on the same planes. Suppose current point is P0. Then we manage to get n points which are the nearest to P0. Suppose P1, P2, …, Pn are the n points we have found and ε denotes three-dimensional neighbour radii. Then we compute ε in x direction as follows:

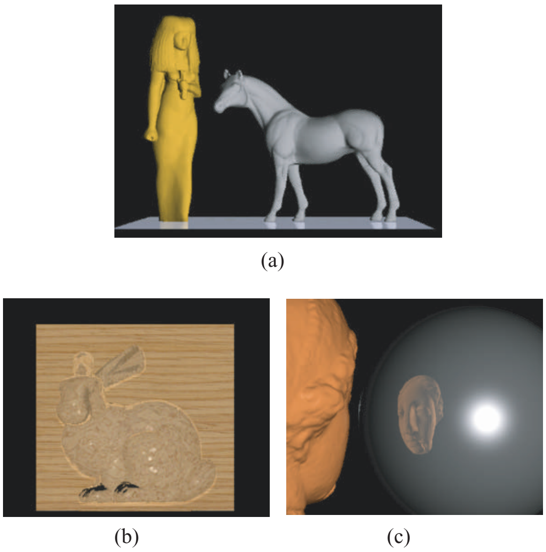
$$\varepsilon_x = \max\{|\mathrm{P1}.x - \mathrm{P0}.x|, \mathrm{P2}.x - \mathrm{P0}.x|, \ldots, |\mathrm{P}n.x - \mathrm{P0}.x|\} \tag{2}$$

ε in other two directions can be computed in a similar way.

3. RENDERING BASED ON POINT AND NEIGHBOUR RADIUS

In this paper a new computational method of intersection based on point and neighbour radius is presented. In the course of ray tracing, each ray is tested for intersection against points. In Algorithm 1, we just use neighbour radii of x direction and y direction as an example. Actually, neighbour radii of two directions are chosen based on the largest component of the incident ray’s direction vector. For example, if z is the largest component, then we chose neighbour radii of x direction and y direction. In order to accelerate further the speed of intersection computation, this paper presents a simplified method that narrates how to calculate intersection between a ray R and a surface point P. It is concretely described as Algorithm 1.

Algorithm 1. dx and dy are temporary variables. nerax and neray respectively denote neighbour radii of x direction and y direction. Current ray is R. Current point is P. The largest component of current ray’s direction vector is z.



(a)

(b) (c)

Figure 1. Rendering effects of our method.

```
obtain coordinates P($x1$,$y1$,$z1$) of current point.
compute intersection point Q($x2$,$y2$,$z1$) between current ray R and
plane $z = z1$
$dx$ = absolute value of ($x1 - x2$)
d$y$ = absolute value of ($y1 - y2$)
if d$x$ < P.nerax and d$y$ < P.neray then
decide that ray R intersects point P and return P
else decide that there are no intersection between R and P.
```

4. IMPLEMENTATION AND CONCLUSIONS

All the cases were tested on a PC with a Pentium 4 processor of 2 GHz and 256 MB RAM. We have implemented our approach on four models: the Cyberware Isis, the Cyberware Venus, the Stanford bunny and the Stanford horse. Figure 1a shows the Cyberware Isis and the Stanford horse rendered using our approach. Figure 1b shows rendering effect of refraction. In Figure 1b, we place a bunny made of glass in front of a wooden plane. Figure 1c shows reflection of the Venus model on a sphere surface. We compared our approach in rendering time with the approach in Ref. [3]. The statistics of performance are shown in Table 1. In the table we take the horse and the Venus as examples and use Approach 1 to denote Schaufler and Jensen's approach. From Ref. [5], we know that the approach in Ref. [3] is about two orders of magnitude faster than ray tracing point set surfaces Ref. [4]. This is why we use only

Table 1. Comparison in rendering time.

Model	Horse[b]	Venus[c]
Number of points	48,485	134,345
Rendering time of our approach (s)	18.5	23.2
Rendering time of approach 1 (s)	25.1	33.6

Note. b and c respectively denote 5 and 6 octree levels.

Schaufler and Jensen's approach for comparison in Table 1. Note that speed is inversely proportional to time. From the table we can know that our approach is about 1.4 times faster. For reducing the number of intersection an octree [6] is used.

In this paper we have presented a new point-based method for rendering with global illumination. During pre-process we add an important attribute that we call 'three-dimensional neighbour radii' to each point. Due to the simplification of our intersection algorithm, we gain a higher rendering speed compared with previous methods. Our approach is faster than the approach in Ref. [3] and thus much faster than the approach in Ref. [4]. We directly use point primitives to represent objects and thus surfaces we use needn't be parameterized. Compared with approach in Ref. [1], ours avoids the process of reconstructing the underlying surface. In Ref. [3], a fixed radius is used to locate neighbouring points. In this paper variation of neighbour radii of a certain point is related to points around it. We consider it to be advantageous to vary neighbour radii of points according to the local density of them.

ACKNOWLEDGEMENTS

We would like to thank our colleagues at the Graphics Laboratory of Jilin University in China. The Isis and the Venus model are courtesy of Cyberware. The bunny and the horse model are courtesy of the Stanford Computer Graphics Laboratory. This work was supported by NSFC (Natural Science Foundation of China) grant number 69883004 and Special Foundation for Doctoral Students grant number 20010183041.

REFERENCES

1. M. Alexa, J. Behr, D. Cohen-Or, S. Fleishman, C. Silva and D. Levin (2001), Point set surfaces. *IEEE Visualization 2001*, San Diego, USA, pp. 21–28.
2. M. Levoy and T. Whitted (1985), The use of points as a display primitive. TR 85–022. CS Department, University of North Carolina at Chapel Hill.

3. G. Schaufler and H. Wann Jensen (2000), Ray tracing point sampled geometry. Rendering Techniques 2000, EG Workshop on Rendering, Brno, Czech Republic, pp. 319–328.
4. A. Adamson and M. Alexa (2003), Ray tracing point set surfaces. *Proceedings of Shape Modeling International 2003*, Washington, USA.
5. A. Adamson and M. Alexa (2003), Approximating and intersecting surfaces from points. *Eurographics Symposium on Geometry Processing*, Aachen, Germany, pp. 230–239.
6. N. Ahuja and J. Veenstra (1986), Octree generation and display. Technical Report UILU-ENG-86-2215. Coordinated Science Laboratory, University of Illinois, Urbana.

STUDY ON PARTNERS SELECTION OF AGILE VIRTUAL ENTERPRISE BASED ON PARTICLE SWARM OPTIMIZATION

You Zhou, Yanchun Liang, Xiaohu Shi and Hongwei Ge
College of Computer Science, Jilin University, Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, Changchun 130012, China

Abstract An implementation scheme for solving complicated multi-restriction optimization problems of decision of agile virtual enterprise partners selection is presented based on particle swarm optimization (PSO) algorithm. The particle-flying space, objective function and fitness function are constructed. Experiments are performed to examine the proposed approach. Comparisons with the analytic hierarchy process (AHP) and genetic algorithm (GA) show that the proposed method is effective and feasible.

Keywords: agile virtual enterprise, partners selection, particle swarm optimization.

1. INTRODUCTION

Agile virtual enterprise (AVE) is the main manufacture pattern in the future. It was proposed in the report of the 21st century manufacture developmental strategy by USA gov in 1991 [1]. AVE emphasizes how to improve enterprise's adaptive abilities to the variable market. To develop new product or new management opportunity, AVE dynamically sets up the virtual organization around the world. The reason to develop AVE is that the chief enterprise is lack of resources, so it has to divide the main project into several sub-projects according to the manufacturing process of production. Some of these sub-projects are done by the partners selected after comprehensive consideration on the relationship between sub-projects and the capability, location and time cost of these partners. The AVE partners selection represents one category of decision on complex combinational optimization problems. The decision directly affects the efficiency of the AVE foundation. In this paper, the AVE

G. R. Liu et al. (eds.), Computational Methods, 1321–1326.

partners selection is handled using the particle swarm optimization (PSO) based algorithm.

2. DESCRIPTION ON THE PARTNER CHOICE

The formalization definitions are as follows:

$$\begin{aligned} P &= \{m_i/i \in \{1, 2, \ldots, NUM(P)\}\} = CMS \cup PMS, && (1) \\ SMRS &= \{r_i/r_i \in \mathrm{Re}l(P)\}; && \\ \mathrm{Re}l(P) &= \{[m_1, m_2, \ldots, m_\mathrm{k}, S]/k \in \{1, 2, \ldots, NUM(P)\}\}, && (2) \\ PGS &= \{pg_i/pg_i \in PGC\}, && (3) \\ CES(m_i) &= \{CE_1, CE_2, \ldots, CE_{\{NUM(CES(m_i))\}}\}, && (4) \\ OCR &= \{[m_i, CE_j]/m_i \in PMS, CE_j \in CES\}. && (5) \end{aligned}$$

Where definition (1) represents that project P consists of several sub-projects m_i which are executed by the chief enterprise (*CMS*) or partner enterprise (*PMS*) and *NUM(P)* is the number of sub-projects of AVE project P. Definition (2) represents that there are multiple relationships between m_i and m_j, such as the development period, cost restriction and control, where r_i is a certain relationship, S is the type of relationship. Definition (3) represents the goal set, where pg_i is one of goals, such as minimum time or minimum cost, etc., and PGC is the type of goal. Definition (4) is the candidate enterprise set of sub-project m_i, where CE_i is the candidate enterprise. Definition (5) is the enterprise optimization combination result. The number of enterprise combination is $\prod_{j=1}^{NUM(MS)}(CES(m_j))$, it increases exponentially with the increase of the number of sub-projects and enterprises.

3. PARTNERS SELECTION ALGORITHM BASED ON PSO

3.1 Design of *D*-dimensional flying space

In PSO, the particle flying space is the solution space [2]. Each location in the space is a potential solution to the problem. pBest is the optimal location of a particle, while gBest is the optimal location in all particles. Particle updates its speed and location continuously according to its pBest and gBest, and searches the optimal solution. To solve the problem of AVE partners selection, D-dimensional flying space is designed. In $NUM(P)$-dimension space, each dimension in the space is a sub-project m_i and each dimension space length could be different with length unit of 1 and each unit as a candidate enterprise CE_{ij}.

3.2 Variety of flying speed and location

The velocity of the particle affects the convergence of the algorithm and the ability of searching the optimal solution. In PSO algorithm, each particle is described by (x_{id}, v_{id}, p_{id}), where x_{id} is the current location, v_{id} is the current speed, p_{id} is the optimal location that the particle has experienced. Each particle has a fitness value, which is determined by the optimal fitness function $f(x_{id})$. The PSO method produces initial particle swarm randomly, and searches the optimal solution through iterations. The particles are updated by the following equations:

$$v_{id}^{n+1} = \omega v_{id}^{n} + c_1 r_1^n (p_{id}^n - x_{id}^n) + c_2 r_2^n (p_{gd}^n - x_{id}^n), \tag{6}$$

$$x_{id}^{n+1} = x_{id}^{n} + v_{id}^{n+1}, \tag{7}$$

$$p_{id} = f(x_{id}^{n+1}) \qquad if\ f(x_{id}^{n+1}) < f(x_{id}^{n}). \tag{8}$$

Where ω is the parameter of inertia weight, which describes the effect on current velocity by the former one, p_{id}^n is the particle's best location in the n iterations, p_{gd}^n is the best location that all particles have encountered in the n iterations, c_1 and c_2 are positive constants and r_1^n and r_2^n are generated randomly in (0, 1).

3.3 Objective function

The goal of the partners selection is to obtain the best combination of AVE. The combination of AVE is determined by the objective of the project and the relationship between sub-projects, such as minimum cost, minimum venture, etc. Therefore, the objective function is a multi-object function. For example, the cost function is:

$$\begin{aligned} Min(C) &= Min(C\_MS + C\_EMS + C\_Link) \\ &= Min\left(\sum_{i=1}^{NUM(MS)} C(m_i) + \sum_{j=1}^{NUM\,(EMS)} C(m_j)\right. \\ &\quad \left. + \sum_{i=1}^{NUM(MS)} \sum_{j=1}^{NUM(EMS)} CL(m_i, m_j, T, E)\right) \end{aligned} \tag{9}$$

Where *C_MS* is the cost of the chief enterprise manufacture, *C_EMS* is the cost of the candidate enterprise manufacture and *C_Link* is the immanent cost when the AVE is constituted, which includes the organization and management of AVE, the distribution of manpower resource and equipment resource. Considering that the algorithm may not make all objects be optimal, the weighted processing technique could be used.

3.4 Fitness function

The fitness function is determined by the *PGS* and the intensity of the relationship among the sub-projects. The definition of the fitness function is as follows:

$$f(x_1, x_2, \cdots, x_{NUM(P)}) = \sum_{pg_l \in PGS} \sum_{x_p = x_q} k_{pg_l} \times pg_l(I(x_p, x_q)) \Bigg/ \left(\sum_{x_p \neq x_q} I(x_p, x_q) + \sum_{s \in SMRS} \sum_{x_p \neq x_q} k_s G_s R_s(x_p, x_q) \right) \tag{10}$$

Where pg_l is one of the project objects, $k_{pg_l}$the weighted value of pg_l, $pg_l(I(x_p, x_q))$ the sum of the objective values when the object is pg_l and sub-project m_i selects enterprise x_p, sub-project m_j selects enterprise x_q, $I(x_p, x_q)$ the relative intensity values between m_i and m_j, k_s the weighted value of relationship type S, G_s the penalty weighted value of relationship type S which chooses a relative larger value according to $\sum I(x_p, x_q)$ and

$$R_s(x_p, x_q) = \begin{pmatrix} 0 & IF\ (x_p, x_q)\ satisfies\ SMR_s(m_i, m_j) \\ 1 & ELSE \end{pmatrix} \tag{11}$$

means that if enterprises x_p and x_q satisfy the relationship between sub-projects m_i and m_j, then it takes 0, else it takes 1.

4. EXPERIMENTAL RESULTS

The performance of the proposed PSO-based algorithm is examined by an AVE partners selection problem. The experimental results are compared with those obtained from the analytic hierarchy process and genetic algorithm. One part of the experiment data are obtained from Ref. [3], which is the basic requirement attribute of sub-projects. The experiment data among the sub-project relationship are shown in Table 1. The figures in parentheses represent the cost relationship requirement and the figures outside of parentheses represent the time relationship requirement. The time requirement represents that m_i is completed before $m_j n$ time unit, and the cost requirement represents that the sum of m_i and m_j cost must be within n cost unit. Project calculates minimum cost

Table 1. Sub-project relationship requirement.

Time\(cost)	s-p1	s-p2	s-p3	s-p4	s-p5	s-p6	s-p7	s-p8
s-p1	0	1\(25.2)	4.33	3.7	1.58	1.64	1	1
s-p2		0	1	5.67	(17.5)	7.37	4.76	
s-p3			0		6\(31.96)		10.13	(18.3)
s-p4				0				20.37
s-p5					0	15	(20)	
s-p6						0		
s-p7							0	
s-p8								0

object and minimum time object. The relative weighted value is 1:1 and the corresponding penalty weighted value is $G_c = 500$ and $G_t = 50$. The range of the velocity updating is $[-3,3]$, the size of the particle swarm is 30, and the number of permitted iterations is 1000.

The PSO-based algorithm has a steady fitness value when it iterates around 165 times. The results are satisfied. The best location in the global space is (1, 2, 17, 4, 11, 1, 5, 9). The partners combination is (0, 12, 33, 39, 47, 0, 61, 73) and the fitness value is 0.016773. The time for the development is 58 and the expended cost is 137. The GA has a steady fitness value when it iterates around 232 times. The results obtained from the proposed PSO-based algorithm are the same as those from the GA. When the AHP is used, the weighted estimates should be performed and the ratio of the cost and time is 1:1. A relative good partners combination is (0, 12, 27, 37, 47, 0, 59, 73). The development time is 60 and the cost is 153. Experimental results show that the AHP algorithm could transfer to a completely quantitative analysis when the stable parameter weight vector is determined. The estimates for the weighted relationship among parameters are determined totally by experts' experience or simple evaluation rules and formulae. The proposed PSO-based algorithm has the same functions as GA, however, the coding in the PSO-based algorithm is quite simple and the implementation is much easier. When the decision objects increase, the speed of convergence of the proposed algorithm is superior to that of the GA.

5. CONCLUSIONS

A PSO-based approach to solve the partners selection is proposed aiming at the partners selection when the AVE is constituted. The decision efficiency can be improved and the requirement of the speediness can be suited. Experimental results show the effectiveness and feasibility of the proposed PSO-based algorithm.

ACKNOWLEDGEMENTS

The authors are grateful to the support of the science-technology development project of Jilin Province of China under Grant no. 20030520 and the doctoral funds of the National Education Ministry of China under Grant no. 20030183060.

REFERENCES

1. R. Nagel and R. Dove R (1991), *21st Century Manufacturing Enterprise Strategy: An Industry-Led View*. Iacocca Institute, Lehigh University, USA.
2. R.C. Eberhart and J. Kennedy (1995), A New Optimizer Using Particle Swarm Theory In: *Proceedings of the 6th International Symposium on Micro Machine and Human Science*. Nagoya, Japan,. pp. 39–43.
3. L. Sha, X.F. Xu and Q.L. Li (2002), An optimization method for the standard of agile virtual enterprise performance base on genetic algorithm. *Computer Integration Manufacture System*, 6, pp.461–466.

ADAPTIVE DIRECTIONAL WEIGHTED MEDIAN FILTERING

Xiaoxin Guo, Zhiwen Xu, Yinan Lu and Yunjie Pang
College of Computer Science and Technology, Jilin University, Qianjin Road 10#,
Changchun 130012, P. R. China

Abstract A novel filtering algorithm called adaptive directional weighted median (ADWM) filtering is proposed in this paper. The ideas of the directional filtering and the weighted median filtering are combined to construct the ADWM filter. The use of the variance of the moving window and the base variance support the adaptivity of the ADWM filter. The experimental results show that the ADWM filter can both reduce random noise and preserve details.

Keywords: ADWM filtering, directional filtering, filter, median filtering.

1. INTRODUCTION

Median filtering is quite popular, employed in various applications of image smoothing, because median filters provide excellent noise-reduction capabilities, with considerably less blurring than linear smoothing filters of similar size. Especially, they are particularly effective in the presence of impulse noise. There are also such problems for median filtering as the damaging of thin lines and sharp corners in the image, the blurring fine details in the presence of Gaussian noise. Some improved and hybrid methods are developed. A combination of adaptive techniques [1–4] and order statistic filter [5] has been applied for noise suppression in images [6, 7].

In this paper, we propose a filtering algorithm called adaptive directional weighted median (ADWM) filtering by combining the ideas of the directional filtering and the weighted median filtering. The ADWM filtering can be accomplished by computing the variance of the moving window and the base variance, comparing between the two, and determining the corresponding directional sub-window and the weight of the centre pixel.

G. R. Liu et al. (eds.), Computational Methods, 1327–1331.

2. ADAPTIVE DIRECTIONAL WEIGHTED MEDIAN FILTER

In directional filtering [8], the directions of sub-windows vary depending on the change of directional edges. Directional filtering has the capability of preserving thin lines. Generally, possible sub-windows consist of five, four for directional filtering and one for nondirectional filtering. The sub-windows are defined as follows [8]

$$\begin{aligned}
W^0_{(n1,n2)} &= \{(n_1 + l_1, n_2 + l_2) \mid -N \le l_1, l_2 \le N\} \\
W^1_{(n1,n2)} &= \{(n_1 + l, n_2) \mid -N \le l \le N\} \\
W^2_{(n1,n2)} &= \{(n_1 + l, n_2 + l) \mid -N \le l \le N\} \\
W^3_{(n1,n2)} &= \{(n_1, n_2 + l) \mid -N \le l \le N\} \\
W^4_{(n1,n2)} &= \{(n_1 + l, n_2 - l) \mid -N \le l \le N\}
\end{aligned} \tag{1}$$

where N determines the length of the linear sub-windows. By using the sub-windows defined above, we select the most uniform sub-window out of the five at each pixel according to a criterion of uniformity [8], and perform filtering in the sub-window. Directional filtering along the edge avoids spatial blurring to a large extent. Therefore, it is very suitable for such applications as image smoothing for the images with obvious texture.

The ADWM filter is defined within the set of sub-windows $W^j_{(n1,n2)}$ for $j = 0, 1, 2, 3, 4$ as follows

$$\begin{aligned}
\hat{s}(n_1, n_2) = Med\Big(&\big\{w_{i1,i2,k} \diamond g(n_1, n_2)\big\} \cup \{g(i_1, i_2)| \\
&\times (i_1, i_2) \in W^k_{(n1,n2)}, (i_1, i_2) \neq (n_1, n_2)\}\Big)
\end{aligned} \tag{2}$$

$$k = \arg\min_j \left\{ uniformity\_criterion\left(W^j_{(n1,n2)}\right)\right\} \tag{3}$$

where Med is the median operator, performing median filtering; $k \diamond g(n_1,n_2)$ denotes the k-times repetition of $g(n_1,n_2)$; the weight of the pixel (n_1,n_2) is given by

$$w_{i1,i2,k} = [2L \cdot R(n_1, n_2)] + 1 \tag{4}$$

where $[\bullet]$ is a round operator, L is a convenient measure and satisfies that $2L + 1$ equals to the number of the pixels within the sub-window, and $R(n_1,n_2)$ is a normalized measure of the noise level in the window around pixel (i_1,i_2), given by

$$R(n_1, n_2) = 1 - V_n^2 \big/ V_x^2(n_1, n_2) \tag{5}$$

$V_n{}^2$ is a sort of base level for the variance in the whole area of the image. Just by resetting negative values to 0, $R(n_1,n_2)$ can be assured to satisfy $0\leq R(n_1,n_2)\leq 1$. $R(n_1,n_2)$ varies between 0 and 1, resulting in the weight of the centre pixel changing over the range from 1 to $2L$, which will directly influence the resulting value.

When one has no a priori value at hand, a method for getting an estimate of $V_n{}^2$ is to find the median for the entire image and then compute the RMS about the median. This strategy works well only if the image has been flat-fielded. In the usual case, however, where no flat field is perfect, we have to turn to the variance of the entire image for noise estimation.

When determining the criterion of uniformity, the variance of pixels within each sub-window may be considered as a selection criterion. If the variance of the pixels within the nondirectional window is the smallest, we decide that there is no edge crossing that pixel. If an edge is present at the pixel, the sub-window with the lowest variance indicates the most likely edge orientation at that pixel.

3. CAPABILITIES OF ADWM FILTERING

As denoted in Equation (5), $R(n_1,n_2)$ has relations with $V_n{}^2$ and $V_x{}^2$. In smooth, slowly varying parts of the image, if the window's variance V_x^2 is similar to the base variance $V_n{}^2$, the current window is dominated by random noise rather than signal, and $R(n_1,n_2)$ approximately equals to 0. In the extreme, when $R(n_1,n_2)=0$, the weight of the centre pixel is 1, and the ADWM filter can be reduced to the straight median filter, which is a stronger 'central indicator'. Since the median filtering is hardly affected by a small number of discrepant values among the pixels in the window, it is very effective at removing various kinds of noise. As a result, the output is the median of the window's values, and noise is strongly suppressed.

Although this strong smoother is very effective in removing salt and pepper and impulse noise because they do not depend on values that are significantly different from typical values in the window, its disadvantage is also very evident. The straight median filter destroys fine detail in the image as a result of the presence of image details relatively small in size compared to the size of the relative flat background. These details will have minimal affect on the value of the median filter, and so will be filtered out. In fact, median filtering is often used just for this purpose. This is an obvious drawback if one is interested in the fine detail.

Fortunately, by adaptively adjusting $R(n_1,n_2)$, we can achieve the goal of retaining image details and other signal of high frequency. When the window's variance $V_x{}^2$ is large compared to the base value $V_n{}^2$, the current window is in a strongly varying part of the image. Then the centre pixel's value is strongly weighted, more or less, depending on how far the window variance is above the base variance. When $R(n_1,n_2)$ approaches or equals to 1, the weight of the centre pixel is approximately $2L + 1$, which equals to the weight of the entire rest of the window. In this case, the value of the centre pixel is assured of being the output of the median operation. The ADWM filter turns out to be the identity filter, since the output equals to the input.

Therefore, the ADWM filter can be varied over the range from the median filter to the identity filter just by adaptively varying the central weight. This corresponds to going from strong noise and detail removal (much smoothing) to none. In other words, the definition and use of $R(n_1,n_2)$ enable the ADWM filter to achieve its adaptivity.

4. EXPERIMENTS AND RESULTS

In this experiment, we use median filter, directional filter and ADWM filter for comparison purposes. The noise considered in this paper is a combination of Gaussian noise (zero mean and 0.5 variance for intensity images whose intensity ranges from 0 to 1) and 'salt and pepper'(2%). Figure 1(a) is the noise-free image; (b) is the noisy one and (c–e) are corresponding outputs using different filters. The corresponding MSE is listed in Table 1. Table 1 shows the applicability and the robustness of the novel filter as compared to the other three filters.

Table 1. Results of comparison (5 × 5 window).

Output	Noisy image	Median filtering	Directional filtering	ADWM filtering
Normalized MSE	0.010381	0.0045971	0.0069207	0.004431
SNR	18.649	102.87	73.305	78.539
PSNR	68.953	385.52	324.35	306.52


(a) (b) (c) (d) (e)

Figure 1. (a) Original image, (b) noisy image, (c) output of median filter, (d) output of directional filter and (e) Output of ADWM filter.

5. CONCLUSIONS

In this paper, we propose the ADWM filter, which is an efficient and convenient nonlinear smoothing method that reduces the blurring of edges. The ADWM filter has the features of directional filtering and centre weighted median filtering. The employment of narrow sub-windows can achieve edge preservation efficiently and contribute to keeping a basic feature oriented from blurring. The novel filter can both remove noise and preserve detail via adaptively determining the weight of the centre pixel according to the lowest variance of pixels within all sub-windows.

6. ACKNOWLEDGEMENT

This work was supported by the foundation of science and technology development of Jilin Province, China under Grant 20040531.

REFERENCES

1. R. Bernstein (1987), Adaptive nonlinear filters for simultaneous removal of different kinds of noise in image. *IEEE Transactions on Circuits and Systems*, pp. 1275–1291.
2. Q. Cai, W. Yang, Z. Liu and D. Gu (1992), Recursion adaptive order statistics filters for signal restoration. In: *Proceedings of IEEE International Symposium on Circuits System,* San Diego, CA, pp. 129–132.
3. S. Haykin (1996), *Adaptive Filter Theory*, 3rd ed. Prentice-Hall, Upper Saddle River, NJ.
4. T. Peli and J.S. Lim (1982), Adaptive filtering for image enhancement. *Optical Engineering*, 21, 1, pp. 108–112.
5. A.C. Bovik, T.S. Huang, and D.C. Munson, (1983), A generalization of median filters using linear combination of order statistics, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-31,6.
6. C. Kotropoulos and I. Pitas (1992), Constrained adaptive LMS L-filters. *Signal Processing*, 26, 3, pp. 335–358.
7. C. Kotropoulos and I. Pitas (1996), Adaptive LMS L-filters for noise suppression in images. *IEEE Transactions on Image Process*, 5, 12, pp. 1596–1609.
8. L.S. Davis (1975), A survey of edge-detection techniques. *Computer Graphics and Image Processing*, 4, 248–270.

AN INTERACTIVE WORKFLOW MANAGEMENT FUNCTION MODEL

Jia-Fei Li, Da-You Liu and Wan-Jun Yu
[1]*College of Computer Science and Technology, JiLin University, Changchun 130021, China*
[2]*Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China*

Abstract In order to solve the problems of low description ability and interactivity in current workflow management systems, an interactive workflow management function model (IWMFM) is proposed. On the basis of using IWMFM to model the business process in the method of hierarchies, an algorithm of workflow process interaction based on the event-driven mechanism is put forward. Combining IWMFM and a formalized orgnization-role model, an algorithm of role analysis based on responsibility is given to implement the flexible role analysis during the interaction. Practical application shows that the proposed model has some advantages in enhancing description ability, reducing modelling complexity, strengthening interactivity between processes and improving flexibility in role analysis.

Keywords: workflow, function model, interactive, event-driven, role analysis.

1. INTRODUCTION

Nowadays, many enterprises want to extend the application of information technology into much more complicated and larger business processes. How to model these processes in an easy way become a key problem [1]. Based on the WfMC reference model [2] and the requirement of workflow management, we proposed an interactive workflow management function model (IWMFM). IWMFM can decompose the complex process definition into simple subprocesses by using a description method of hierarchies. Figure 1 shows a disassembled process definition. Through event-driven mechanism, the decomposed processes can interact with each other. During the interaction, we need to analyse the tasks participants in the decomposed processes from role to

G. R. Liu et al. (eds.), Computational Methods, 1333–1338.

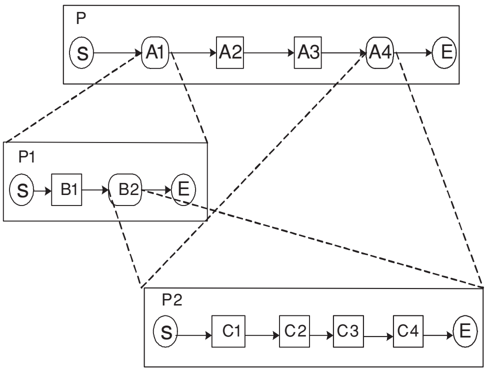


Figure 1. A decomposed process definition.

user or team according to the information included in the event. In this way, the same business logic as the former complicated process is completed and the modelling process is simplified. IWMFM combines with an orgnization-role model to implement the flexible role analysis during the interaction.

2. PROCESS MANAGEMENT

Process management mainly includes workflow modelling, workflow engine, event manager and organization manager and worklist handler.

2.1 Process modelling

The model element of process definition is defined as follows:

Definition 1. procDef = [procName, procAttr], where

(1) procName: is the name of the process template, event manager maintain it in the type linkage item.
(2) procAttr: is the process's attributes, e.g., process ID, description, global variables, router type, transition conditions, creator, deadline, etc.
Lets δ: $N \to D$ is the mapping from N to D, where
(3) procName set N; procDef set D;
(4) $d = \delta(n)$ is a process definition that has the name of n. The number of process definitions that have the name of n is limited to one and each process definition can own one unique name, hence δ is of one to one relationship.

2.2 Event manager

Event is the occurrence of a special status or condition which is significative to some workflow processes, e.g., changement of a process state, complement of an activity, generation of a business object, etc. The event is defined as follows:

Definition 2. evt = [evtName, evtParas, procName], where

(1) evtName: is the name of the event. It is unique in the system.
(2) evtParas: is the information that the event receiver need to handle, e.g., the information for role analysis, business object, etc.
(3) procName: is the name of the process definition that generates the event.

Event manger is responsible for dispatching the event to the event receiver (a process instance) when the event is generated and the maintenance of the type linkage item.

2.3 Workflow engine

The workflow engine is the core of the workflow management system, interprets the process definition, controls the creating, activating, suspending and terminating of process instance, takes charge of the states transition of process and activities, navigates the activities running, records history events and manages transactions.

3. WORKFLOW INTERACTION

After modelling a complex business process by the method of hierarchies, the decomposed subprocesses in different level need to interact with each other to complete the same business logic as the former complicated process. In IWMFM, the interaction is achieved by event manager. Through the mechanism of publish-subscribe [3], event manager knows the generation of the events, searches for the corresponding receiver. Because the event is published by a process instance and the receiver is a process definition, the workflow process interaction is implemented. The relationship between the event and the receiver is maintained in a type linkage item. The type linkage item is defined as follows:

Definition 3. typeLinkage = [evtName, procName, status], where

(1) evtName: is the name of a event.
(2) procName: represents a receiver and is the name of the process definition that receive the event.

```
Algorithm WPI(evt)
/*the workflow process interaction algrithm*/
 WPI1 [Initialization]
     FOR i =1 TO n
         tpLks[i] ← Λ.
  WPI2 [Workflow process interaction]
     FOR j =1 TO n
       (IF(ζ(evt)[j] ∈ TL AND ζ(evt)[j].status = 1) THEN
             (tpLks[j] ← ζ(evt)[j].
                 procName1 ← tpLk[j].procName.
                 procDef1 ← δ(procName1).
                 IF (procDef1 ≠ Λ ) THEN
                (//RBRA is the reponsibility-based role
                 // analysis algorithm and is defined later.
                  roleMap←RBRA(evt.evtParas.item,
                              evt.evtParas.value,orgRole).
                 //start a workflow instance
                 StartWf(procDef1,roleMap)
                )
        ) ■
```

Figure 2. Workflow process interaction algorithm.

(3) status: identifies whether the linkage item is active or not.

Lets $\zeta: E \rightarrow TL$ is the mapping from E to TL, E is a evt set, TL is a typeLinkage set. Obviously, ζ is of one to many relationship. We give the workflow process interaction algorithm in the language of ADL [4] in Figure 2.

4. ORGANIZATION-ROLE MANAGEMENT

Orgnization-role management in IWMFM mainly implements the flexible role analysis during the interaction. RBAC96 is an accepted and important access control model [5], but RBAC96 does not define the authorization about data object and user team which can simplify the mapping from role to user. We extend the RBAC_0 of RBAC96 to xRBAC and use the xRBAC as the organization-role management model and give a role analysis algorithm based on responsibility.

Definition 4. organization-role = $<U, R, T, P, O, \theta, \alpha, \beta, \gamma, \lambda, \Sigma>$, where

(1) User set U; Role set R; Team set T; Permission set P; Object set O.
(2) θ: $U \rightarrow R$, is the mapping from U to R. $r = \theta(u)$ is a role that u plays. The number of roles that one user can play is limited to one, but one role can be played by different users, hence θ is of many to one relationship.
(3) α: $T \rightarrow R$, is the mapping from T to R. $r = \alpha(t)$ is a role that t plays, i.e., every member in t plays role r. Similar to θ, α is also a many to one relationship.

(4) β: $T \rightarrow U$, is the mapping from T to U. $U' = \beta(t)$ is the user set in which every user is the member of t.

(5) γ: $U \rightarrow T$, is the mapping from U to T. $T' = \gamma(u)$ is the team set in which u belongs to every team.

(6) $\lambda \subseteq R \times O \times P$. $(r, o, p) \in \lambda$, means that role r is granted access on object o with privilege p. The connection between permission and user is made by role. According to Definition 4, u has the privileges that the privileges of his roles plus the privileges of his teams. The privilege set of u is $\{p|r = \theta(u) \wedge (r, o, p) \in \lambda\} \cup \{p|T' = \gamma(u) \wedge t \in T' \wedge r = \alpha(t) \wedge (r, o, p) \in \lambda\}$.

(7) $\Sigma : R \rightarrow 2^A$. $A = U \cup T \cup R$. If $\Sigma(r) \subseteq U \cup T$, then we get the participants of workflow activity directly. If $\Sigma(r) \subseteq R$, then we analyse the participants by orgnization-role object.

Responsibility is the duty of an agent which refers to a user, a role or a team, e.g., if the sum of an order is over \$5000, the manager has the responsibility to deal with it. Otherwise, the employer can handle the order. The judgment condition of responsibility is defined as follows:

Definition 5. condition = [item, operator, scale], where

(1) item: identifies the content to judge, e.g., the sum of the order, the leave time, etc.

(2) operator: is the operator to use in judgment, e.g., >, <, =, etc.

(3) scale: is the standard value to compare with the item's actual value.
Lets ψ: $I \rightarrow C$ is the mapping from item to condition, I is an item set, C is a condition set. ψ is a one to many relationship.

Definition 6. responsibility: Π: $C \rightarrow A$ is the mapping from C to A, condition set C, $A = U \cup T \cup R$. (There is a hypothesis that limit one agent can associate with one condition only and the standards of different conditions that have the same item are identical.) We give the responsibility-based role analysis algorithm in Figure 3.

5. CONCLUSION

IWMFM can softly model a complex business process in a structure of hierarchy and the different process instance can conveniently interact with each other by using the workflow interaction algorithm. Moreover, the responsibility-based role analysis algorithm can be used to implement the flexible role analysis during the interaction. IWMFM was applied in JQ/cERP and was proved to satisfy the workflow management requirements in the manufacturing enterprises.

```
Algorithm RBRA(item,value,orgRole.roleMap)
 RBRA1 [Initialization]
     FOR i = 1 TO n DO
         roleMap[i] $\leftarrow \Lambda$ .
 RBRA2 [Get the condition to use]
     $tc \leftarrow \psi$(item)[0].
     $st \leftarrow$ tc.scale.
     //compare the value with the scale
     $op \leftarrow$ Compare(value,scale).
     $c \Leftarrow$ <item,op,st>.
 RBRA3 [Role analyzing based on responsibility]
     IF ($\Pi(c) \in U \cup T$) THEN
         roleMap $\leftarrow \Pi(c)$. //role analysis
             ELSE IF ($\Pi(c) = r'' \in R \wedge \Sigma(r'') = r' \in R$) THEN
               roleMap $\leftarrow \{u | u \in U \wedge \theta(u) = r'\} \cup \{t | t \in T \wedge \alpha(t) = r'\}$.
              ELSE  roleMap[j].$= \Sigma(r'')$ ▌
```

Figure 3. Responsibility-based role analysis algorithm.

ACKNOWLEDGEMENTS

This work is supported by the National Natural Science Foundation of China under Grant No. 60173006; the National High-Tech Research and Development Plan of China under Grant No. 2003AA118020; NSFC Major Research Program 60496321, Basic Theory and Core Techniques of Non Canonical Knowledge; and the Major Program of Science and Technology Development Plan of Jilin Province under Grant No. 20020303.

REFERENCES

1. W. van der Aalst and K. van Hee (2002), *Workflow Management Models, Methods, and Systems.* MIT Press, Cambridge, Massachusetts.
2. WfMC TC00-1003 Issue (January 1995), *Workflow Management Reference Model*. Workflow Management Coalition, Crown Walk, Winchester, Hampshire, UK.
3. E. Gamma, R. Helm, R. Johnson and John Vlissides (1995), *Design Patterns: Elements of Reuseable Object-Oriented Software.* Addison Wesley Longman, Inc.
4. D-Y. Liu and Y-F. Li, *Priciples of Data Structure.* Jinlin University Press, Changchun, Jilin, China.
5. R. Sandhu, E. Coyne, H. Feinstein, et al (1996), Role-based access control models. *IEEE Computers*, 29, 2, pp. 38–47.

PACKET FAIR SCHEDULING ALGORITHM BASED ON WEIGHTS DYNAMIC COMPENSATION

Zhang Wei, Liu Yan-heng, Yu Xue-gang and Xie A-lian
Key Laboratory for Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, China

Abstract PFQ is approximation of GPS in packet switching networks and it can provide packet flows with QoS and fairness. Based on research of fairness of GPS and details of WFQ and WF^{2}Q, this paper implements CFQ algorithm with the idea of maintaining the proportional fairness as GPS. CFQ introduces the concept of deflection ratio and uses it in computing of timestamp. During scheduling, system makes micro-adjustment to weight according to flow's deflection ratio. Analysis and NS2 simulation show that CFQ reduces the fairness difference among flows with different QoS requirement obviously and provide bounded QoS performance similar to GPS.

Keywords: IntServ, packet fair queuing, weighted fairness, compensation.

1. INTRODUCTION

GPS [1] based on bit stream has ideal fairness. In packet networks, current PFQ algorithms [2, 3] can provide performance similar to GPS by simulating the service curve of GPS. WFQ [1] and WF^{2}Q [4] are the most excellent PFQ algorithms. Based on fairness analysis of GPS, the performance evaluation criteria [2] and wireless compensation models [5], this paper introduces CFQ (Compensation-based weighted Fair Queuing) algorithm with the idea of maintaining the proportional fairness of service as GPS.

G. R. Liu et al. (eds.), Computational Methods, 1339–1344.

Table 1. Deflection ratio statistic of WFQ/WF^{2}Q during time slots.

Time slots	(0, 1)	(0, 2)	(0, 3)	(0, 4)	(0, 5)	(0, 6)	(0, 7)	(0, 8)	(0, 9)	(0, 10)
WFQ Δ_1	13.99	24.24	32.09	38.28	43.29	47.43	50.91	53.87	56.43	58.65
WFQ Δ_2	−1.4	−2.43	−3.21	−3.83	−4.33	−4.74	−5.09	−5.39	−5.64	−5.86
WF^{2}Q Δ_1	18.18	0	5.19	0	3.19	0	2.30	0	1.8	0
WF^{2}Q Δ_2	−1.4	57.58	50.91	29.09	27.75	18.19	17.49	12.12	11.70	8.26

2. DEFINITION OF DEFLECTION RATIO

When all flows are continually backlogged, flow's weight and received service of GPS have relation as follows:

$$\frac{W_i(\tau, t)}{W_j(\tau, t)} = \frac{\Phi_i}{\Phi_j} \quad i, j \in B(t) \tag{1}$$

where W symbolizes service received by flow during time slots, Φ symbolizes scheduled weight value and B symbolizes set of busy sessions. This formula shows that received services of flows are strictly proportional to their weights. It's proportional to fairness [2]. Packets are not dividable that will destroy this fairness and result in deflection during some time slots. A ladder variational deflection ratio Δ is defined as follows to show the extent of fairness deflection:

$$\Delta_i\left(t_j\right) = \Delta\frac{W_i}{\phi_i} = \frac{\dfrac{W_i}{\phi_i} - \dfrac{1}{N_B}\displaystyle\sum_{k=1}^{N_B}\frac{W_k}{\phi_k}}{\displaystyle\sum_{k=1}^{N_B}\frac{W_k}{\phi_k}} \times 100\% \tag{2}$$

where N_B symbolizes the number of busy flows and j symbolizes the number of events occurring during time slots from time zero to now. According to scheduling order of WFQ and WF^{2}Q in Ref. [4], deflection ratio statistics is listed in Table 1. It's obvious that there exists a large difference among flows with different weight.

3. DYNAMIC WEIGHT COMPENSATION UNDER RESTRICTION OF DEFLECTION RATIO

GPS's ideal fairness was based on the attribute that stream is infinitely dividable, and we consider that weight value is also infinitely dividable. During packet scheduling, the fairness deflection can be fed back to weight, and then make a micro-adjustment to weight according to their deflection ratio. Because weight participates in computing of virtual time stamp, so it can be assured that flows with fairness damaged during previous time slots can obtain advantageous deflection in choice of next packet. It shows as following, formula (3):

$$\Phi_i\left(t_j\right) = \Phi_i \times \left(1 - \Delta_i\left(t_j\right)\right) \tag{3}$$

where Δ is a decimal between 0 and 1. This scheme names CFQ. In this scheme, the virtual time V and timestamp F_i^k are computed as following, formula (4) and (5). Where, the δ called compensation index, which has the form as (6):

$$V(0) = 0, \quad V\left(t_{j-1} + \mu\right) = V\left(t_{j-1}\right) + \frac{\mu \cdot r}{\sum_{i \in B_j} \Phi_i\left(t_{j-1}\right)} \mu \le t_j - t_{j-1}, \quad j = 2, 3, \ldots \tag{4}$$

$$S_{i,CFQ}^k = \max\left\{F_{i,CFQ}^{k-1}, V\left(A_i^k\right)\right\} F_{i,CFQ}^k = \delta_j \cdot F_{i,CFQ}^{k-1} + \frac{L_i^k}{\phi_i\left(t_j\right)} \tag{5}$$

$$\delta_j = \frac{\phi_i\left(t_{j-1}\right)}{\phi_i\left(t_j\right)} \quad \text{if} \quad S_{i,CFQ}^k = F_{i,CFQ}^{k-1} \quad \text{else} \quad \delta_j = 1 \tag{6}$$

The scheduling process of CFQ is described as follows:

S1. Initialize weight registers and service register. Set flow weights with defined parameters corresponding to classes.
S2. Check the state of busy session aggregation B. If B changes, go to S7, or continue.
S3. If a packet A_i^k arrives, record the values of i, k and packet length L.
S4. If a packet dequeues:
  a) Add Ltm/$L_{\max}$ to service register W_i.
  b) Update weight register ϕ_i.
  c) Update timestamp F_i^k of HOL packets.
S5. Round robin and choose HOL packet with least timestamp.
S6. Go to S2.
S7. End and exit.

4. NS2 SIMULATION AND PERFORMANCE ANALYSIS

According to packet arrival in Ref. [4], the CFQ scheduling order is shown in Figure 1. It's obvious that the dequeuing order just accords with the original intention, namely, compensation mechanism works.

Figure 2 is the topology of NS2 simulation. It's a simple scene with only three flows.

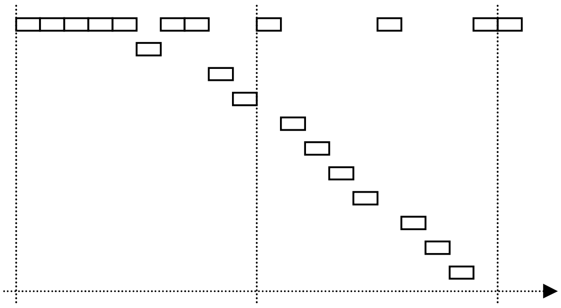


Figure 1. CFQ scheduling order.

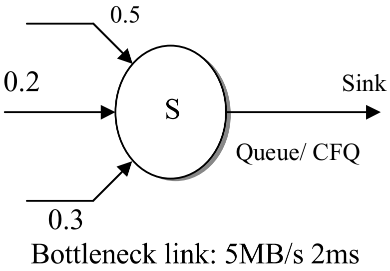


Figure 2. NS2 network topology.

4.1 Analysis of deflection ratio

The compare of deflection ratio is shown in Figure 3. WFQ and WF^{2}Q have obvious deflection to certain flows and CFQ has a notable action of compensation. According to long-term fairness definition [2], generalized by formula (7), the CFQ has better fairness.

$$F_{\delta,i} = \frac{1}{\Delta t}\int_{t}^{t+\Delta t} f_{\delta,i}\left(\delta_{i,1}, \delta_{i,2}, \ldots, \delta_{i,n}; t\right) \tag{7}$$

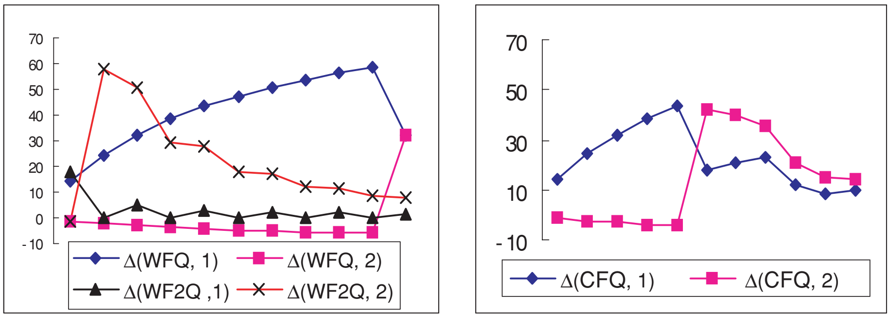


Figure 3. Deflection ratio analysis.

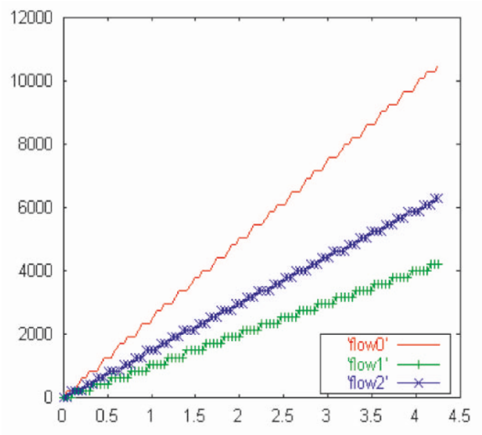


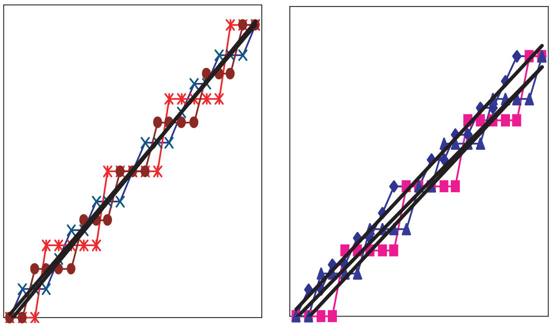
Figure 4. Long-term service curve.

4.2 Service curve

When weights are 0.5, 0.2 and 0.3, the service summation illustrates in Figure 4. CFQ can simulate the proportional fairness of GPS accurately during long-term.

4.3 Curve of fairness maintaining

According to service curve, the fairness curve of CFQ, namely formula (1), changes with time slots is plotted in Figure 5. The Figure 5(a) is in CFQ and Figure 5(b) is in WFQ. It’s evident that CFQ obviously reduces the fairness difference among flows with different class and QoS requirement.



(a) (b)

Figure 5. Fairness compare.

5. CONCLUSION AND FUTURE WORK

This paper implements CFQ algorithm based on the idea of maintaining the proportional fairness of service of GPS, introduces the concept of deflection ratio and uses it in computing of timestamp as an effect factor. NS2 simulation result shows that CFQ can obviously reduce fairness difference among flows with different QoS requirement. Wireless QoS scheduling is a hot topic in recent years. Adopting CFQ as reference algorithm and combining it with compensation model are the future work.

REFERENCES

1. A.K. Parekh (1992), A generalized processor sharing approach to flow control in integrated services networks: the single node case. In: *IEEE INFOCOMM'92*. IEEE Computer Society Press, pp. 915–924.
2. L. Chuang, S. Zhiguang and R. Fengyuan (2004), *Quality of Service of Computer Networks*. Tsinghua University Publishing Company, Beijing, China, pp. 176–177.
3. M. Zheng and C. Jian (2003), A study on the performance of PFQ algorithms. *Acta Electronica Sinica*, 31, 10, pp. 1555–1557.
4. J.C.R. Bennett and H. Zhang (1996), WF^2Q: worst-case fair weighted fair queuing. In: *Proceedings of IEEE INFOCOM'96*, San Francisco, CA, pp. 120–128.
5. S. Bucheli (2000), Compensation modeling for QoS support on a wireless network. M.S. Theses, Ecole Polytechnique de Lausanne/University of Illinois, pp. 14–20.

EXTENDED ADAPTIVE WEIGHTED AVERAGING FILTER MODEL

Xiaoxin Guo, Zhiwen Xu, Yinan Lu, Yunjie Pang
College of Computer Science and Technology, Jilin University, Qianjin Road 10#,
Changchun 130012, P.R.China.

Abstract In this paper, a relatively flexible filter model called extended adaptive weighted averaging (EAWA) filter model is proposed, by which some particular filters can be designed via selecting an appropriate pixel of interest (POI) and defining a weight distribution function for the filter according to some specified requirements.

Keywords: AWA filter, filter, image sequence, pixel of interest.

1. INTRODUCTION

The classical AWA filter [1, 2] is a motion compensated filter, which computes a weighted average of the image values within the spatio-temporal support along the motion trajectory. The filter is therefore particularly well suited for efficient filtering of sequences containing segments with varying scene contents due to rapid zooming and changes in the view of the camera.

In this paper, we propose the extended adaptive weighted averaging (EAWA) filter model based on the former AWA filter used for noise filtering. The reason that the EAWA filter is called a *model* is that the EAWA filter can turn to be a practical filter only when some quantities (functions, parameters and domain) are defined definitely; otherwise, the filter is indefinite. The EAWA filter model supplies some possibilities to design filters. The following factors strengthened the flexibility and the generality of the model: (1) the weight distribution function to control the distribution of the weights; (2) the pixel of interest (POI) to position the center of weight distribution; (3) the boundary threshold T to distinguish the desirable pixels from the undesirable ones; (4) the plateau threshold ε to influence the ability of noise reduction and the adaptivity of the weight

G. R. Liu et al. (eds.), Computational Methods, 1345–1349.

distribution. Therefore, the flexible model is able to facilitate transforming itself to some practical filters.

2. MODELLING

Based on the method of AWA filtering, the EAWA filter is defined within the support $S_{n1,n2,k}$ as follows

$$\hat{s}(n_1, n_2, k) = \sum_{(i_1,i_2,l)\in S_{n_1,n_2,k}} w(i_1, i_2, l)g(i_1, i_2, l) \tag{1}$$

where $g(i_1, i_2, l)$ is the noisy pixel within the support $S_{n1,n2,k}$, $\hat{s}(n_1, n_2, k)$ denotes the estimate of the ideal image $s(n_1, n_2, k)$ applying this model, and

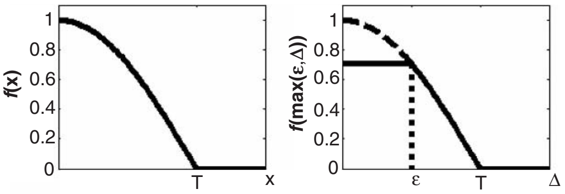
$$w(i_1, i_2, l) = h_{n_1,n_2,k}(i_1, i_2, l) \cdot \left(\sum_{(i_1,i_2,l)\in S_{n_1,n_2,k}} h_{n_1,n_2,k}(i_1, i_2, l) \right)^{-1} \tag{2}$$

are the weights within the support $S_{n1,n2,k}$, and

$$h_{n_1,n_2,k}(i_1, i_2, l) = f(\max\{\varepsilon, |C(n_1, n_2, k) - g(i_1, i_2, l)|\}) \tag{3}$$

The quantity $\varepsilon(\varepsilon \geq 0)$ is the parameters of the filter. $C(n_1, n_2, k)$ denotes the value of the POI within the support $S_{n1,n2,k}$. Which pixel is defined as the POI within the support $S_{n1,n2,k}$ can reflect what kind of values of signal (or pixels) is desirable to be preserved after filtering noise. $f(x)$ is nonnegative within its domain ($x \geq 0$) and nonincreasing as x increases (see Figure 1), which determines the distribution of the weights, called the weight distribution function. Now, the above definition is noted with such properties as follows:

1. Take $C(n_1, n_2, k) = g(n_1, n_2, k)$ and $f(x) = 1/(1 + ax^2)$, where a ($a > 0$) is a penalty parameters [1], the AWA filter can be formed. Therefore, the AWA filter is just one case in this model, and the EAWA filter is an extension to the AWA filer.



(a) Graph of *f*(*x*) **(b) Graph of *f*(max(e,Δ))**

Figure 1. Illustration of the weight distribution function $f(x)$ (T is a boundary threshold, ε is a plateau threshold, $\Delta = |C(n_1,n_2,k) - g(i_1,i_2,l)|$).

2. The EAWA filter enhances the ability to control the weights of pixels within the support $S_{n1,n2,k}$. In some cases, the weight of zero is desirable because it can eliminate the interference of undesirable pixels. Unfortunately, the weights within the support $S_{n1,n2,k}$ of the AWA filter is always positive. By defining an appropriate weight distribution function $f(x)$ for an EAWA filter, it is possible to settle the problem. Distributing positive weights to the desirable pixels, and zero weights to the undesirable ones is one of the simplest ways. Hence, the EAWA filter is considered to be more flexible in selecting the desirable pixels and rejecting the undesirable. In addition, the curve of the weight distribution of the AWA filter is fixed, whereas the EAWA filter escapes the confine of the fixed weight distribution function, and allows us to define the weight distribution function as we require. Therefore, in different applications the weights can be adjusted by designing and adopting different weight distribution function $f(x)$. The flexibility of defining $f(x)$ is not confined to the existing of enough matching pixels within the support $S_{n1,n2,k}$. Here, an example for $f(x)$ is given in Figure 1a, where T is a constant, and serves as a boundary threshold, by which the desirable pixels can be distinguished from the undesirable ones. $f(x)$ with the threshold T divides $S_{n1,n2,k}$ into two parts, the subsupport $U_{n1,n2,k}$, given by

$$U_{n1,n2,k} = \{(i_1, i_2, l) | (i_1, i_2, l) \in S_{n1,n2,k}, |g(i_1, i_2, l) - C(n_1, n_2, k)| \leq T\} \tag{4}$$

   and its complement within the support $S_{n1,n2,k}$. If $f(x)$ is not a function with a compact support, the boundary threshold T is considered to be infinity.
3. The same weight distribution function does not necessarily result in the same filter. The POI is also influential on the model. For the AWA filter, high weights are fixedly distributed around the centre pixel. In contrast, by allowing to define POI $C(n_1,n_2,k)$, the EAWA filter model makes it more flexible to distribute the weights around its POI, which may be an arbitrary value, probably even not any value of existing pixels. When we change the POI, actually we also alter the centre of the weight distribution. Besides, the procedure of selecting the POI can be considered as a pre-filtering stage. So we say that the identical weight distribution function does not mean the identical filter, and that the filters vary depending on their POIs and their subsupport $U_{n1,n2,k}$.
4. By appropriately selecting the parameter ε, for instance, setting ε equal to the value of the standard deviation of noise, ε can be used as a criterion of the consistent intensity of POI to measure whether a region is stationary or nonstationary. According to the criterion, the EAWA filter model distribute high weights among the pixels whose intensities are consistent with POI

(stationary image regions), and to distribute low ones among the pixels beyond the consistent intensity of POI (nonstationary image regions).

At the same time, the parameter ε as a plateau threshold (see Figure 1b) plays an important part in the relationship of the efficiency of noise suppression and the adaptivity of the weight distribution. ε controls the ability of the two aspects. When ε approaches the boundary threshold T, the ability of noise reduction increases, whereas the adaptivity decreases. When ε approaches zero, the situation is just the reverse. Note that Ref. (3) can be written as follows:

$$h_{n_1,n_2,k}(i_1,i_2,l) = \begin{cases} f(\varepsilon) & \textit{if } |C(n_1,n_2,k) - g(i_1,i_2,l)| < \varepsilon \\ f(|C(n_1,n_2,k) - g(i_1,i_2,l)|) & \textit{otherwise} \end{cases} \tag{5}$$

It can be shown that if the difference between the values $C(n_1, n_2, k)$ and $g(i_1, i_2, l)$ is less than ε, all the weights attain the same value $f(\varepsilon)$. However, if the difference is greater than ε, the contribution of $g(i_1, i_2, l)$ is weighted down

Table 1. Application of the EAWA filter model.

Gradient inverse weighted smoothing scheme [3]

$C(n_1, n_2) = g(n_1, n_2)$

$$f(x) = \begin{cases} 1 & \textit{if } 0 \le x < 1 \\ 1/(Vx) & \textit{if } 1 \le x \le T \\ 0 & \textit{otherwise} \end{cases}$$

where

$$V = \sum_{(i_1,i_2)\in V_{n1,n2}} \frac{1}{|C(n_1,n_2) - g(i_1,i_2)|}$$

$$V_{n1,n2} = \{(i_1,i_2)|(i_1,i_2) \in S_{n1,n2}, (i_1,i_2) \neq (n_1,n_2)\}$$

and the support $S_{n1,n2}$ is a 3×3 square region.

Edge-preserving filter [5]

$C(n_1, n_2) = g(n_1, n_2)$

$$f(x) = \begin{cases} 1 \ \textit{if } 0 \le x \le \sigma_{\min} \\ 0 \quad \textit{otherwise} \end{cases}$$

where $\sigma_{\min} = \min_{j=0}^{5}\{\sigma(U^j_{n1,n2})\}$, $U^0_{n1,n2}$ is an n × n square region, and

$$U^1_{n1,n2} = \{(i_1,i_2)|(i_1,i_2) \in U^0_{n1,n2}, (i_1 - n_1) - (i_2 - n_2) = 0\}$$

$$U^2_{n1,n2} = \{(i_1,i_2)|(i_1,i_2) \in U^0_{n1,n2}, (i_1 - n_1) + (i_2 - n_2) = 0\}$$

$$U^3_{n1,n2} = \{(i_1,i_2)|(i_1,i_2) \in U^0_{n1,n2}, i_1 - n_1 = 0\}$$

$$U^4_{n1,n2} = \{(i_1,i_2)|(i_1,i_2) \in U^0_{n1,n2}, i_2 - n_2 = 0\}$$

k-Nearest neighbour averaging filter [4]

$C(n_1, n_2) = g(n_1, n_2)$

$$f(x) = \begin{cases} 1 & \textit{if } 0 \le x \le \sigma_{\min} \\ 0 & \textit{otherwise} \end{cases}$$

where $\sigma_{\min}$ denotes the minimum standard deviation with respect to the neighbourhoods E^k. And E^k refers to the k neighbourhood in an $n \times n$ square region around a point P, whose standard deviation is considered as an index of the uniformity in the local region.

Order-statistic filter

$$C_{\max}(n_1,n_2,k) = \max_{(i_1,i_2,l)\in S_{n_1,n_2,k}} \{g(i_1,i_2,l)\}$$

$$C_{\mathrm{med}}(n_1,n_2,k) = \operatorname*{med}_{(i_1,i_2,l)\in S_{n_1,n_2,k}} \{g(i_1,i_2,l)\}$$

$$C_{\min}(n_1,n_2,k) = \min_{(i_1,i_2,l)\in S_{n_1,n_2,k}} \{g(i_1,i_2,l)\}$$

$$f(x) = \begin{cases} 1 & \textit{if } 0 \le x \le T \\ 0 & \textit{otherwise} \end{cases}$$

When T approaches zero, we can obtain a maximum, a median and a minimum filter corresponding to three POIs, respectively.

by $f(|C(n_1,n_2,k)-g(i_1,i_2,l)|) < f(\varepsilon)$ in that $f(x)$ is a nonincreasing function. Therefore, the parameter ε should be selected carefully through experiments.

3. APPLICATION

With the model, we can design some particular filters via defining an appropriate POI $C(n_1,n_2,k)$ and an appropriate weight distribution function $f(x)$. We give some examples to demonstrate its generality and flexibility (See Table 1). These examples describe some classic noise filters in the form of the EAWA filter model. We note that designing and defining $f(x)$ and $C(n_1, n_2, k)$ for the EAWA filter will directly determine its performance for a specific filtering.

4. CONCLUSION

A filter model called the EAWA filter model derived from the AWA filter is introduced in this paper. The EAWA filter model is a relatively flexible and general filter model, by which some particular filters can be designed via selecting an appropriate POI and defining a weight distribution function for the filter according to some specified requirements to perform a specific filtering.

ACKNOWLEDGEMENT

This work was supported by the foundation of science and technology development of Jilin Province, China under Grant 20040531.

REFERENCES

1. M.K. Ozkan, M.I. Sezan and A.M. Tekalp (August 1993), Adaptive motion-compensated filtering of noisy image sequences. *IEEE Transactions on Circuits and Systems for Video Technology*, 3, pp. 277–290.
2. A.M. Tekalp (1998), *Digital Video Processing*. PTR: Prentice Hall, pp. 272–276.
3. D.C.C. Wang, A.H. Vagnucci and C.C. Li (1981), Gradient inverse weighted smoothing scheme and evaluation of its performances. *Computer Graphics and Image Processing*, 15, pp.167–181.
4. L.S. Davis and A. Rosenfeld (1978), Noise cleaning by iterated local averaging. *IEEE Transactions on Systems, Man and Cybernetics*, 8, pp. 705–710.
5. L.S. Davis (1975), A survey of edge-detection techniques. *Computer Graphics and Image Processing*, 4, pp. 248–270.

AN IMPROVED QUANTUM-INSPIRED EVOLUTIONARY ALGORITHM FOR CLUSTERING GENE EXPRESSION DATA

W.G. Zhou, C.G. Zhou, G.X. Liu, H.Y. Lv and Y.C. Liang
College of Computer Science and Technology, Jilin University, Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, Changchun 130012, P. R. China

Abstract Microarray technologies have made it straightforward to monitor simultaneously the expression pattern of thousands of genes. So an important task is to cluster gene expression data to identify groups of genes with similar patterns and hence similar functions. In this paper, an improved quantum-inspired evolutionary algorithm (IQEA) is first proposed for minimum sum-of-squares clustering. We have suggested a new representation form and added an additional mutation operation in IQEA. Experiment results show that IQEA appears to be much more robust in finding optimum or best-known solutions and be superior to conventional k-means and self-organizing maps clustering algorithms even with a small population.

Keywords: gene expression, clustering, quantum-inspired evolutionary algorithm.

1. INTRODUCTION

In the field of bioinformatics, clustering algorithms have received renewed attention due to the breakthrough of microarrays data. Microarrays experiments allow for the simultaneous monitoring of the expression patterns of thousands of genes. Since a huge amount of data is produced during microarray experiments, clustering methods are essential in the analysis of gene expression data. The goal is to extract the fundamental patterns inherent in the data and to partition the elements into subsets referred to as clusters. In gene expression, elements are usually genes. The vector of each gene contains its expression levels under each of the monitored conditions. Several clustering algorithms have been proposed for gene expression data analysis, such as hierarchical

G. R. Liu et al. (eds.), Computational Methods, 1351–1356.

clustering, self-organizing maps, k-means, and some graph theoretic approaches. In this paper, we propose an improved quantum-inspired evolutionary algorithm (IQEA) for clustering gene expression data. The IQEA has been shown to perform well and be superior to k-means and self-organizing maps clustering algorithms even with a small population.

2. MINIMUM SUM-OF-SQUARES CLUSTERING

Clustering can be considered as a combinatorial optimization problem, in which an assignment of data vectors to clusters is desired, such that the sum of distance square of the vectors to their cluster mean (centroid) is minimal. Let P_k denote the set of all partitions of X with X $= \{x_1, \ldots, x_n\}$ denoting the data set of vectors with $x_i \in R^m$ and C_i denoting the ith cluster with mean x_i. Then we can formulate the problem as the search for an assignment p of the vectors to the clusters with $C_i = j \in \{1, \ldots, n\} | p[j] = i$. Thus the objective function becomes:

$$\min_p \sum d^2(x_i, \hat{x}_{p[i]}). \tag{1}$$

This combinatorial optimization problem is called the minimum sum-of-squares clustering (MSSC) problem as shown in the work by Merz [1] and has been proven to be NP-hard. Since the novel quantum-inspired evolutionary algorithm has been shown to be effective as compared to the conventional genetic algorithm, the application of QEA to the MSSC appears to be promising.

3. QUANTUM-INSPIRED EVOLUTIONARY ALGORITHM

Quantum-inspired evolutionary algorithm (QEA) is based on the concept and principles of quantum computing such as a quantum bit and superposition of states. Like other evolutionary algorithms, QEA is also characterized by the representation of the individual, the evaluation function and the population dynamics. A Q-bit is defined as the smallest unit of information in QEA, which is defined with a pair of numbers (α, β). It may be in the '1' state, in the '0' state, or in any superposition of the two. A Q-bit individual is a string of Q-bits. The state of a Q-bit can be changed by the operation with a quantum gate, such as NOT gate, Rotation gate, and hadamard gate, etc. Rotation gate is often used to update the Q-bit as follows:

$$\begin{bmatrix} \alpha_i' \\ \beta_i' \end{bmatrix} = \begin{bmatrix} \cos(\Delta\theta_i) & -\sin(\Delta\theta_i) \\ \sin(\Delta\theta_i) & \cos(\Delta\theta_i) \end{bmatrix} \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix}. \tag{2}$$

Table 1. The angle parameters used for rotation gate.

x_i	b_i	$f(x) \geq f(b)$	$\Delta\theta_i$	x_i	b_i	$f(x) \geq f(b)$	$\Delta\theta_i$
0	0	false	$\theta 1$	1	0	false	$\theta 5$
0	0	true	$\theta 2$	1	0	true	$\theta 6$
0	1	false	$\theta 3$	1	1	false	$\theta 7$
0	1	true	$\theta 4$	1	1	true	$\theta 8$

The angle parameters used for rotation gate are shown in Table 1. Where x_i and b_i are the *i*th bit of the best solution b and the binary solution x respectively. The structure of QEA is shown in previous work by Han [2].

4. THE GENE EXPRESSION DATA SETS

The first data set denoted as HL-60 is described in the work by Tomayo [3] and it contains data from macrophage differentiation experiments. It consists of 7229 genes and expression levels at four time points. We apply a variation filter which discarded all genes with an absolute change in maximum and minimum expression level less than 30. The number of genes which pass the filter is 2792. The vectors are then normalized to have mean 0 and variance 1. The second data set denoted as Yeast is described in the work by Cho [4]. It consists of 6602 yeast genes measured at 17 time points over two cell cycles. The 90-minute and 100-minute time points are excluded because of difficulties with scaling. Afterwards, we use a variation filter to discard all genes with an absolute expression level change less than 100, and an expression level of max/min <2.0. The number of genes that pass the filter is 2947. Again, the vectors are normalized to have mean 0 and variance 1.

5. IQEA FOR CLUSTERING GENE EXPRESSION DATA

In the experiments, we compare the improved IQEA with k-means and self-organizing maps (SOM) algorithms. IQEA and k-means algorithms are implemented in Matlab 6.5. The self-organizing maps algorithm is available in the software package Gene Cluster 2.0.

5.1 Initialization, Representation and Fitness Function

The k-means algorithm is first run ten times and so ten initial solutions are produced for each data set. Each solution is composed of *n* mean vectors where

n is the number of clusters. Given the mean vectors, the cluster memberships can be calculated by calculating the nearest distance of each gene vector with all mean vectors. The representation used in the IQEA is straightforward. There are n^*10 mean vectors in all for the ten initial solutions. Then we number each mean vector as 1, 2, 3 ... n^*10 and put them into the set V where $V = (v_1, v_2, \ldots, v_{n^*10})$. So n^*10 Q-bits are used in each Q-bit individual. A Q-bit individual has the following form in the tth generation:

$$\begin{bmatrix} \alpha_1^t & \alpha_2^t & \cdots & \alpha_n^t & \alpha_{n+1}^t & \cdots & \alpha_{2\times n}^t & \cdots & \alpha_{9n+1}^t & \cdots & \alpha_{10\times n}^t \\ | & | & | & | & | & | & | & | & | & | & \\ \beta_1^t & \beta_2^t & \cdots & \beta_n^t & \beta_{n+1}^t & \cdots & \beta_{2\times n}^t & \cdots & \beta_{9n+1}^t & \cdots & \beta_{10\times n}^t \end{bmatrix} \tag{3}$$

where α is initially given as a random number between 0 and 1 initially and β could be computed according to Equation (2). The fitness function used in the IQEA is the MSSC error provided in Equation (1).

5.2 Make and Repair Operation

To obtain the binary string, the step of 'Make(x)' by observing the states of Q-bit can be implemented for each Q-bit individual as follows:

```
If (random (0, 1) < $|\beta_i|^2$ && $k < n$) then $x(i) \leftarrow 1; k \leftarrow k + 1$;
else $x(i) \leftarrow 0$;
```

where the variant k is used to guarantee that the number of '1' in each binary string must be less than or equal to n. Afterwards, if the number of '1' is less than n in some strings, an additional mutation operation 'Repair (x)' should be performed to be sure that there should be and only be n '1' in each binary string as follows:

```
while k < n do
randpos $\leftarrow$ random (1, $n^*10$); if x (randpos) = 0 then x (randpos) $\leftarrow$ 1; $k \leftarrow k + 1$;
end
```

5.3 Evaluated and Updated Operation

Then the evaluated operation is executed to calculate the fitness value according to Equation (1) for each Q-bit individual and so the best individual can be selected. The updated operation is used to update Q-bit states of each individual by rotation gate in Equation (3). The angle parameter in Table 1 is

Table 2. Experiment results for the two data sets.

Dataset	Algorithm	Best obj	Avg obj	Excess
HL-60	IQEA	1514.3	1598.5	5.56%
	K-means	1514.6	1604.3	5.92%
	SOM	1523.2	1624.0	6.62%
Yeast	IQEA	15741.0	15860.0	0.76%
	K-means	15752.0	15905.0	0.97%
	SOM	15801.0	16002.0	1.27%

set as follows according to experiment observation:

$$\text{If } \alpha^*\beta > 0 \text{ then } \theta_3 = 0.01\ \pi; \theta_5 = 0.01\ \pi;$$
$$\text{If } \alpha^*\beta < 0 \text{ then } \theta_3 = 0.01\ \pi; \theta_5 = 0.01\ \pi;$$

6. EXPERIMENT RESULTS

In all experiments, IQEA is run with a population size of 20. The k-means and IQEA algorithms are all terminated when the maximum generation of 50 is arrived. The three algorithms are all run 10 times, and the best and the average objective value (fitness value) are produced. In Table 2, the results are displayed for the described two data sets. Excess value is the percentage of the average objective value above the best objective value. It is obvious that IQEA performs better than the other two algorithms.

7. CONCLUSIONS

In this paper, an improved quantum-inspired evolutionary algorithm (IQEA) by using a new representation form and adding an additional mutation operation is proposed for clustering gene expression data. We present experimental evidence that the proposed algorithm is effective and produces better solutions than the conventional k-means and self-organizing maps clustering algorithms even with a small population.

REFERENCES

1. P. Merz (2003), Analysis of gene expression profiles: an application of memetic algorithms to the minimum Sum-of-squares clustering problem. *Biosystem*, 72, pp. 99–109.

2. K. Han and J. Kim (2002), Quantum-inspired evolutionary algorithm for a class of combinatorial problem. *IEEE Transactions on Evolutionary Computation*, 6, pp. 580–593.
3. P. Tamayo and D. Slonim (1999), Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *PNAS*, 96, pp. 2907–2912.
4. R.J. Cho, E.A. Winzeler and R.W. Davis (1998), A genome-wide transcriptional analysis of the mitotic cell cycle. *Molecular Cell*, 2, pp. 65–73.

AUTOMATIC BUFFER OVERFLOW DETECTION BASED ON OPERATION SEMANTIC

Dongfan Zhao and Lei Liu
College of Computer Science and Technology, Jilin University, Changchun 130012, China

Abstract Buffer overflow is the most dangerous attack method that can be exploited. According to the statistics of CERT (Computer Emergency Readiness Team), buffer overflow accounts for 50% of the current software vulnerabilities, and this ratio is going up. Considering a subset of C language and Mini C, this paper presents an abstract machine model that can realize buffer overflow detection, which is based on operation semantic. Thus, the research on buffer overflow detection can be built on strict descriptions of operation semantic. Not only the correctness can be assured, but also the system can be realized and extended easily.

Keywords: buffer overflow detection, abstract machine, program analysis.

1. BUFFER OVERFLOW DETECTION TECHNOLOGY

Buffer overflow detection approaches can be divided into two categories: static approaches and dynamic approaches. Both dynamic approaches, such as StackGuard [1] and StackShield [2], and static approaches, such as ITS4 [3] and Rats [4], have their own disadvantages and limitations. Dynamic approaches increase the system operation spending, and could become a service rejected attack easily. This paper presents an approach to eliminate vulnerabilities of buffer overflow on the source code level, which investigates the problem from a static point of view.

A buffer is a piece of continuous space allocated in the memory. In general, buffer overflow could occur at anytime if the program writes more information into the buffer than the space where it is allocated. In order to detect a buffer overflow, a buffer is defined as (Alloc, NullPos). Alloc represents the

G. R. Liu et al. (eds.), Computational Methods, 1357–1361.

size of the buffer. NullPos represents the relative address in the buffer, whose initial value is zero. A dual set (IDE, Offset) is defined for each pointer. IDE represents the buffer, which the pointer points to. Offset represents the offset to the beginning address of the buffer, which the pointer points to. If there are operations on buffers, pointers or strings, the possible changes of the integer couple (Alloc, NullPos) of each buffer and the dual set (IDE, Offset) of each pointer are tracked. If NullPos ≥ Alloc, there could be an occurrence of buffer overflow.

Char arrays, pointers and string operation functions in C language are the main elements, which result in buffer overflow. The Name of a char array can be handled as a pointer, so operations on the buffer can be implemented by pointer operations. Based on this idea, an abstract machine model is applied to detect buffer overflow, which enables the research on buffer overflow detection to base on strict operation semantic descriptions. This approach not only assures the correctness of the buffer overflow detection, but also simplifies the implementation and extension of the system.

2. AN ABSTRACT MACHINE BASED DETECTION APPROACH

2.1 The definition of an abstract machine

An abstract machine (MS) consists of seven parts. They are programs (Prog), buffers (BBuffer), pointers (PPointer), a reservation (Reserv), a static environment (SEnv), a dynamic environment (DEnv) and alarms (Alarm). If p, μ, θ, r, ε, σ and α represent these seven parts, MS = (p, μ, θ, r, ε, σ, α). A detailed definition is as follows:

MS ::= Prog × BBuffer × PPointer × Reserv × SEnv × DEnv × Alarm
BBuffer ::= $(\mathrm{IDE}\rightarrow\mathrm{Alloc} \times \mathrm{NullPos})^*$
PPointer ::= $(\mathrm{IDE}\rightarrow\mathrm{IDE} \times \mathrm{Offset})^*$
Reserv ::= $((\mathrm{LAB}|\mathrm{VLAB}) \times \mathrm{STAT})^*$
SEnv ::= $(\mathrm{IDE}\rightarrow\mathrm{var}\ (\mathrm{TYPE}) + \mathrm{func}\ (\mathrm{PARAM}^* \times \mathrm{BLOCK} \times \mathrm{SEnv}))^*$
DEnv ::= $(\mathrm{IDE}\rightarrow\mathrm{int}\ (\mathrm{INT}) + \mathrm{bool}\ (\mathrm{BOOL}))^*$
Alarm ::= INT

Because redirection statements need to be considered, virtual labels are used to describe statements without labels. These virtual labels (VLAB) assure every statement has a label. If an initial state is $\mathrm{MS}^0 = (\mathrm{PROG}^0,[],[],[],[],[],[])$, the final state is $\mathrm{MS} = ([],\_,\_,\_,\_,\_,\alpha)$.

2.2 Main state transfer rules of an abstract machine

2.2.1 Variable declarations

An array declaration generates a buffer and its pointer. The size of the buffer is the size of the array. The position of the end note is zero. The pointer points to the beginning address of the generated buffer.

$(\mathrm{strdec}(\mathrm{IDE}\times \mathrm{N}\times \mathrm{DIM}_1\times\ldots\times \mathrm{DIM}_n) : \mathrm{VDECL} : \mathrm{FUNCL},\mu,\theta,r,\varepsilon,\sigma,\alpha) \rightarrow$
$(\mathrm{VDECL} : \mathrm{FUNCL},\mu \oplus (\mathrm{IDE}_k\rightarrow \mathrm{DIM}_n\times 0),$
$\theta\oplus(\mathrm{IDE}_k\rightarrow \mathrm{IDE}_k\times 0),r,\varepsilon(\mathrm{IDE}_k\rightarrow \mathrm{STRING}), \sigma, \alpha)\ \ k = 1..\mathrm{DIM}_1\times\ldots\times \mathrm{DIM}_{n-1}$

When a pointer is declared, the pointed buffer is marked as NULL and the Offset is initialized as zero.

$(\mathrm{pointdec}(\mathrm{IDE}):\mathrm{VDECL}:\mathrm{FUNCL},\mu,\theta,r,\varepsilon,\sigma,\alpha)$
$\rightarrow(\mathrm{VDECL}:\mathrm{FUNCL},\mu,\theta \oplus (\mathrm{IDE}\diamond \mathrm{NULL} \times 0),r,\varepsilon(\mathrm{IDE}\rightarrow \mathrm{POINT}),\sigma,\alpha)$

If there are other variable declarations, only the relevant static environments need to be changed.

$(\mathrm{otherdec}(\mathrm{IDE}\times \mathrm{TYPE}):\mathrm{VDECL}:\mathrm{FUNCL},\mu,\theta,r,\varepsilon,\sigma,\alpha)$
$\rightarrow(\mathrm{VDECL}:\mathrm{FUNCL},\mu,\theta,r,\varepsilon(\mathrm{IDE}\rightarrow \mathrm{TYPE}), \sigma,\alpha)$

2.2.2 Function declarations

$(\mathrm{func}(\mathrm{FIDE}\times \mathrm{PARAM}^*\times \mathrm{BLOCK}): \mathrm{FUNCL},\mu,\theta,r,\varepsilon,\sigma,\alpha) \rightarrow$
$(\mathrm{BLOCK}:\mathrm{FUNCL},\mu, \theta,r,\varepsilon(\mathrm{FIDE}\rightarrow \mathrm{func}(\mathrm{PARAM}^*\times \mathrm{BLOCK}\times \varepsilon)),\sigma,\alpha)$

2.2.3 Statements

$(\mathrm{skip}:\mathrm{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha) \rightarrow (\mathrm{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)$
$(\mathrm{label}(\mathrm{LAB}):\mathrm{COMM}:\mathrm{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)$
$\rightarrow(\mathrm{COMM}:\mathrm{COMML}, \mu,\theta, (\mathrm{LAB}\times \mathrm{COMM}): r, \varepsilon,\sigma,\alpha)$
$(\mathrm{expr}(\mathrm{EXPR}):\mathrm{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)$
$\rightarrow(\mathrm{EXPR}:\mathrm{COMML}, \mu,\theta, (\mathrm{VLAB}\times \mathrm{expr}(\mathrm{EXPR})): r, \varepsilon,\sigma,\alpha)$
$(\mathrm{if}(\mathrm{EXPR}\times \mathrm{COMML1}\times \mathrm{COMML2}):\mathrm{COMML},\mu, \theta,r,\varepsilon,\sigma,\alpha)$
$\rightarrow(\mathrm{expr}(\mathrm{EXPR})\rightarrow(\mathrm{COMML1}:\mathrm{COMML},\mu,\theta,$
$(\mathrm{VLAB}\times \mathrm{if}(\mathrm{EXPR}\times \mathrm{COMML1}\times \mathrm{COMML2})):r,\varepsilon,\sigma,\alpha),$
$(\mathrm{COMML2}:\mathrm{COMML},\mu,\theta,$
$(\mathrm{VLAB}\times \mathrm{if}(\mathrm{EXPR}\times \mathrm{COMML1}\times \mathrm{COMML2})):r,\varepsilon,\sigma,\alpha))$
$(\mathrm{while}(\mathrm{EXPR}\times \mathrm{COMML1}):\mathrm{COMML}, \mu,\theta,r,\varepsilon,\sigma,\alpha)$
$\rightarrow(\mathrm{exp}(\mathrm{EXPR})\rightarrow(\mathrm{COMML1}:\mathrm{while}(\mathrm{EXPR}\times \mathrm{COMML1}):$

$\text{COMML},\mu,\theta,(\text{VLAB}\times\text{while}(\text{EXPR}\times\text{COMML})): r,\varepsilon,\sigma,\alpha),$
$(\text{COMML},\mu,\theta,(\text{VLAB}\times\text{while}(\text{EXPR}\times\text{COMML1})):r,\varepsilon,\sigma,\alpha))$
$(\text{goto}(\text{LAB}):\text{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)$
$\rightarrow(\text{lab}(\text{LAB},\text{COMML}),\mu,\theta,(\text{VLAB}\times\text{jmp}(\text{LAB})):r,\varepsilon,\sigma,\alpha)$

2.2.4 Expressions

$(\text{const}(\text{CONSTANT}):\text{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)\rightarrow(\text{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)$
$(\text{var}(\text{IDE}):\text{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)\rightarrow(\text{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)$
$(\text{cal}(\text{EXPR}):\text{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)\rightarrow(\text{EXPR}:\text{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)$
$((\text{assspec}(\text{VAR}\times\text{EXPR}):\text{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)\rightarrow(\text{COMML},\mu',\theta',r,\varepsilon,\sigma',\alpha')$
Where $(\mu',\theta',\sigma',\alpha')=\text{UpdateAss}(\text{VAR}\times\text{EXPR})$

If there are other assign statements, the states need to be updated. Details are omitted.

$((\text{assother}(\text{EXPR1}\times\text{EXPR2}):\text{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)\rightarrow$
$(\text{COMML},\mu,\theta,r,\varepsilon,\sigma',\alpha)$
$(\text{call}(\text{PIDE}\times\text{EXPR}^*):\text{COMML},\mu,\theta,r,\varepsilon(\text{PIDE}\times\text{PARAM}^*\times\text{BLOCK}\times\varepsilon),\sigma,\alpha)$
$\rightarrow((\text{assother}(\text{PARAM}\times\text{EXPR}))^*:\text{BLOCK}:\text{ENDFUNC}:$
$\text{COMML},\mu\Delta\mu1,\theta\Delta\theta1,r,\varepsilon\Delta\rho,\sigma\Delta\beta,\alpha)$
$(\text{call}(\text{STRPIDE}\times\text{EXPR}):\text{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)\rightarrow(\text{COMML},\mu',\theta,r,\varepsilon,\sigma',\alpha')$
Where $(\mu',\sigma',\alpha')=\text{UpdateFunc}(\text{STRPIDE}\times\text{EXPR}^*)$
$(\text{ENDFUNC}:\text{COMML},\mu\Delta\mu1,\theta\Delta\theta1,r,\varepsilon\Delta\rho,\sigma\Delta\beta,\alpha)\rightarrow(\text{COMML},\mu,\theta,r,\varepsilon,\sigma,\alpha)$

3. A BUFFER OVERFLOW DETECTION EXAMPLE

The approach of Wagner et al. [7] cannot detect buffer overflow in the following code, however, the approach described in the paper can detect the buffer overflow properly.

The program code is as follows.

char s[20], *p, t[10]; strcpy(s,“Hello”); p = s + 5; strcpy(p,“world!”); strcpy(t,s);

According to NullPos > Alloc of t, it can be detected that the statement strcpy (t, s) results in buffer overflow of t.

4. CONCLUSIONS

This paper presents an abstract machine model for buffer overflow detection from operation semantic point of view. The strict operation semantic description eliminates different meanings, simplifies the implementation and extension of the system and assures the correctness. Following the approach presented in this paper, we developed a prototype system, which is benefit from Flex and Accent. This prototype system primarily proves the correctness and feasibility of this approach.

REFERENCES

1. C. Cowan, C. Pu, D. Maier, H. Hintony, J. Walpole, P. Bakke, S. Beattie, A. Grier, P. Wagle and Q. Zhang, *StackGuard: Automatic Adaptive Detection and Prevention of Buffer-Overflow Attacks.* Department of Computer Science and Engineering, Oregon Graduate Institute of Science & Technology.
2. StackShield technical info file v0.7 Vendicator (January 2001), Available online at http://www.angelfire.com/sk/stackshield.
3. J. Viega, J.T. Bloch, T. Kohno and G. McGraw (2000), *ITS-4: A Static Vulnerability Scanner for C and C++ Code.* Reliable Software Technologies.
4. Secure Software Soliutions (September 2001), *Rough Auditing Tool for Security, RATS 1.3*. Available online at http://www.securesw.com/rats.
5. J. Wilander and M. Kamkar (2003), *A Comparison of Publicly Available Tools for Dynamic Buffer Overflow Prevention.* Department of Computer and Information Science, Linköpings universitet.
6. J. Wilander and M. Kamkar (2003), *A Comparison of Publicly Available Tools for Static Intrusion Prevention.* Department of Computer and Information Science, Linköpings universitet.
7. D. Wagner, J.S. Foster, E.A. Brewer and A. Aiken (2000), *A First Step Towards Automated Detection of Buffer Overrun Vulnerabilities.* University of California, Berkeley.

QUANTUM-INSPIRED EVOLUTIONARY ALGORITHM FOR TRAVELLING SALESMAN PROBLEM

X.Y. Feng, Y. Wang, H.W. Ge, C.G. Zhou and Y.C. Liang
College of Computer Science and Technology, Jilin University, Key Laboratory
of Symbol Computation and Knowledge Engineering of the Ministry of Education,
Changchun 130012, China

Abstract A novel quantum coding mechanism is proposed to solve the travelling salesman problem (TSP) based on the quantum-inspired evolutionary algorithm. It adopts Q-bit individual to encode the visited sequence of the cities and employs the quantum rotation gate to adjust the population dynamically. Experimental results of 14 cities show that the proposed approach is feasible and effective for small-scale TSPs, which indicates a promising novel approach for solving TSPs.

Keywords: quantum-inspired evolutionary algorithm, Q-bit individual, quantum computing, travelling salesman problem.

1. INTRODUCTION

Quantum computing was proposed by Benioff and Feynman in the early 1980s. Quantum computing can solve many difficult problems in the field of classical computation based on the concepts and principles of quantum theory, such as superposition of quantum states, entanglement and intervention. Due to its unique computational performance, the quantum computing has attracted extensive attention of researchers [1, 2]. Han [3] proposed the quantum-inspired evolutionary algorithm (QEA), inspired by the concept of quantum computing. Like other evolutionary algorithms, QEA is also characterized by the representation of the individual, the evaluation function and the population dynamics. However, instead of binary, decimal or symbolic representation, QEA used a Q-bit, defined as the smallest unit of information and a Q-bit individual as a string of Q-bits. Besides, a Q-gate is introduced as a variation operator to

G. R. Liu et al. (eds.), Computational Methods, 1363–1367.

promote the optimization of the individuals. Han and Yang [4, 5] have applied the QEA to some optimization problems. The results show that the performance of the QEA is better than the traditional evolutionary algorithm.

Travelling salesman problem (TSP) is a well-known NP-hard combinatorial optimization problem, which is the generalization and simple form of many complicated problems in different fields. It is easy to be described but hard to deal with. In this chapter, a novel quantum coding mechanism is proposed to solve it, in which the Q-bit is adopted to represent the population diversity, the observation process is used to obtain a binary string and the Q-gate is applied to update the population. The best solution can be given when the termination condition is satisfied. The binary strings in the observation process represent the visited sequence of the cities. The test results show that it is feasible to apply the QEA to solve the TSP.

2. THE PROPOSED APPROACH

2.1 Brief introduction to QEA

In QEA, the smallest unit of information is called a Q-bit [3], which is defined as $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$, where α and β are complex numbers that specify the probability amplitudes of the corresponding states. The moduli $|\alpha|^2$ and $|\beta|^2$ are the probabilities that the Q-bit exists in state '0' and state '1', respectively, which satisfy that $|\alpha|^2 + |\beta|^2 = 1$. A Q-bit individual as a string of m Q-bit is defined as

$$\left[\begin{array}{c|c|c|c} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \beta_1 & \beta_2 & \cdots & \beta_n \end{array}\right], \quad \text{where } |\alpha_i|^2 + |\beta_i|^2 = 1 \quad (i = 1, 2, \ldots, n).$$

The state of a Q-bit can be changed by operating a quantum gate. There are several quantum gates. It can be chosen according to the problem. In this chapter, the quantum rotation gate is adopted.

2.2 Quantum coding of TSP

According to the characteristic of the TSP, the quantum population is defined as $P(t) = \{p_1^t, p_2^t, \ldots, p_k^t\}$ where t is the number of the generation and k is the scale of population. $p_i^t = q_{i1}q_{i2}\ldots q_{im}$ where $i = 1, 2, \ldots, k$, m is the number of the cities. q_{ij} is a Q-bit individual. We adopt the coding mechanism of Han [3] to define q_{ij} as $\left[\begin{array}{c|c|c|c} \alpha_{ij_1} & \alpha_{ij_2} & \cdots & \alpha_{ij_n} \\ \beta_{ij_1} & \beta_{ij_2} & \cdots & \beta_{ij_n} \end{array}\right]$ where n is the number of

Q-bit, i.e., the string length of the Q-bit individual, $j = 1, 2, \ldots, m$, $\left|\alpha_{ij_s}\right|^2 + \left|\beta_{ij_s}\right|^2 = 1$ and $s = 1, 2, \ldots, n$. m and n satisfy that $2^n \geq m$.

2.3 Fitness function

The solution is obtained by observing the individuals. When observing the state of p_i^t, q_{ij} will be represented by a binary string $X_{ij} = (x_{j1}x_{j2}\ldots x_{jn})$, where the value $x_{js}(s = 1, 2, \ldots, n)$ is 0 or 1 determined by the probability $\left|\alpha_{ij_s}\right|$ or $\left|\beta_{ij_s}\right|$, respectively.$X_{ij}$ denotes the visited sequence of the jth city. Then the individual p_i^t is described as $X_{i1}X_{i2}\ldots X_{im}$. Sorted $X_{i1}X_{i2}\ldots X_{im}$, we get the sequence $X = X_{ic_1}X_{ic_2}\ldots X_{ic_m}$ where $c_t = 1, 2, \cdots m(t = 1, 2, \ldots m)$. Then the solution is a binary string of length m*n defined as X. The sequence of the visited cities is $c_1 \to c_2 \to \cdots \to c_n$. The cost is $T_d = \sum_{i=1}^{n-1} d(c_i, c_{i+1}) + d(c_n, c_1)$ where $d(c_i, c_j)$ denotes the distance between city c_i and city c_j. The fitness function is taken as $f = -T_d$.

2.4 Implementation

The procedure of the proposed approach based on QEA is summarized as follows:

1. Generate an initial population randomly, and encode each individual with the proposed quantum coding mechanism.
2. Calculate the fitness value of each individual.
3. Replace the local best value with the arising better one and find the global best value from the local ones.
4. Update the population by QEA until the maximum iterative times are reached.

3. EXPERIMENTAL RESULTS

The performance of the proposed algorithm for TSP is examined by the benchmark problem with 14 nodes from the standard TSPLIB [6]. The experiment is implemented on a PC (PentiumIV-2.6GHz, 512M RAM, Windows XP, Matlab6.5). The related parameters of the approach are set as follows: $k = 20$, $m = 4, n = 14$.

The obtained optimal solution is

$$1 \to 10 \to 9 \to 11 \to 8 \to 13 \to 7 \to 12 \to 6 \to 5 \to 4 \to 3 \to 14 \to 2$$

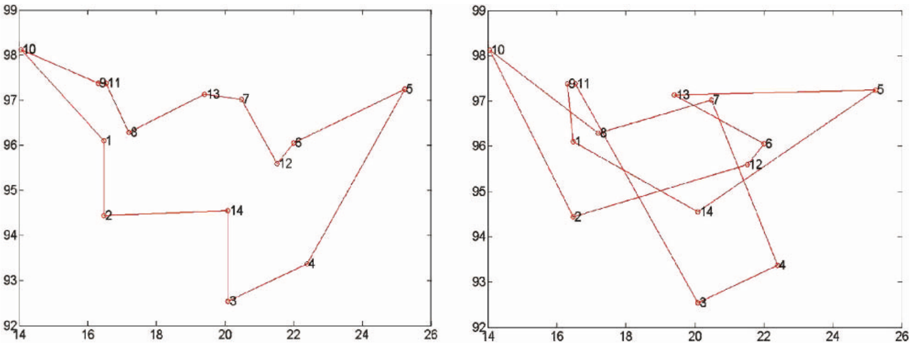


Figure 1. (a) Initial solution. (b) The best solution.

and its cost is 30.8785, which is equal to that obtained by Wang and Huang [7]. The initial random solution with cost 49.7073 and the best one obtained by using the proposed algorithm are shown in Figures 1a and b, respectively.

The performance of the proposed algorithm is finally analysed to show its validity clearly. The scale of the solution space for the benchmark problem is 14!/(14*2) = 3,113,510,400 while the scale of the searched space is the product of the number of the individual, the Q-bit individual and the running iterations, which is equal to 56*20*2000 = 224,000. It is easy to conclude that the searched space is only 0.072% of the solution space, which shows the high convergent efficiency of the proposed algorithm.

4. CONCLUSIONS

A novel quantum coding mechanism on the travelling salesman problem is proposed based on the QEA in this chapter. The test results show that the proposed approach can obtain the best solution by searching a small-size proportion of the solution space. The behaviour of the performance shows worse when the number of the cities increases. We ascribed this to the binary string coding which we used to represent the visit orders of the cities. The study on the limitation of the binary coding is in progress.

ACKNOWLEDGEMENT

This work was supported by the science-technology development project of Jilin Province of China under Grant No. 20030520, the doctoral funds of the National Education Ministry of China under Grant No. 20030183060 and the key science-technology project of the National Education Ministry of China under Grant No. 02090.

REFERENCES

1. P.W. Shor (1994), Algorithms for quantum computation: Discrete logarithms and factoring. In: *Proceedings of the 35th Annual Symposium on Foundations of Computer Science,* IEEE Press, Piscataway, NJ, pp. 124–134.
2. L.K. Grover (1996), A fast quantum mechanical algorithm for database search. In: *Proceedings of the 28th ACM Symposium on Theory of Computing*, pp. 212–219.
3. K.H. Han and J.H. Kim (2002), Quantum-inspired evolutionary algorithm for a class of combinatorial optimization. *IEEE Transactions on Evolutionary Computation*, 6, 6, pp. 580–593.
4. K.H. Han and J.H. Kim (2004), Quantum-inspired evolutionary algorithms with a new termination criterion, Hε gate, and two-phase scheme. *IEEE Transactions on Evolutionary Computation*, 8, 2, pp. 156–169.
5. J.A. Yang, B. Li, Z.Q. Zhuang and Z.F. Zhong (2003), Quantum genetic algorithm and its application research in blind source separation, *Mini-Micro Systems (in Chinese)*, 24, 8, pp. 1518–1523.
6. URL (1997), Standard TSPLIB: http://elib.zib.de/pub/Packages/mp-testdata/tsp/tsplib/tsplib.html.
7. K.P. Wang, L. Huang, C.G. Zhou and W. Pang (2003), Particle swarm optimization for traveling salesman problem. In: *Proceeding of the 2nd ICMLC, Xi'an,* Vol. 3, IEEE Press, pp. 1583–1585.

AN IMPROVED COEVOLUTION ALGORITHM FOR FUZZY MODELLING

Y.N. Lu, Y.L. Liu and Y.C. Liang
College of Computer Science and Technology, Jilin University, Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, Changchun 130012, China

Abstract A modified coevolution algorithm is proposed and applied to fuzzy modelling. Two populations are evolved in the evolutionary processes with parallel running: membership functions by genetic algorithm (GA) and rule sets by genetic programming (GP). Based on rough set theory, the significance of the attributes of the sample data sets is used to benefit the evolution. Moreover, a tree-pruning operator is introduced in GP to limit the code bloat.

Keywords: fuzzy modelling, coevolution, significance of attribute, tree pruning.

1. INTRODUCTION

Coevolution refers to the simultaneous evolution of two or more species with coupled fitness. It favours the evolutionary problems with some of the following features: (1) the solution is complex, (2) the problem is clearly decomposable, (3) the genomes of the components are different types of values, (4) strong interdependencies among the components of the solution and (5) component-ordering drastically affects the fitness. Potter [1] developed a model, where a number of populations explore different decompositions of the problem. Each species represents a subcomponent of a potential solution. The fitness of each individual depends on how well it cooperates with the individuals of other species to solve the problem.

Mendes et al. [2] discovered fuzzy classification rules with genetic programming and coevolution, where the individual encoding schema incorporates several syntactical restrictions . Pena-Reyes and Sipper [3] applied cooperative coevolutionary approach to fuzzy modelling. These algorithms deal with the

G. R. Liu et al. (eds.), Computational Methods, 1369–1373.

simultaneous search for two species: the operational and the connective parameters of a fuzzy inference system. The operational parameters evolved by genetic algorithm (GA) define the membership functions, and the connective parameters evolved by genetic programming (GP) or GA define the rule set. These two species are separate but coupled. Coevolutionary approach performs well in fuzzy modelling.

Genetic fuzzy system, which integrates the fuzzy logic and evolutionary algorithm, has been widely applied to the problems in classification and control [4]. The main problem of fuzzy system is the curse of dimensionality. As a powerful methodology of soft computing, rough set theory has been a useful tool of attribute reduction [5]. In this paper, we introduce the attribute significance concept based on the rough set theory to the evolving process, and also introduce a tree-pruning operator into GP to limit the code bloat.

2. COEVOLUTIONARY FUZZY MODELLING

As mentioned above, two cooperating species are separately created in coevolutionary fuzzy modelling. Like the work of Mendes [2], the individuals of the first GA species define the membership functions. Each variable is associated with its own trapezoidal membership functions. Every individual tree in GP species corresponds to a set of rule antecedents encoded in disjunctive normal form (DNF), and the rule consequents are implied. The non-terminal nodes express the logical functions of {OR, AND, NOT}, and the terminal nodes express the conditions of the form: ‘Val_i is Lab_{ij}’.

The pseudo-code of the algorithm is given below.

```
begin Coevolution for Fuzzy Modelling
   g := 0
   S := 1, initialize GA population $P_s(0)$
   S := 2, initialize GP population $P_s(0)$ with $P_1(0)$
   Evaluate $P_1(0)$, $P_2(0)$
   While not done do
      g := g + 1;
      For each species s(s = 1,2)
         $E_s(g)$ = elite-select $P_s(g - 1)$
         $P_s'(g)$ = Select[$P_s(g)$]
         $P_s''(g)$ = ApplyGeneticOperators[$P_s'(g)$]
         Evaluate population $P_s''(g)$, and prune tree if s = 2
         $P_s(g)$ = Introduce[$E_s(g)$,$P_s''(g)$]
      End for
   End while
End
```

In the algorithm, while initializing, each GP individual is associated with a GA individual. We use the GA individual for the fuzzy discretization of attribute (variable in fuzzy system) values. Then we evaluate the significance of the variables by a rough set theory algorithm [5]. In our algorithm, some attributes are removed from the variable set if their significance smaller than a threshold T_s. Besides, the variables with higher significance are more possible to appear in the GP individuals' terminal nodes.

3. GENETIC OPERATORS AND FITNESS FUNCTION

We use the tournament selection and elite strategy. Crossover and mutation of GA species are conventional. In GP species, crossover and mutation must guarantee that the generated offspring is always syntactically-valid. In the evolving process, we adopt a new mutation operated on GP species based on the significance of the variables. While mutating a GP individual, a GA individual is randomly chosen from the GA cooperators, and is used to evaluate the significance of the attributes. Then in the GP individual, the terminal nodes having variables with low significances are replaced by a new generated tree only having the variables with high significances.

The fitness of the individual depends on F_i, the fitness of the fuzzy systems combined this individual with the ith cooperators of another species. Some of the cooperators are the fittest individuals of the population, and the others are selected randomly. The fitness F_i bases on two criteria: (1) sensitivity S_e, or true-positive ratio; (2) specificity S_p [6]. The fitness of fuzzy system is given by

$$F_i = S_e * S_p \tag{1}$$

The fitness of the individual is $F = \sum_{i=1}^{n} F_i/n$, where n is the number of cooperators.

4. TREE PRUNING

Code bloat in GP greatly affects the performance of the system [7]. A tree-pruning operator is designed to counteract it. This operator also can speed up the evolution and improve the intelligibility of the GP result. In the pruning operator, if all of the extents to which the cases match one condition part in the individual are smaller than a threshold T_v, the node representing this condition part is marked with 0, otherwise 1. The operator is applied to the individuals whose fitness are higher than T_f with the probabilityP_h. The process is shown below:

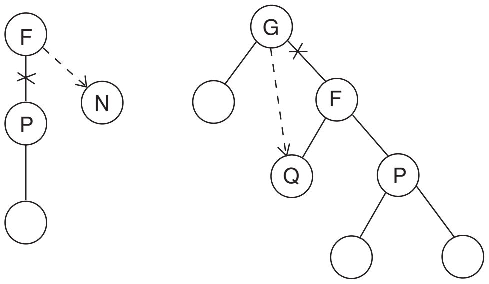


Figure 1. Tree pruning.

1. Before evaluating the GP individual, mark each function node with 0.
2. In the evaluating process, if the extent to which a case matches one node is higher than the threshold T_v, mark the node with 1.
3. After evaluating, search for the node P satisfying the following conditions:
   (1) Marked with 0
   (2) The number of descendants is higher than a thresholdC_v
4. Prune the tree as the method shown in Figure 1. If P is a single-branch node (refer to the left figure), the subtree rooted at P is replaced by a randomly generated new terminal node N; if not (refer to the right figure), then P's father node F, P and the subtree are replaced by P's sibling node.

5. EXPERIMENTS

We apply the improved algorithm to the sleep apneas problem. It has 150 cases, each of which is described by 15 variables: age, weight, height, blood pressure, rhythm of the heart, etc. The data set was divided into a training set and testing set. Our system gets a best fuzzy rule set with five variables. Considering that parameters of accuracy (A_{cc}), positive predictive value (PPV) and average rule size (A_{rs}) could be used to measure the performance of the algorithms [3], we also use them to examine the validation of the system. The results listed in Table 1 show that the improved method increases the accuracy and the PPV of the fuzzy system, and also decreases the size of the rule set.

Table 1. Performance comparison.

Method	A_{cc}	PPV	A_{rs}
Cooperation	76.85%	61.15%	14
Improved cooperation	79.54%	65.24%	5

6. CONCLUSIONS

In this paper we describe an improved coevolution algorithm for fuzzy modelling. The system uses two evolutionary algorithms: a genetic algorithm (GA) evolving the membership function population and a genetic programming (GP) evolving the fuzzy rule set generation. The significance of attributes and the tree pruning to the algorithm are introduced. The effectiveness of the proposed algorithm is examined using a medical data set.

ACKNOWLEDGEMENT

This work was supported by the science-technology development projects of Jilin Province of China (20040531, 20030520), the doctoral funds of the National Education Ministry of China (20030183060) and the key science-technology project of the National Education Ministry of China (02090).

REFERENCES

1. M.A. Potter and K.A. DeJong (2000), Cooperative coevolutioin: Anarchitecture for evolving coadapted subcomponents. *Evolutionary Computation*, 8, 1, pp. 1–29.
2. R.R.F. Mendes, F.B. Voznika, A.A. Freitas and J.C. Nievola (2001), Discovering fuzzy classification rules with genetic programming and co-evolution. *Lecture Notes in Artificial Intelligence*, 2168, pp. 314–325.
3. C.A. Pena and M. Sipper (2001), Fuzzy CoCo: Balancing accuracy and interpretability of fuzzy models by means of coevolution. *IEEE Transactions on Fuzzy Systems*, 9, 5, pp. 727–737.
4. L. Magdalena, O. Cordon, F. Gomide, F. Herrera and F. Hoffmann (2004), Ten years of genetic fuzzy systems: Current framework and new trends. *Fuzzy Sets & Systems*, 141, 1, pp. 5–31.
5. Z. Pawlak (1998), Rough sets theory and its applications to data analysis. *Cybernetics & Systems*, 29, 199, pp. 661–688.
6. M.L. Wong and K.S. Leung (2000), *Data Mining Using Grammar Based Genetic Programming and Applications*. Kluwer Academic Publishers.
7. T. Soule (1998), Code growing in genetic programming. PhD Thesis, University of Idaho, Moscow, Idaho, USA.

A COMPUTATIONAL METHOD FOR SOLVING CAUCHY PROBLEMS OF ELLIPTIC OPERATORS

Y.C. Hon, T. Wei and L. Ling
Department of Mathematics, City University of Hong Kong, Hong Kong SAR, China

Abstract In this talk we combine a meshless *method of fundamental solution* (MFS) with different regularization methods to solve Cauchy problems of elliptic differential operators. The main idea of MFS is to approximate the unknown solution by a linear combination of fundamental solutions whose singularities are located outside the computational domain. Three regularization methods based on the singular value decomposition with four different choice strategies for regularization parameters are proposed.

Keywords: method of fundamental solution, discrete ill-posed problem, inverse problem.

1. INTRODUCTION

The method of fundamental solutions (MFS) is a technique for finding the numerical solution of certain elliptic boundary value problem. MFS belong to the general class of boundary collocation methods. The unknown solution is represented as a linear combination of fundamental solutions with their source points located outside the computational domain. The boundary conditions are satisfied by collocation or by least squares fitting. Unlike finite element method (FEM) and boundary element method (BEM), the MFS does not involve integral evaluation and hence provides an efficient computational alternative for problems in higher dimension with irregular domains.

MFS has recently been used extensively for solving linear partial differential equations of various kinds; for examples, Laplace equation [1, 2], harmonic equation [3], biharmonic equations [4], elastostatics problems [5] and

G. R. Liu et al. (eds.), Computational Methods, 1375–1384.

wave scattering problems [6, 7]. Moreover, MFS was applied to the solution of non-homogeneous linear and non-linear Poisson equations [8–11]. Details can be found in the review papers of Fairweather and Karageorghis [12] and Golberg and Chen [13]. All these studies focus on the well-posed *forward* problems in which the Dirichlet or Neumann data on the whole boundary are known. On the other hand, boundary conditions are usually complicated and incomplete for *inverse* problems. The truly meshfree MFS is an excellent candidate for solving these problems.

The Cauchy problem is a typical inverse problem which is severely ill-posed in nature. That is, the solution does not depend continuously on the boundary data. A small error in the given data may result in an enormous error in the numerical solution. The Cauchy problem arises from many branches of science and engineering such as non-destructive testing [14], steady-state inverse heat conduction problem [15], electro-cardiology [16] and steady-state inverse heat conduction [15]. For the Cauchy problem for the Laplace equation, a number of numerical methods have been researched over the past years [17–23]. When boundary collocation method is employed for MFS, the resultant linear systems will be severely ill-conditioned and standard numerical techniques for solving linear algebraic equations do not work well. Special techniques are required to solve such problems [24].

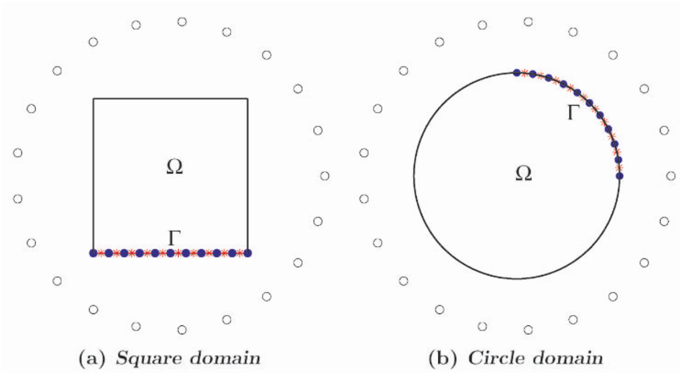
In this paper, numerical methods based on fundamental solution and regularization techniques are proposed to solve the ill-posed Cauchy problem. The outline is as follows: in 'Section 2', the Cauchy problem is introduced. The key ideas of MFS are reviewed in 'Section 3' and we show how MFS can be used to solve the Cauchy problem with noise-free data. In 'Section 4', the coupling of MFS with regularization techniques for noisy data are considered.

2. CAUCHY PROBLEMS FOR ELLIPTIC OPERATORS

Let Δ be the Laplace operator in R^d. Three elliptic differential operators are investigated:

$$\begin{aligned} \text{Laplace:} \quad & \mathrm{L}w = \Delta w, \\ \text{Helmholtz:} \quad & \mathrm{L}w = \Delta w + k^2 w, \\ \text{Modified Helmholtz:} \quad & \mathrm{L}w = \Delta w - k^2 w, \end{aligned} \tag{1}$$

for some wave number $k > 0$. Let Ω be a bounded and simply connected domain in R^d for $d = 2, 3$ with Lipschitz boundary $\partial\Omega$ and let $\Gamma \subset \partial\Omega$ be an open part of the boundary. Let L be a second order elliptic operator in R^d.



(a) *Square domain* (b) *Circle domain*

Figure 1. Cauchy Problem for two domains Ω . Dots (•) are collocation points for Dirichlet data, stars (∗) represent collocation points for Neumann data and circles (○) are source points.

The Cauchy problem aims to find a distribution function $w \in C^2(\Omega) \cap C^1(\bar{\Omega})$ which satisfies,

$$Lw = 0, \quad \text{in } \Omega, \tag{2}$$

$$w|_\Gamma = \varphi, \quad \text{and} \quad \frac{\partial w}{\partial n}|_\Gamma = \psi, \tag{3}$$

from discrete (and noisy) observed Dirichlet and Neumann boundary data φ and ψ on *part of the boundary* Γ, see Figure 1. In (3), $\frac{\partial w}{\partial n}$ denotes the outward normal derivative of w.

It is well known that the Cauchy problem presented above is extremely ill-posed. That is, any small noise to the given data may induce an enormous error to the numerical solution. However, a solution that continuously depends on the given data can be obtained under additional a priori conditions. This is commonly referred as *conditional stability* for Laplace's equation [25, 26].

3. MFS FOR CAUCHY PROBLEMS

A fundamental solution to the differential operator L is a functional G such that

$$\mathrm{L}G(P, Q) = \delta(P - Q),$$

where δ is a Dirac-delta function. From [27], the fundamental solutions to the elliptic operators in (1) are given in Table 1 where P and Q are points in R^d and $|P - Q|$ denotes their Euclidean distance R^d. In Table 1, H_0^2 and H_1^2 are Hankel functions of the second kind with order 0 and 1, respectively; K_0 and K_1 are modified Bessel functions of the second kind with order 0

Table 1. Fundamental solutions for elliptic operators in R^2 and R^3.

L	R^2	R^3
Laplace:	$\frac{1}{2\pi}\log\|P-Q\|$	$-\frac{1}{4\pi\|P-Q\|}$
Helmholtz:	$\frac{i}{4}H_0^2(k\|P-Q\|)$	$-\frac{\exp(-i\cdot k\|P-Q\|)}{4\pi\|P-Q\|}$
Modified Helmholtz:	$-\frac{1}{2\pi}K_0(k\|P-Q\|)$	$-\frac{\exp(-k\|P-Q\|)}{4\pi\|P-Q\|}$

and 1, respectively. When the singular *source point* $Q \notin \bar{\Omega}$ is located outside the domain, the fundamental solution satisfies the elliptic equation in Ω.

The idea of MFS [28] is to express the unknown solution w in (2) as a linear combination of fundamental solutions,

$$w \approx W := \sum_{i=1}^{M} \lambda_j G(\cdot, Q_j), \quad \text{in } \Omega \cup \partial\Omega, \tag{4}$$

where $\lambda := \{\lambda_j\}_{j=1}^M$ are expansion coefficients to be determined and $\{Q_j\}_{j=1}^M$ is a set of M source points in the exterior of $\bar{\Omega}$. By construction, W satisfy (2) in the interior, Ω, for any $\lambda_i \in$ C. The values of λ in (4) can then be determined merely from the boundary conditions (3) by using the methods of *direct collocation*, *linear least-squares fit* or *integral fit* of residual. Since the source points Q can be placed anywhere outside of the domain Ω, the MFS is truly meshfree. A comparison between MFS and the general class of meshfree methods can be found in [29].

When the boundary conditions are given over the whole boundary, the problem is well-posed. In this case, Cheng [30] gave a convergence result of MFS on a circular domain in R^2. Unlike forward problem, the boundary data in the Cauchy problem are known only on part of the boundary. Suppose that there are N_D and N_N *exact* Dirichlet and Neumann data available on Γ. Let $N = N_D + N_N$. We choose a set of N *collocation points* $\{P_k\}_{k=1}^N$ on Γ where $M \le N$. The unknown expansion coefficient λ can then be determined from the following linear system by using direct collocation ($M = N$) or least-squares fit ($M < N$) :

$$W(P_k) = \sum_{i=1}^{M} \lambda_j G(P_k, Q_j), \quad \text{for all} \quad k = 1, \dots, N. \tag{5}$$

In matrix form, (5) can be expressed as

$$A\lambda = b, \tag{6}$$

where

$$A = \begin{pmatrix} G(P_{k_1}, Q_j) \\ \dfrac{\partial G}{\partial n}(P_{k_2}, Q_j) \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} \varphi(P_{k_1}) \\ \psi(P_{k_2}) \end{pmatrix}, \tag{7}$$

for $k_1 = 1, \ldots, N_D,\ k_2 = N_D + 1, \ldots, N$ and $j = 1, \ldots, M$.

The main analysis tool here is the singular value decomposition (SVD), see [31] for instance. The *reduced* SVD of an overdetermined N-by-M matrix A is of the form

$$A = \sum_{i=1}^{M} u_i \sigma_i v_i^* = U \Sigma V^*, \tag{8}$$

where $U = [u_1, u_2, \ldots, u_M]$is the N-by-M left singular matrix, $V = [v_1, v_2, \ldots, v_M]$ is the M-by-M right singular matrix and $\Sigma = \operatorname{diag}(\delta_1 \ldots, \delta_M)$ is a M-by-M diagonal matrix with non-negative and non-increasing diagonal elements such that

$$\sigma_1 \geq \ldots \geq \sigma_M \geq 0.$$

The numbers δ_i are the singular values of A whereas the vectors u_i and v_i are the left and right singular vectors of A respectively. The vector v_i^* is the Hermitian or complex conjugate transpose of v. From (8), the solution to (6) is given by

$$\lambda = \sum_{i=1}^{M} \left(\frac{u_i^* b}{\sigma_i} \right) v_i. \tag{9}$$

The above procedure also applies to the case when $M = N$, i.e., (6) is a square matrix system. Once the values of λ are determined, the numerical solution to the Cauchy problem (2)–(3) is given by (4).

4. METHODOLOGY FOR NOISY DATA

In practical application, the data are observed as noisy values $\tilde{\varphi}$ and $\tilde{\psi}$ that satisfy

$$|\tilde{\varphi}(P) - \varphi(P)| \leq \rho \quad \text{and} \quad |\tilde{\psi}(P) - \psi(P)| \leq \rho, \quad \text{for all} \quad P \in \Gamma, \tag{10}$$

instead of the exact values given in (3). Here, ρ denotes the noise level of the measurement. Throughout the paper, we reserve the notation '$\sim$' to indicate the presence of noise. Since the Cauchy problem is ill-posed, most standard numerical methods fail to produce meaningful solution in solving the directly Equation (6) in the presence of noise. It is necessary to employ some techniques to *stabilize* the numerical solution.

Let $\tilde{b}$ be the observed data and b be the exact data in (7). From (10), we have the noise vector $\varepsilon := \tilde{b} - b$ and $\|\varepsilon\| \leq \rho$. The resulting error in the numerical solution due to the presence of noise is

$$\tilde{\lambda} - \lambda = \sum_{i=1}^{M} \left(\frac{u_i^* \varepsilon}{\sigma_i} \right) v_i. \tag{11}$$

From (11), we can see that the error corresponding to the small eigenvalues will be magnified—a solution to this problem is using the regularization method having the following general form:

$$\tilde{\lambda}_\alpha = \sum_{i=1}^{M} f_i \left(\frac{u_i^* \varepsilon}{\sigma_i} \right) v_i, \tag{12}$$

where $f_i = f_i(\alpha)$ is referred as the *filter factors* that takes value 0 and 1 when $\alpha \to 0$ and $\alpha \to \infty$, respectively. Regularization methods always include a so-called *regularization parameter* $\alpha \in (0, \infty)$ which controls the degree of approximation. Let $\tilde{\lambda}_\alpha$ be the regularized solution to the noisy Cauchy problem with regularization parameter α. The idea is to decompose (11) into

$$\tilde{\lambda}_\alpha - \lambda = \underbrace{\sum_{i=1}^{M} (f_i - 1) \frac{u_i^* b}{\sigma_i} v_i}_{(a)} + \underbrace{\sum_{i=1}^{M} f_i \frac{u_i^* \varepsilon}{\sigma_i} v_i}_{(b)}. \tag{13}$$

The first term (13a) is the error introduced by the regularization method whereas the second error term (13b) is due to the noise in the given boundary data. The filter factors of the regularization methods we investigated are listed in Table 2. When $\alpha \to \infty$ or $f_i = 1$ for all i, (13) becomes (11) and no regularization is needed. Conversely, a full regularization will minimize (13b) but maximize (13a) in the same time. The determination of a suitable value of the regularization parameter α is crucial and is still under intensive research. References for these methods can be found in [24] for TR, [32] for DSVD and [33] for TSVD.

Table 2. Filter factors of the regularization methods.

Regularization method	Filter factors
Tikhonov Regularization (TR)	$f_i = \frac{\sigma_i^2}{(\sigma_i^2 + \alpha^2)}$
Damped SVD (DSVD)	$f_i = \frac{\sigma_i}{(\sigma_i + \alpha)}$
Truncated SVD (TSVD)	$f_i = \begin{cases} 1 & \text{if } \sigma_i \geq \alpha \\ 0 & \text{otherwise} \end{cases}$

The TSVD methods are a discrete regularization method. The condition $\sigma_i \geq \alpha$ in Table 2 can be replaced by $i \geq m$ for some positive integer $m = m(\alpha)$, $1 \leq m \leq M$. The integer m is often referred as the *discrete regularization parameter*. The continuous dependence on (discrete) regularization parameter plays an important role in many methods for choosing the 'optimal' parameter. In this case, discrete computed values are interpolated by spline curve and the optimal regularization parameter is the point closest to the continuous solution.

Four experimental methods for choosing the regularization parameters are:

Discrepancy principle (DP)

Based on some a priori estimate on the error level ρ in (10), the DP method [34] gives the optimal regularization parameter such that

$$\left\| A\,\tilde{\lambda}_\alpha - \tilde{b} \right\| = \rho.$$

An overestimate or underestimate of ρ will result in an overregularized or underregularized solution. The following three methods do not require any knowledge of ρ but rely on information at the right-hand vector $\tilde{b}$.

L-curve criterion (LC)

This method was firstly introduced by Lawson and Hanson [35]. Hansen and O'Leary [20] investigated the properties of regularized systems under different values of the regularization parameter α and defined the L-*curve* by

$$\left(\log\left(\left\|\,\tilde{\lambda}_\alpha\right\|^2\right),\ \log\left(\left\|A\,\tilde{\lambda}_\alpha - \tilde{b}\right\|^2\right)\right) \quad \text{for} \quad \alpha > 0. \tag{14}$$

An optimal regularization parameter is then corresponding to the 'corner' of the L-curve. For computational purpose, the point with maximum curvature will be chosen.

Generalized cross-validation (GCV)

The GCV method [36] determines the optimal regularization parameter by minimizing the following continuous GCV function

$$G(\alpha) = \frac{\left\|A\,\tilde{\lambda}_\alpha - \tilde{b}\right\|^2}{(\text{trace}\,(I_N - AA_\alpha^\oplus))^2} \quad \text{for} \quad \alpha > 0, \tag{15}$$

where $A_\alpha^\oplus$ is the *generalized* pseudoinverse for the employed regularization methods, i.e., the M-by-N matrix that produces the regularized solution $\lambda_\alpha = A_\alpha^\oplus \tilde{b}$.

Quasi-optimality criterion (QO)

Similar to GCV, the QO method [34] minimizes the following function to obtain the optimal regularization parameter:

$$Q(\alpha) = \alpha \left\| \frac{d}{d\alpha} \tilde{\lambda}_\alpha \right\| = \left(\sum_{i=1}^{M} \left(f_i(1 - f_i) \frac{u_i^* \tilde{b}}{\sigma_i} \right)^2 \right)^{1/2} \quad \text{for} \quad \alpha > 0. \qquad (16)$$

ACKNOWLEDGEMENT

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 1185/03E).

REFERENCES

1. A. Bogomolny (1985), Fundamental solutions method for elliptic boundary value problems. *SIAM Journal on Numerical Analysis*, 22, 4, pp. 644–669.
2. R. Mathon and R.L. Johnston (1977), The approximate solution of elliptic boundary-value problems by fundamental solutions. *SIAM Journal on Numerical Analysis*, 14, 4, pp. 638–650.
3. Y.-S. Smyrlis and A. Karageorghis (2003), Some aspects of the method of fundamental solutions for certain biharmonic problems. *Computer Modeling in Engineering and Sciences*, 4, 5, pp. 535–550.
4. A. Karageorghis and G. Fairweather (1987), The method of fundamental solutions for the numerical solution of the biharmonic equation. *Journal of Computational Physics*, 69, pp. 433–459.
5. A. Poullikkas, A. Karageorghis and G. Georgiou (2002), The method of fundamental solutions for three-dimensional elastostatics problems.*Computers and Structures*, 80, 3–4, pp. 365–370.
6. P.S. Kondapalli, D.J. Shippy and G. Fairweather (1992), Analysis of acoustic scattering in fluids and solids by the method of fundamental solutions. *The Journal of the Acoustical Society of America*, 91, 4 Pt 1, 1844–1854.
7. P.S. Kondapalli, D.J. Shippy and G. Fairweather (1992), The method of fundamental solutions for transmission and scattering of elastic waves. *Computer Methods in Applied Mechanics and Engineering*, 96, 2, pp. 255–269.
8. K. Balakrishnan and P.A. Ramachandran (2000), The method of fundamental solutions for linear diffusion-reaction equations. *Mathematical and Computer Modelling*, 31, 2–3, pp. 221–237.
9. K. Balakrishnan and P.A. Ramachandran (2001), Osculatory interpolation in the method of fundamental solution for nonlinear passion problems. *Journal of Computational Physics*, 172, 1, pp. 1–18.
10. M.A. Golberg (1995), The method of fundamental solutions for Poisson's equation. *Engineering Analysis with Boundary Elements*, 16, 3, 205–213.

11. P. W. Partridge and B. Sensale (2000), The method of fundamental solutions with dual reciprocity for diffusion and diffusion-convection using subdomains. *Engineering Analysis with Boundary Elements*, 24, 9, pp. 633–641.
12. G. Fairweather and A. Karageorghis (1998), The method of fundamental solutions for elliptic boundary value problems. *Advances in Computational Mathematics*, 9, 1–2, pp. 69–95.
13. M.A. Golberg and C.S. Chen (1999), The method of fundamental solutions for potential, Helmholtz and diffusion problems. In: *Boundary Integral Methods: Numerical and Mathematical Aspects, Volume 1 of Computational Engineering.* WIT Press/Comput. Mech. Publ., Boston, MA, pp. 103–176.
14. G. Alessandrini (1993), Stable determination of a crack from boundary measurements. *Proceedings of the Royal Society of Edinburgh. Section A*, 123, 3, pp. 497–516.
15. N.M. Al-Najem, A.M. Osman, M.M. El-Refaee and K.M. Khanafer (1998), Two dimensional steady-state inverse heat conduction problems. *International Communications in Heat and Mass Transfer*, 25, 4, pp. 541–550.
16. P. Colli Franzone and E. Magenes (1979), On the inverse potential problem of electrocardiology. *Calcolo*, 16, 4, pp. 459–538.
17. D.D. Ang, N.H. Nghia and N.C. Tam (1998), Regularized solutions of Cauchy problem for the Laplace equation in an irregular layer: a three dimensional case. *Acta Mathematica Vietnamica*, 23, pp. 65–74.
18. F. Berntsson and L. Eldén (2001), Numerical solution of a Cauchy problem for the Laplace equation. *Inverse Problems*, 17, 4, pp. 839–853 [Special issue to celebrate Pierre Sabatier's 65th birthday (Montpellier, 2000)].
19. J. Cheng, Y.C. Hon, T. Wei and M. Yamamoto (2001), Numerical computation of a Cauchy problem for Laplace's equation. *Zeitschrift fuer Angewandte Mathematik und Mechanik*, 81, 10, pp. 665–674.
20. D.N. Hào and D. Lesnic (2000), The Cauchy problem for Laplace's equation via the conjugate gradient method. *IMA Journal of Applied Mathematics*, 65, 2, pp. 199–217.
21. Y.C. Hon and T. Wei (2001), Backus-Gilbert algorithm for the Cauchy problem of the Laplace equation. *Inverse Problems*, 17, 2, pp. 261–271.
22. Y.C. Hon and Z. Wu (2000), A numerical computation for inverse boundary determination problem. *Engineering Analysis with Boundary Elements*, 24, 7–8, pp. 599–606.
23. H.-J. Reinhardt, H. Han and D.N. Hào (1999), Stability and regularization of a discrete approximation to the Cauchy problem for Laplace's equation. *SIAM Journal on Numerical Analysis*, 36, 3, pp. 890–905 (electronic).
24. A.N. Tikhonov and V.Y. Arsenin (1977), *Solutions of Ill-Posed Problems*. V. H. Winston & Sons, Washington, DC and John Wiley & Sons, New York (Translated from the Russian, Preface by translation editor Fritz John, Scripta Series in Mathematics).
25. J. Cheng and M. Yamamoto (1998), Unique continuation on a line for harmonic functions. *Inverse Problems*, 14, 4, pp. 869–882.
26. L.E. Payne (1960), Bounds in the Cauchy problem for the Laplace equation. *Archive for Rational Mechanics and Analysis*, 5, pp. 35–45.
27. Prem K. Kythe (1996), *Fundamental Solutions for Differential Operators and Applications*. Birkhäuser Boston Inc., Boston, MA.
28. M. Katsurada and H. Okamoto (1996), The collocation points of the fundamental solution method for the potential problem. *Computers and Mathematics with Applications*, 31, 1, pp. 123–137.
29. J. Li, Y.C. Hon and C.S. Chen (2002), Numerical comparisons of two meshless methods using radial basis functions. *Engineering Analysis with Boundary Elements*, 26, pp. 205–225.

30. R.S.C. Cheng (1987), Delta-Trigonometric and Spline Methods using the Single-Layer Potential Representation. PhD Thesis, University of Maryland.
31. P.A. Ramachandran (2002), Method of fundamental solutions: singular value decomposition analysis. *Communications in Numerical Methods in Engineering*, 18, 11, pp. 789–801.
32. M.P. Eckstrom and R.L. Rhodes (1974), On the application of eigenvector expansions to numerical deconvolution. *Journal of Computational Physics*, 14, pp. 319–340.
33. P.C. Hansen (1987), The truncated SVD as a method for regularization. *BIT*, 27, 4, pp. 534–553.
34. V.A. Morozov (1984), *Methods for Solving Incorrectly Posed Problems*. Springer-Verlag, New York (Translated from the Russian by A.B. Aries, Translation edited by Z. Nashed).
35. C.L. Lawson and R.J. Hanson (1994), *Solving Least Squares Problems*. Prentice-Hall Series in Automatic Computation, Prentice-Hall Inc., Englewood Cliffs, NJ.
36. G. Wahba (1990), *Spline Models for Observational Data*, *volume 59 of CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

NUMERICAL DETERMINATION OF THE RESONANCE FREQUENCIES AND EIGENMODES USING THE MFS

Carlos J.S. Alves and Pedro R.S. Antunes
CEMAT & Departamento de Matemática, Instituto Superior Técnico, Lisboa, Portugal

Abstract In this work we present a numerical algorithm for the determination of the eigenvalues and eigenfunctions associated to the Dirichlet problem for the Laplacian, in a bounded or in an exterior domain. The determination of higher eigenfrequencies is a well-known numerical problem, that has been addressed with other numerical methods. Here we propose to use the method of fundamental solutions. Since the MFS produces highly ill conditioned matrices, a particular technique was derived to overcome the difficulty of determining accurately those eigenfrequencies. Extensive numerical simulations will be presented.

Keywords: eigenfrequencies, resonance, acoustic waves, method of fundamental solutions.

1. INTRODUCTION

The determination of the eigenvalues and eigenfunctions associated to the Laplace operator in a bounded domain Ω is a well-known problem with applications in acoustics. For simple shapes, such as rectangles or circles in 2D, this leads to straightforward computations, without the need of a numerical algorithm. However, when the shape is non-trivial, that computation requires the use of a numerical method for PDEs. A standard finite differences method can produce good results when dealing with a particular type of shapes defined on rectangular grids, while for other type of shapes the finite element method or the boundary element method are more appropriated [1]. These classical methods require extra computational effort; in one case, the construction of the mesh and the associated rigid matrix, and in the other, the integration of weakly singular kernels. Here we propose to use the method of fundamental solutions (MFS). The MFS has been mainly applied to boundary problems in PDEs, starting in

G. R. Liu et al. (eds.), Computational Methods, 1385–1389.

the 1960s [2]. The application of the MFS to the calculation of the eigenvalues has been introduced by Karageorghis [3], and applied for simple shapes. In that work it is presented a comparison with the boundary element method used by Mey [1], and the results obtained for simple shapes (circles, squares), show a better performance for the MFS. The application of other meshless methods to the determination of eigenfunctions and eigenmodes has also been subject to recent research, mainly using radial basis functions [4].

In this work we consider the application of the MFS to general shapes. In that case the choice of the source points in the MFS becames more important to retrieve higher eigenfrequencies. We are able to obtain good results with a particular algorithm associating the source points to the shape. Having determined an approximation of the eigenvalue, we apply an algorithm based on the MFS to obtain the associated eigenmodes.

2. DETERMINATION OF EIGENFREQUENCIES

Let $\Omega \subset \mathbb{R}^2$ be a bounded connected domain with regular boundary $\partial\Omega$. For simplicity we will consider the 2D-Dirichlet eigenvalue problem for the Laplace operator. This is equivalent to obtain the resonance frequencies κ that verify the Helmholtz equation

$$\begin{cases} \Delta u + \kappa^2 u = 0 & in\ \Omega, \\ u = 0 & on\ \partial\Omega, \end{cases} \tag{1}$$

for a non-null function u. As an application, this corresponds to recover the resonance frequencies $\kappa > 0$ associated with a particular shape of a drum Ω. A fundamental solution Φ_κ of the Helmholtz equation verifies $(\Delta + \kappa^2)\Phi = -\delta$, where δ is the Dirac delta distribution. In the 2D case, we take $\Phi_\kappa(x) = \frac{i}{4}H_0^{(1)}(\kappa|x|)$ where $H_0^{(1)}$ is the first Hänkel function. A density result in [5] allows to justify the approximation of a $L^2(\partial\Omega)$ function using a sequence of functions

$$u_m(x) = \sum_{j=1}^{m} \alpha_{m,j}\Phi_\kappa(x - y_{m,j}) \tag{2}$$

that converges to $u|\Gamma$ in $L^2(\partial\Omega)$. Definining m collocation points $x_i \in \partial\Omega$ and m source points $y_{m,j} \in \hat{\Gamma}$, we obtain the system

$$\sum_{j=1}^{m} \alpha_{m,j}\Phi_\kappa(x_i - y_{m,j}) = 0 \quad (x_i \in \partial\Omega). \tag{3}$$

Therefore a straighforward procedure is to find the values κ for which the $m \times m$ matrix

$$A(\kappa) = [\Phi_\kappa(x_i - y_j)]\,[\alpha_j]_{m\times m} \tag{4}$$

has a null determinant. However, an arbitrary choice of source points may lead to worst results than the expected with the MFS applied to simple shapes. We will choose the points $x_1, \ldots., x_m \in \partial\Omega$ and $y_1, \ldots., y_m \in \hat{\Gamma}$ in a particular way. Given the m points x_i on $\partial\Omega$, we take m point sources

$$y_i = x_i + \tilde{\mathbf{n}}_i$$

where $\tilde{n}_i$ is approximately normal to the boundary $\partial\Omega$ on x_i. To obtain the vector $\tilde{n}_i$ we just consider $\tau^- = x_i - x_i - 1$, $\tau^+ = x_i - x_i + 1$ and calculate $\mathbf{n}^-$, $\mathbf{n}^+$ which are normal to τ^- and τ^+ (respectively) and pointing outwards Ω. Then we take $\tilde{n}_i = \frac{1}{2}(\mathbf{n}^- + \mathbf{n}^+)$. By some experimental criteria, we will usually take $|\tilde{n}_i| = \beta$. Source points taken too far from the boundary only presented better results in some particular cases, using simple shapes. The components of the matrix $A(\kappa)$ are complex numbers, so the determinant is also a complex number. We consider the real function $g(\kappa) = |Det[A(\kappa)]|$. It is clear that the function g will be very small in any case, since the MFS is highly ill conditioned and the determinant is quite small. To avoid machine precision problems the code was built in *Mathematica*. If κ is an eigenfrequency, κ is a point of minimum where $g(\kappa) = 0$ and therefore the derivative changes sign. We will make use of the rough approximation $g'(w) \approx \frac{g(w)-g(w-\varepsilon)}{\varepsilon}$ for a small $\varepsilon > 0$. To approximate the eigen-frequencies, where a clear change on the sign of the derivative g' is attained, we used the simple bissection method, which revealed to be quite accurate in the search of high eigenfrequencies.

3. DETERMINATION OF EIGENMODES

To obtain an eigenfunction associated with a certain resonance frequency ω we use a collocation method on $n + 1$ points, with $x_1, \ldots, x_n$ on $\partial\Omega$ and an extra point $x_{n+1} \in \Omega$. Then, the approximation of the eigenfunction is given by

$$\tilde{u}(x) = \sum_{k=1}^{n+1} \alpha_k \Phi_\omega(x - y_k). \tag{5}$$

To exclude the solution $\tilde{u}(x) \equiv 0$, the coefficients α_k are determinated by the resolution of the system $\tilde{u}(x_i) = \delta_{i,n+1}$, $i = 1, \ldots, n + 1$ where $\delta_{i,j}$ is the delta of Kronecker.

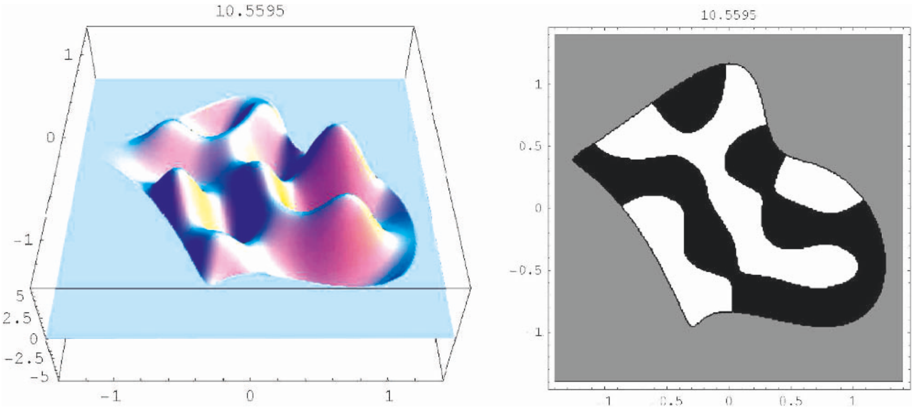


Figure 1. Plots of the 22th eigenmode and nodal domains associated to Ω.

4. SIMULATION WITH DIRICHLET BOUNDARY CONDITION

We consider a domain Ω with a non-trivial boundary given by the parametrization

$$t \longmapsto \left(\cos(t) - \frac{\cos(t)\sin(2t)}{2}, \sin(t) + \frac{\cos(4t)}{6} \right).$$

In Figure 1 we present the plot of the eigenfunction and nodal domains associated to the 22th eigenvalue.

5. CONCLUSIONS

In this brief account we presented the MFS method with an algorithm for the choice of source points. The single example for a particular situation of mixed Dirichlet/Neumann illustrates the good results already obtained for other type of boundary conditions and non simply connected domains.

ACKNOWLEDGEMENTS

This work was partially supported by FCT-POCTI/FEDER, project POCTI-MAT/34-735/00 and NATO-PST.CLG.980398.

REFERENCES

1. GDe Mey (1976), Calculation of the eigenvalues of the Helmholtz equation by an integral equation. International Journal of Numerical Method of Engineering 10, pp. 59–66.
2. VD, Kupradze MA, Aleksidze (1964), The method of functional equations for the approximate solution of certain boundary value problems. *USSR Computational Mathematics and Mathematical Physics*, 4, pp. 82–126.
3. Karageorghis A (2001), The method of fundamental solutions for the calculation of the eigenvalues of the helmholtz equation. Applied Mathematic, Letters, 14, 7, pp. 837–842.
4. J T, Chen M H, Chang K H, Chen and S R Lin (2002), Boundary collocation method with meshless concept for acoustic eigenanalysis of two-dimensional cavities using radial basis function. *Journal of Sound and Vibration*, 257, 4, pp. 667–711.
5. C.J.S. Alves (2000), Density results for the Helmholtz equation and the method of fundamental solutions. *In: Advances in Computational Engineering & Sciences*, Vol. I, S.N. Atluri and F.W. Brust (eds.). Technical Science Press, pp. 45–50.
6. A. Bogomolny (1985), Fundamental solutions method for elliptic boundary value problems. SIAM Journal of Numerical Analysis 22, 4, pp. 644–669.

SCATTERED NODE COMPACT FINITE DIFFERENCE-TYPE FORMULAS GENERATED FROM RADIAL BASIS FUNCTIONS

Grady Wright[1] and Bengt Fornberg[2]
[1]*Department of Mathematics, University of Utah, USA*
[2]*Department of Applied Mathematics, University of Colorado, USA*

Abstract In standard equispaced finite difference (FD) formulas, symmetries can make the order of accuracy relatively high compared to the number of nodes in the FD stencil. With *scattered* nodes, such symmetries are no longer available. Thus, the number of nodes in the stencils can be relatively large compared to the resulting accuracy. The generalization of compact FD (CFD) formulas that we propose for scattered nodes and radial basis functions (RBFs) achieves the goal of reducing the number of stencil nodes without a similar reduction in accuracy. We analyze the accuracy of these new compact RBF–FD formulas by applying them to some model problems, and study the effects of the shape parameter that arises in, for example, the multi-quadric radial function.

Keywords: radial basis functions, partial differential equations, compact finite difference method, mesh-free.

1. INTRODUCTION

An obvious approach for circumventing the geometric inflexibility of the standard finite difference method (FDM) for solving partial differential equations (PDEs) is to allow the nodes of the FD stencils to be placed freely, so that a good discretization of the physical domain of the problem can be obtained. However, this natural *mesh-free* idea raises questions of how the weights of the resulting *scattered* node FD formulas should be computed. It has recently (and what appears to be independently) been proposed by Shu et al. [1], Tolstykh et al. [2], Cecil et al. [3] and Wright [4] that RBF interpolants be used for computing these weights. We refer to this idea as the RBF–FD method. The following are some reasons for using RBFs: (1) for the appropriate choice

G. R. Liu et al. (eds.), Computational Methods, 1391–1395.

of radial function $\phi(r)$, the RBF interpolation method is well-posed in all dimensions (unlike polynomial interpolation); (2) RBF interpolants can be very accurate at approximating derivatives; and (3) certain types of radial functions $\phi(r)$ feature a 'shape' parameter ε that allows them to vary from being nearly flat ($\varepsilon \to 0$) to sharply peaked ($\varepsilon \to \infty$). The recent work of, for example, Fornberg et al. [5] shows that all classical FD formulas can be recovered by 'flat' RBF interpolants (when the nodes are arranged accordingly).

In standard equispaced finite difference (FD) formulas, symmetries can make the order of accuracy relatively high compared to the number of nodes in the FD stencil. With scattered nodes, such symmetries are no longer available. Thus, the number of nodes in the stencils can be relatively large compared to the resulting accuracy. To circumvent this problem, we propose a generalization of compact finite difference (CFD) formulas first introduced by Collatz [6]. The basic idea behind this method is to keep the stencil size fixed and to also include in the FD formula a linear combination of derivatives of u at surrounding nodes. In the case of 1-D and equispaced nodes, the weights for these CFD formulas are typically derived using Padé approximants. For scattered nodes in one and higher dimensions, and for RBFs, this Padé approach is no longer available. Instead, we propose a method based on Hermite RBF interpolation.

Without loss of generality, we limit the discussion to RBF–FD and RBF–CFD formulas for the d-dimensional Laplacian ∇^2.

2. HERMITE RBF INTERPOLATION

Since the RBF–CFD formulas are ultimately obtained from Hermite RBF interpolants, we review in this section a method for solving the Hermite interpolation problem. We note that the standard RBF–FD formulas are ultimately obtained from standard RBF interpolants, which turn out to be a special case of Hermite RBF interpolants.

Let σ be a vector containing some combination of $m \leq n$ distinct numbers from the set $\{1, \ldots, n\}$. Given a set of distinct data points $\underline{x}_i \in \mathbb{R}^d$, $i = 1, \ldots, n$, and corresponding (scalar) data values $u(\underline{x}_i)$, $i = 1, \ldots, n$, and $\nabla^2 u(\underline{x}_{\sigma_i})$, $i = 1, \ldots, m$, the Hermite RBF interpolation method we consider is to find an interpolant of the form

$$s(\underline{x}) = \sum_{i=1}^{n} \lambda_i \phi(||\underline{x} - \underline{x}_i||) + \sum_{j=1}^{m} \alpha_j \, \nabla^2 \, \phi(||\underline{x} - \underline{x}_{\sigma_j}||) + \beta. \tag{1}$$

Here ϕ(r) is some *radial* function and $||\cdot||$ is the standard Euclidean norm. This method is similar to the Hermite–Birkhoff method proposed by Wu [7]. Imposing the Hermite interpolation constraints and the additional constraint

$\Sigma_{i=1}^{n}\lambda_i = 0$ leads to the following symmetric linear system of equations (in block form)

$$\underbrace{\left[\begin{array}{c|c|c} \Phi & \nabla^2\Phi & 1 \\ \hline \nabla^2\Phi & \nabla^4\Phi & 0 \\ \hline 1 & 0 & 0 \end{array}\right]}_{A} \left[\begin{array}{c} \lambda \\ \hline \alpha \\ \hline \beta \end{array}\right] = \left[\begin{array}{c} u \\ \hline \nabla^2 u \\ \hline 0 \end{array}\right]. \tag{2}$$

For the appropriate choice of ϕ, A is guaranteed to be non-singular [8]. Note that for $m = 0$, the Hermite problem reduces to the standard RBF interpolation problem.

In this study, we focus on the multi-quadric (MQ) radial function, $\phi(\mathrm{r}) = \sqrt{1 + (\varepsilon r)^2}$, since it can produce very accurate interpolants, and it features a free shape parameter ε that can be adjusted to significantly improve the resulting accuracy of the interpolants. We note also that the linear system (2) is guaranteed to be non-singular for the MQ radial function (and ε non-zero).

3. RBF–FD FORMULATION

In this section we describe how to generate the RBF–FD and RBF–CFD formulas. Without loss of generality, we consider a stencil consisting of n (scattered) nodes $\underline{x}_1, \ldots, \underline{x}_n$ and are interested in approximating $\nabla^2 u(\underline{x}_1)$. Here we let $\nabla^2 u(\underline{x}_1) := \nabla^2 u(x)|$.

For RBF–FD formulas the goal is to find weights c_i such that, $\nabla^2 u(\underline{x}_1) \approx \Sigma_{i=1}^{n} c_i u(\underline{x}_i)$. This is accomplished by solving the linear system

$$A[c|\mu]^T = \Big[\, \underbrace{\nabla^2\phi(||\underline{x} - \underline{x}_1||) \cdots \nabla^2\phi(||\underline{x} - \underline{x}_n||)}_{B(\underline{x})} \mid 0 \,\Big]^T, \tag{3}$$

where A is the matrix in (2) (with $m = 0$) and μ is a dummy value related to β in (1).

For the RBF–CFD formulas the goal now is to increase the accuracy of the approximation without increasing the stencil size. We accomplish this by using nodes where u and $\nabla^2 u$ are given exactly. Let σ be a vector containing some combination of $0 < m < n$ distinct numbers from the set $\{2, \ldots, n\}$, then we seek to find weights c_i and $\tilde{c}_{\sigma_j}$ such that $\nabla^2 u(\underline{x}_1) - \Sigma_{j=1}^{m} \tilde{c}_{\sigma_j} \nabla^2 u(\underline{x}_{\sigma_j}) \approx \Sigma_{i=1}^{n} c_i u(\underline{x}_i)$. This is accomplished by solving the linear system

$$A[c|\tilde{c}|\mu]^T = \big[\, B(\underline{x}) \mid \nabla^4\phi(||\underline{x} - \underline{x}_{\sigma_1}||) \;\cdots\; \nabla^4\phi(||\underline{x} - \underline{x}_{\sigma_m}||) \mid 0\big]^T, \tag{4}$$

where A is the matrix in (2) and $B(\underline{x})$ is given in (3).

For small (in magnitude) values of the shape parameter ε, the linear systems (3) and (4) will be extremely ill-conditioned. To bypass this problem, we use the Contour-Padé algorithm [9], which allows for the stable computation of the RBF–FD and RBF–CFD weights for all $\varepsilon \geq 0$.

4. APPLICATION: POISSON'S EQUATION

To illustrate the improved accuracy of the RBF–CFD formulas, we apply them to the model problem

$$\nabla^2 u = f \quad \text{in } \Omega = \left\{(x, y) \mid x^2 + y^2 < 1\right\} \text{ and } u = g \quad \text{on } \partial\Omega. \tag{5}$$

Note that $\nabla^2 u$ is given analytically as f in the interior. For the experiment that follows, f and g are computed from the known solution

$$u(\underline{x}) = u(x, y) = \frac{25}{25 + (x - 0.2)^2 + 2y^2}\,. \tag{6}$$

The domain is discretized using the $N = 200$ points shown in Figure 1 (a). To measure the error (which depends on ε), we use $E(\varepsilon) = \max_{i=1,\ldots,N} |\bar{u}(\underline{x}_i, \varepsilon) - u(\underline{x}_i)|$ (i.e., the max norm), where $\bar{u}$ is the approximate solution. We call the ε where $E(\varepsilon)$ reaches a minimum the 'optimal' ε.

Figure 1 (b) contains the results using the $n = 9$ ($m = 0$) node RBF–FD formulas, and the RBF–CFD formulas using $n = 9$ $m = 5$, and $n = 10$ $m = 9$. Looking at the error for the standard $n = 9$ solution and the compact $n = 9$ $m = 5$ solution, we see that the accuracy is vastly improved. As we should expect, the accuracy can be further improved by increasing n and m, as illustrated

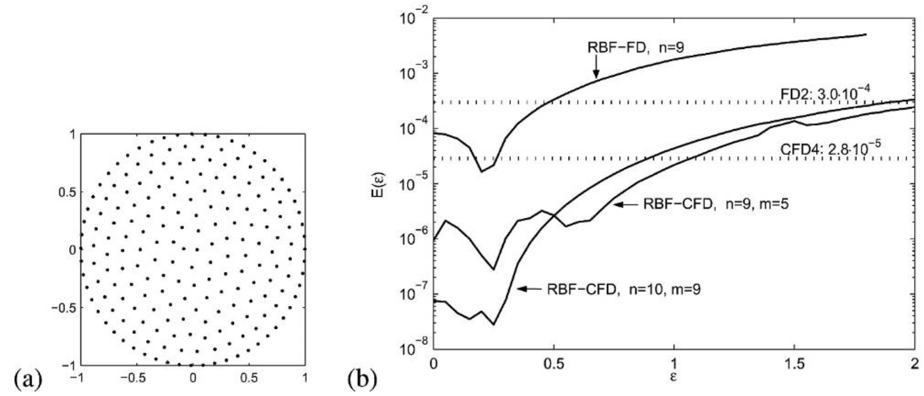


Figure 1. (a) 200 point unstructured discretization of the unit disk. (b) The error as a function of ε for various numerical solutions of (5).

by the $n = 10$ $m = 9$ solution. For this example, however, any improvements appear to be lost for ε approximately > 0.5. The figure also illustrates that the optimal value of ε is small (in magnitude), and non-zero, as is typically the case for the RBF interpolation problem. Also included in the figure (see the dotted lines) are the results for the standard FD solutions based on a uniform polar mesh with approximately the same number of boundary and interior points as the unstructured mesh. FD2 marks the results for the standard 5-node second-order FD scheme, while CFD4 marks the results for the standard compact 9-node fourth-order FD scheme. Comparing the non-compact FD2 and $n = 9$ RBF–FD solution, we see that approximately for $\varepsilon < 0.48$, the RBF solution is clearly better. Comparing the CFD4 solution and the $n = 9$ $m = 5$ RBF–CFD solution, which both happen to use the same number of nodes and derivative values in their respective stencils, we see that the RBF solution is better for all values of ε approximately <1.05.

In all cases, the RBF–FD and RBF–CFD solutions to this problem were computed using successive over-relaxation (SOR). Once the optimal relaxation parameter was found, this iterative method turned out to be quite computationally effective.

REFERENCES

1. C. Shu, H. Ding and K.S. Yeo (2003), Local radial basis function-based differential quadrature method and its application to solve two-dimensional incompressible Navier-Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 192, pp. 941–954.
2. A.I. Tolstykh, M.V. Lipayskii and D.A. Shirobokov (2003), High-accuracy discretization methods for solid mechanics. *Archives of Mechanics*, 55, pp. 531–553.
3. T. Cecil, J. Qian and S. Osher (2004), Numerical methods for high dimensional Hamilton-Jacobi equations using radial basis functions. *Journal of Computational Physis*, 196, pp. 327–347.
4. G. Wright (2003), *Radial Basis Function Interpolation: Numerical and Analytical Developments*. Ph. D. thesis, University of Colorado, Boulder.
5. B. Fornberg, G. Wright and E. Larsson (2002), Some observations regarding interpolants in the limit of flat radial basis functions. *Computers & Mathematics with Applications*, 47, pp. 37–55.
6. L. Collatz (1960), *The Numerical Treatment of Differential Equations*. Springer-Verlag: Berlin.
7. Z. Wu (1992), Hermite-Birkhoff interpolation of scattered data by radial basis functions. *Approximation Theory and its Applications*, 8, 2, pp. 1–10.
8. X. Sun (1994), Scattered Hermite interpolation using radial basis functions. *Linear Algebra and its Applications*, 207, pp. 135–146.
9. B. Fornberg, and G. Wright (2004), Stable computation of multiquadric interpolants for all values of the shape parameter. *Computers & Mathematics with Applications*, 48, pp. 853–867.

EXPLOSION SIMULATION BY SMOOTHED PARTICLE HYDRODYNAMICS

Kobashi Wataru and Matsuo Akiko
Department of Mechanical Engineering, Keio University, Yokohama 223-8522, Japan

Abstract The Smoothed Particle Hydrodynamics (SPH) method is used to simulate the dynamics of the underwater explosion (UNDEX) surrounded by an iron wall. One-dimensional vertical elastic oscillation of an iron stick and oscillation of an explosive gas caused by UNDEX surrounded by rigid walls are examined to verify the solid model and the UNDEX model focusing on the interaction between the explosive gas and the surrounding water, and the results agree well with the theoretical value and the previous study. In the simulation result, the shock wave impacts the iron wall, and the wave propagates in the iron wall and reflects at the interface. The SPH method qualitatively reproduces the dynamics of UNDEX surrounded by the iron wall. A potential of the SPH method for a prediction of the explosion damage is demonstrated in the present study.

Keywords: smoothed particle hydrodynamics, particle method, underwater explosion, underwater shock wave, structural analysis, coupled problem.

1. INTRODUCTION

The SPH method has the following nice features; meshfree, Lagrangian and particle method. These features allow us to treat heterogeneous states; detonation gas, water and solid. The SPH method has been used for a simulation of UNDEX [1] which mainly focused on the interaction between the explosive gas and the surrounding water, or the high velocity impact (HVI) [2] UNDEX is now used in broad fields: medical, industrial and military field. For the purpose of these applications of UNDEX, it is important to predict the damage of a structure caused by UNDEX. The objective of the present study is to indicate a potential of the SPH method for a prediction of the explosion damage. For verification of the accuracy into the simulation result, one-dimensional vertical elastic oscillation of an iron stick and oscillation of an explosive gas

G. R. Liu et al. (eds.), Computational Methods, 1397–1403.

caused by UNDEX surrounded by rigid walls are examined in advance of main topic.

2. GOVERNING EQUATIONS

Following inviscid compressible equations reproduce UNDEX phenomenon.

$$\frac{D\rho}{Dt} = -\rho\nabla\cdot\mathbf{v}, \quad \frac{D\mathbf{v}}{Dt} = \frac{1}{\rho}\Delta\sigma,$$
$$\frac{De}{Dt} = \frac{1}{\rho}\sigma : \nabla\otimes\mathbf{v}, \quad P = P(\rho, e) \tag{1}$$

ρ, e, σ, $\mathbf{v}$ and t represent density, internal energy, stress tensor, velocity and time respectively. The stress tensor appearing in Equations (1) is defined in terms of an isotropic part which is the pressure P and the deviatoric stress S($\sigma = -P\mathbf{I} + \mathbf{S}$). **I** represents an unit tensor. The following equation is a constitutive equation of elastic material.

$$\frac{D\mathbf{S}}{Dt} = 2\mu\left(\dot{\varepsilon} - \frac{1}{3}Tr(\dot{\varepsilon})\mathbf{I}\right) + \mathbf{w}\cdot\mathbf{S} - \mathbf{S}\cdot\mathbf{w} \tag{2}$$

$\mathbf{S}$, $\dot{\varepsilon}$, $\mathbf{w}$ and μ represent deviatoric stress tensor, strain rate tensor, rotation tensor and share module respectively. The plastic flow regime is determined by the von Mieses criterion.

3. EQUATION OF STATE

Each equation of states for detonation gas, water and iron are the JWL equation [3], the Mie–Gruneisen equation for water [4] and the Mie–Gruneisen equation for solid [2], respectively. See references for each parameter of these equations.

The JWL equation:

$$P = A\left(1 - \frac{\omega\eta}{R_1}\right)e^{-\frac{R_1}{\eta}} + B\left(1 - \frac{\omega\eta}{R_2}\right)e^{-\frac{R_2}{\eta}} + \omega\eta\rho_0 e \tag{3}$$

The Mie–Gruneisen equation for water:

$$\begin{cases} P = a_1\mu + a_2\mu^2 + a_3\mu^3 + (b_0 + b_1\mu + b_1\mu^2)\rho_0 e, & \mu > 0 \\ P = a_1\mu + (b_0 + b_1\mu)\rho_0 e, & \mu < 0 \end{cases} \tag{4}$$

The Mie–Gruneisen equation for solid:

$$P = \left(1 - \frac{1}{2}\Gamma\mu\right)P_H(\rho) + \Gamma\rho e,$$

$$P_H = \begin{cases} a_0\mu + b_0\mu^2 + c_0\mu^3, & \mu > 0 \\ a_0\mu, & \mu < 0 \end{cases} \tag{5}$$

4. NUMERICAL VERIFICATIONS

An iron stick supposed elastic material impacts into the rigid wall at 100 m/s as shown in Figure 1. Initially, 2525 particles located as 25 in x direction and 101 in y direction are distributed at equal intervals. The impact causes vertical elastic oscillation of the iron stick. Figure 2 shows a displacement of a particle located at $x = 0.0$, $y = 0.5$. The exact solution is obtained by following equation.

$$dx(x,t) = \frac{8l_x v}{c\pi^2}\sum_{m=1}^{\infty}\frac{1}{(2m+1)^2}\,sin(2m+1)\frac{\pi ct}{2l_x}\,sin(2m+1)\frac{\pi x}{2l_x} \tag{6}$$

$c = \sqrt{E/\rho}$, $l_x = 0.24$, $l_y = 1.0$, $x = 0$. In Figure 2 the result of the large artificial viscosity (blue line) shows good agreement with the exact solution about the oscillation and the phase until 0.13 ms. However the oscillation peak

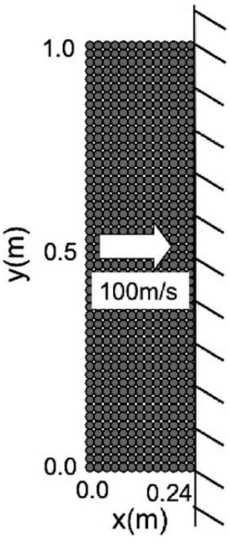


Figure 1. Initial distribution.

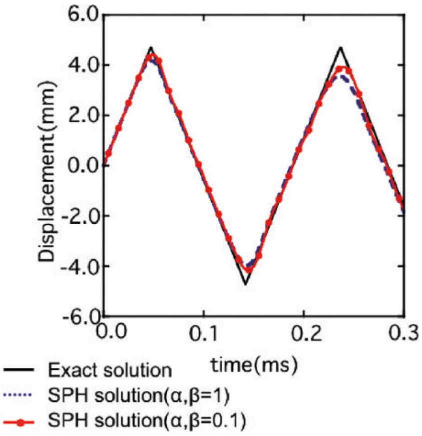


Figure 2. Displacement of a particle at $x = 0.0$, $y = 0.5$.

becomes small at 0.13 ms and 0.23 ms. These errors are caused by artificial viscosity because the result of the smaller artificial viscosity (red line) is closer to the exact solution.

A square shaped TNT charge (0.1 m × 0.1 m) explodes as a high-density gas in water surrounded by rigid walls in two-dimensional space as shown in Figure 3. Initially, 1,0201 particles located as 101 in each direction are distributed at equal intervals. Gas particles are 121 (11 × 11) at the center in the computational domain and the other particles are used for water. In this examination, the strong shock wave propagates through water and is reflected from the rigid walls. The gas bubble expands within the surrounding water and is compressed by a reflection shock wave from the rigid walls. Figure 4 shows a radial oscillation of an explosive gas bubble. From this figure the SPH solution shows good agreement with the Dytran solution [1].

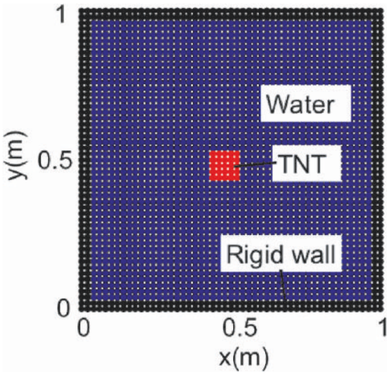


Figure 3. Initial distribution.

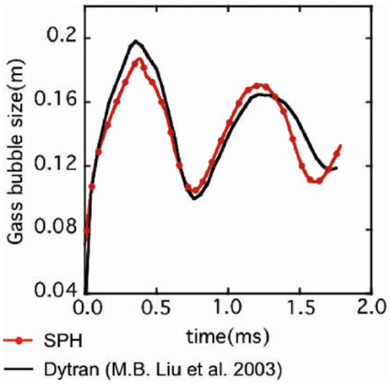


Figure 4. Explosive gas bubble evolution.

5. DYNAMIC SIMULATION THAT THE SHOCK WAVE CAUSED BY UNDEX IMPACTS AN IRON WALL

A square shaped TNT charge (0.05 m × 0.05 m) explodes as a high-density gas in water near an iron wall as shown in Figure 5. Initially, 1,0201 particles located as 101 in each direction are distributed at equal intervals. Gas particles are 36(6 × 6) at the center in the computational domain, and iron particles are 2525(101 × 25) at left side, and the other particles are used for water. The initial conditions are listed in Table 1. Figure 6 shows the pressure distribution of shock waves in water and the iron wall caused by the impact of UNDEX. An initial shock wave propagating outwards, a reflection wave from the iron wall, an elastic–plastic wave in the iron wall and a TNT/water interface can all be seen from this figure. With the shock wave propagating through water, the

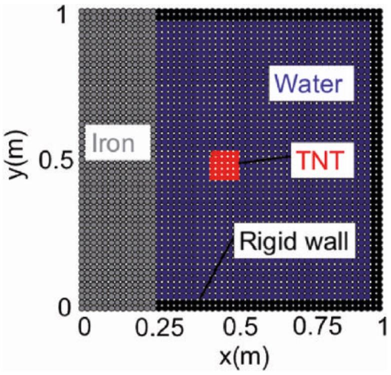


Figure 5. Initial distribution.

Table 1. Initial condition.

	TNT (Gas)	Iron (Solid)	Water (Liquid)
ρ_0(kg/m^3)	1630	7850	1000
P_0(GPa)	8.38	1×10^{-4}	1×10^{-4}
E_0(MJ/kg)	4.29	0.25	0.25

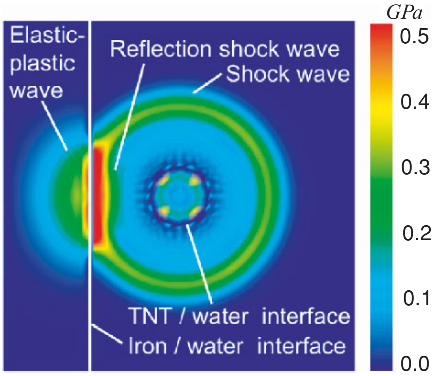


Figure 6. Pressure distribution at 0.14 ms.

gas bubble expands within the surrounding water. The shock wave reaches the iron wall at 0.1 ms, and is reflected from it. At the same time the elastic–plastic wave propagates in the iron wall at higher speed than the underwater shock wave. Figure 7 shows x–t diagram of the pressure distribution at the center line ($y = 0.5$) of the computational domain. It shows that a compression wave propagating within the iron wall towards the left end is caused by the impact of UNDEX, and reflects to become a tensile wave at 0.15 ms. The present

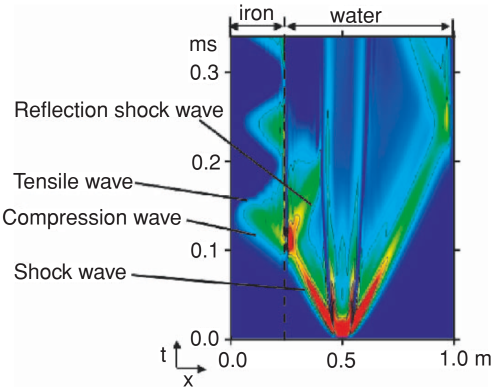


Figure 7. x–t diagram of pressure distribution.

simulation result indicates a potential of the SPH method for a prediction of the explosion damage.

6. CONCLUSIONS

A series of simulation were carried out to clarify the ability of the SPH method for the dynamics that the shock wave caused by UNDEX impacted the iron wall. Two kinds of examinations were performed to confirm the reliability of the simulation; one is one-dimensional vertical elastic oscillation of an iron stick, and the other is oscillation of an explosive gas caused by UNDEX surrounded by rigid walls. The examined results agreed well with the theoretical value and the previous study. In the simulation result, the pressure distribution in the computational domain showed that the strong shock wave caused by UNDEX impacts the iron wall to lead the elastic–plastic wave in the iron wall and the reflection wave in water respectively. The *x*–*t* diagram showed that the compression and tensile wave propagates one after the other in the iron wall. In conclusion, we were confident that the SPH method was one of the leading methods in dynamic explosion simulation.

ACKNOWLEDGEMENT

We are grateful to Prof. H. Noguchi of Keio University for helpful suggestions and comments.

REFERENCES

1. M.B. Liu, G.R. Liu, K.Y. Lam and Z. Zong (2003), Smoothed particle hydrodynamics for numerical simulation of underwater explosion. *Computational mechanics*, 30, pp. 106–118.
2. L.D. Libersky, A.G. Petschek, T.C. Carney, J.R. Hipp and F.A. Allahdadi (1993), High strain Lagrangian hydrodynamics. *Journal of Computational Physics*, 109, pp. 67–75.
3. B.M. Dobratz (1981), *LLNL Explosive Handbook* (UCRL-52997). Lawrence Livermore National Laboratory, Livermore, CA.
4. Y.S. Shin, M. Lee, K.Y. Lam and K.S. Yeo (1998), Modeling mitigation effect of water shield on shock waves. *Shock and Vibration*, 5, pp. 225–234.

CLASSIC TAYLOR-BAR IMPACT TEST REVISITED USING 3D SPH

C.E. Zhou, G.R. Liu and X. Han
Centre for Advanced Computations in Engineering Science, Department of Mechanical Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260

Abstract This paper presents an investigation of effects of various parameters used in 3D SPH simulation of classical Taylor-bar test. The investigation focuses on density evolution approaches, material strength models, smoothing length adaptation schemes and sensitivity of artificial viscosity constants. Based on the comparisons on the numerical results obtained, some conclusions are made which are useful in future SPH applications to solid impact and penetration problems.

Keywords: SPH, impact, Taylor-bar test, numerical simulation, high velocity impact.

1. INTRODUCTION

The meshfree methods have attracted much attention in recent decades, due to their flexibility in solving some challenging problems, such as crack propagation. Among various meshfree methods, the SPH [1, 2] is a potential and promising alternative for many problems, such as high velocity impacts. The classic Taylor-bar impact test is revisited in this paper by 3D SPH calculation, in hopes of providing informative data for formulation options in future SPH.

2. SPH FORMULATIONS WITH VARIATIONS

The meshfree and Lagrangian nature makes the SPH a good candidate tool to solve the hypervelocity impact (HVI) problems for material with strength. In HVI situations, the solid materials behaviour like fluids, which is governed by the conservation laws of hydrodynamics. Following the standard procedure

G. R. Liu et al. (eds.), Computational Methods, 1405–1409.

[3], a set of formulation can be written as follows

$$\begin{cases} \dfrac{D\rho_i}{Dt} = \displaystyle\sum_{j=1}^{N} m_j \left(\mathbf{v}_i^{\beta} - \mathbf{v}_j^{\beta}\right) \frac{\partial W_{ij}}{\partial \mathbf{x}_i^{\beta}} \\ \dfrac{D\mathbf{v}_i^{\alpha}}{Dt} = \displaystyle\sum_{j=1}^{N} m_j \left(\frac{p_i}{\rho_i^2} + \frac{p_j}{\rho_j^2}\right) \frac{\partial W_{ij}}{\partial \mathbf{x}_i^{\beta}} - \sum_{j=1}^{N} m_j \left(\frac{\mathbf{S}_i^{\alpha\beta}}{\rho_i^2} + \frac{\mathbf{S}_j^{\alpha\beta}}{\rho_j^2}\right) \frac{\partial W_{ij}}{\partial \mathbf{x}_i^{\beta}} \\ \dfrac{De_i}{Dt} = \dfrac{1}{2} \displaystyle\sum_{j=1}^{N} m_j \left(\frac{p_i}{\rho_i^2} + \frac{p_j}{\rho_j^2}\right) \left(\mathbf{v}_i^{\beta} - \mathbf{v}_j^{\beta}\right) \frac{\partial W_{ij}}{\partial \mathbf{x}_i^{\beta}} + \frac{1}{\rho_i} \mathbf{S}_i^{\alpha\beta} \dot{\boldsymbol{\varepsilon}}_i^{\alpha\beta} \\ \dfrac{D\mathbf{x}_i^{\alpha}}{Dt} = \mathbf{v}_i^{\alpha} \end{cases} \tag{1}$$

where the summation convention is adopted. Dependent variables include the *scalar* density ρ, pressure p, specific internal energy e, velocity *vector* $\mathbf{v}^{\alpha}$, strain rate *tensor* $\dot{\boldsymbol{\varepsilon}}_i^{\alpha\beta}$ and traceless deviatoric stress *tensor* $\mathbf{S}_i^{\alpha\beta}$, while the spatial coordinates x and time t are the independent variables. The equations are revolved as time steps forward in the moving Lagrangian frame.

It is worth noticing that the calculation of work done by the traceless deviatoric stress in the energy revolution equation is only valid in the elastic range. However, the plastic yielding can be dominant in an HVI problem. Therefore, one must calculate incremental plastic work during every time step and incorporate it into the energy equation. For this calculation, a yield criterion, such as von Mises, is necessary. A simple way to calculate this plastic work is first to estimate the effective plastic strain increment $\Delta\varepsilon^{p}_{eff}$ as:

$$\Delta\varepsilon^{p}_{eff} = (\sigma^{*}_{eff} - Y_0)/3G \tag{2}$$

where σ^{*}_{eff}, Y_0 and G are the provisional von Mises flow stress, yield strength and shear modulus, respectively. One can then estimate the incremental plastic work during the current time step n using

$$\Delta W_p^n = \frac{1}{2}\left(\sigma^{n+1}_{eff} + \sigma^{n}_{eff}\right) \Delta\varepsilon^{p}_{eff} \left(m/\rho^{n+1/2}\right) \tag{3}$$

where σ^{n+1}_{eff} are the effective stress calculated from deviatoric stress $\mathbf{S}_i^{\alpha\beta}$.

In the course of solution, an appropriate material strength model, such as Johnson–Cook model [4], is of importance in the SPH application to solid mechanics. Also, an Equation of State (EOS), such as Mie-Gruneisen EOS [5], is required to establish the relationships between temperature, volume and pressure for a given substance. In addition, the Radial Return Method [6] can be used to scale back the deviatoric stress to yield surface.

However, an SPH simulation can be performed in many different settings. In this paper, some variations are examined to reveal some of the important effects. The Johnson–Cook yield model and the perfectly plastic yield model are both tested; the density is evolved with using either summation or continuity approach; the smoothing length evolution schemes include the one proposed by Benz [7] and Monaghan and Lattanzio [8]; and the sensitivity of the artificial viscosity parameters in the Monaghan type artificial viscosity is also investigated.

3. NUMERICAL INVESTIGATION

We consider a cylindrical bar, made of oxygen-free high-conductivity (OFHC) copper or Armco iron, impacts perpendicularly onto a flat rigid surface. The original length of cylinder was 2.54 cm, and the initial diameter was 0.76 cm. The cylinder was initially, travelling at a speed of 221 m/s, in contact with the rigid surface, while the time at conclusion of the test was set to be 90 μs since the kinetic energy was almost dissipated by then. All material constitutive constants in the testing follow those given in Johnson and Holmquist [9]. Cases of simulation with different combinations of variations were detailed in Table 1.

Table 1. Case definition.

Case no.	Material	Density evolution	Yield model	Smoothing length	Artificial viscosity parameters	
					α	β
1	Armco iron	Summation	Johnson–Cook	Constant, 0.64 mm	1.0	1.0
2	Armco iron	Continuity	Johnson–Cook	Constant, 0.64 mm	1.0	1.0
3	Armco iron	Continuity	Perfectly plastic	Constant, 0.64 mm	1.0	1.0
4	OFHC	Continuity	Perfectly plastic	Constant, 0.64 mm	1.0	1.0
5	OFHC	Continuity	Johnson–Cook	Constant, 0.64 mm	1.0	1.0
6	OFHC	Continuity	Johnson–Cook	Variable, Benz	1.0	1.0
7	OFHC	Continuity	Johnson–Cook	Variable, d = 2.5[%]	1.0	1.0
8	OFHC	Continuity	Johnson–Cook	Variable, d = 2.0	1.0	1.0
9	OFHC	Continuity	Johnson–Cook	Variable, d = 1.5	1.0	1.0
10	OFHC	Continuity	Johnson–Cook	Variable, d = 0.8	1.0	1.0
11[#]	OFHC	Continuity	Johnson–Cook	Variable, d = 1.0	1.0	1.0
12	OFHC	Continuity	Johnson–Cook	Variable, d = 1.0	1.0	0.5
13	OFHC	Continuity	Johnson–Cook	Variable, d = 1.0	1.0	2.0
14	OFHC	Continuity	Johnson–Cook	Variable, d = 1.0	0.5	1.0
15	OFHC	Continuity	Johnson–Cook	Variable, d = 1.0	2.0	1.0
16	OFHC	Continuity	Johnson–Cook	Variable, d = 1.0	0.5	0.5
17	OFHC	Continuity	Johnson–Cook	Variable, d = 1.0	2.0	2.0

Note: Only the highlighted cells contain the variations taken during that comparison, i.e., other choices are identical. Variations with associated cases are indicated below: 1–2: Density evolution; 2–5: Yield model; 5–11: Smoothing length; 11–17: Artificial viscosity.

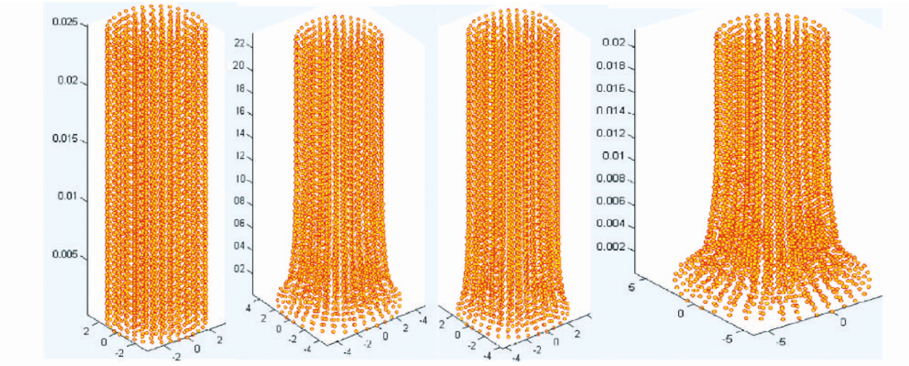
[#]: Standard case.

[%]: Smoothing length Monaghan-Lattanzio Scheme based on the particle density.

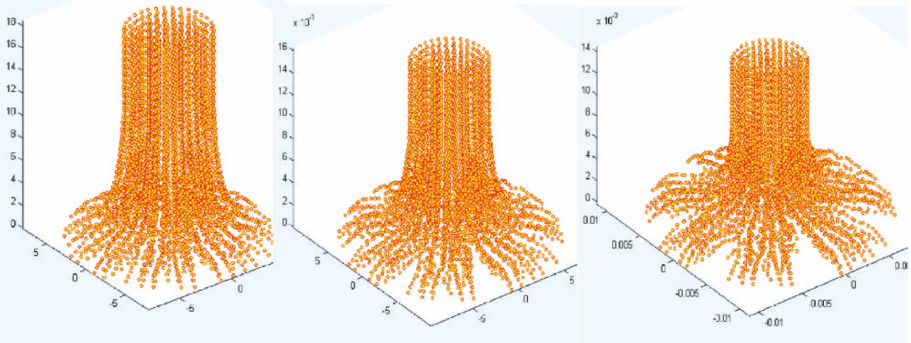
4. RESULTS AND CONCLUSIONS

3D rendering of the SPH calculated cylinders at some selected durations are given in Figure 1 for the standard case 11. The comparisons made conclusions obtained from all other cases are summarized below, in terms of the scheme variation proposed.

1. The plastic work can be dominant during energy evolution in the SPH application to solid impact problems which involves significant plastic yielding.
2. For solid problems with strong discontinuity, e.g., HVI, the continuity density approach is preferred since it yielded better ability to handle material boundaries, in terms of deformed cylinder shapes; and had a much better computational efficiency, in terms of CPU time logged.
3. While more physically based strength models are pursued, the empirical Johnson–Cook models yielded satisfactory results during this numerical investigation, compared to the simplified plastic yield model.



(a) Initial (b) 5 us (c) 10 us (d) 20 us



(e) 40 us (f) 60 us (g) 90 us

Figure 1. 3D rendering of deformed cylinders.

4. Benz's particle adaptation scheme is easier to implement but with a similar accuracy, compared to Monaghan type in which smoothing factor has a significant influence on both accuracy and efficiency of the final solution.
5. The numerical results are sensitive to the artificial bulk viscosity parameter (α) while the change of shear viscosity parameter (β) does not make much difference in the final shape of cylinders. The suggested values go to their recommendations of one.

REFERENCES

1. J.J. Monaghan (1988), An introduction to SPH. *Computer Physics Communications*, 48, pp. 89–96.
2. G.R. Liu (2002), *Meshfree Methods: Moving Beyond the Finite Element Method*. CRC Press, USA.
3. G.R. Liu and M.B. Liu (2003), *Smoothed Particle Hydrodynamics*. World Scientific, Singapore.
4. G.R Johnson and W.H. Cook (1983), A constitutive model and data for metals subjected to large strains, high strain rates and high temperatures. In: *Proceedings of the Seventh International Symposium on Ballistics*, Hague, Netherlands, pp. 541–547.
5. J.A. Zukas (1990), *High Velocity Impact*. John Willey & Sons, New York.
6. M.L. Wilkins (1984), Modelling the behaviour of materials. *Proceedings of the International Conference on Structural Impact and Crashworthiness*, Vol. 2, London, UK, pp. 243–277.
7. W. Benz (1990), Smoothed particle hydrodynamics: a review. In:*Numerical Modeling of Non-linear Stellar Pulsation: Problems and Prospects*. Kluwer Academic, Boston.
8. J.J. Monaghan and J.C. Lattanzio (1985), A refined particle method for astrophysical problems. *Astronomy and Astrophysics*, 149, pp. 135–143.
9. G.R. Johnson and T.J. Holmquist (1988), Evaluation of cylinder-impact test data for constitutive model constants. *Journal of Applied Physics*, 64, 8, pp. 3901–3910.

A MESH FREE METHOD BASED ON AN OPTIMIZATION TECHNIQUE AND THE MOVING LEAST SQUARES APPROXIMATION

H. Dalayeli[1], H. Hashemolhosseini[2], M. Farzin[3], I. Pillinger[4]
and P. Hartley[4]
[1]*Mechanical Engineering Department, Malek-ashtar University of Technology, Shahin-shahr, Isfahan, 83145/115, Iran*
[2]*Mining Engineering Department, Isfahan University of Technology, Isfahan, 84156, Iran*
[3]*Mechanical Engineering Department, Isfahan University of Technology, Isfahan, 84156, Iran*
[4]*Mechanical Engineering, School of Engineering, The University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*

Abstract A mesh free method based on an optimization technique and the moving least-squares approximation is presented. In this method, the problem domain is modelled by a set of properly scattered points. Approximation functions are used to represent field variables and their derivatives. Like other mesh free methods, no mesh of elements is required. The other advantages are: prescribed displacements and tractions are imposed directly as a result of the minimization procedure, and there is no need to integrate the governing equations. The theory is developed for two-dimensional problems. As numerical examples of the proposed scheme, a cantilever beam and a Poisson equation are considered. The numerical results compare very well with the analytical solutions for these examples.

Keywords: optimization technique; MLS approximation; mesh free methods.

1. INTRODUCTION

The finite element method (FEM) is the most widely used method for the numerical solution of partial differential equations. There are, however, some shortcomings in using the FEM: the need to generate a discretised geometry (mesh) and the requirement that the elements must satisfy certain shape criteria, can limit its usefulness. In recent years, many researchers have focussed their attention on the element free or mesh free method, as it does not require a

G. R. Liu et al. (eds.), Computational Methods, 1411–1416.

structured mesh, and the approximate solution is constructed in terms of a set of scattered nodes, at least for field variable interpolation. There have been a number of mesh free methods developed so far [1–3]. In mesh free methods, approximation functions are constructed using techniques such as the moving least-squares (MLS) approximation, in which the influence of a node emanates from a weight function defined on an arbitrary compact subset of the domain. Because the MLS approximation function lacks the Kronecker delta function property, the implementation of prescribed boundary values often gives difficulties. Several methods have been introduced for the imposition of prescribed boundary conditions [3–5]. In the present work, nodal displacements are determined which minimize, in a least square sense, the error in the equilibrium equations, boundary displacements and boundary loads. The MLS procedure is employed to approximate field variable and their derivatives. As no background mesh is needed to perform the numerical integration, this new approach can be classed as a 'truly' mesh free method. In the following section, a brief description of the MLS approximation will be given. The formulation of the new method in two dimensions will then be described. Numerical examples will be presented and finally conclusions will be drawn about the scope of this work.

2. MOVING LEAST-SQUARES APPROXIMATION

The moving least-squares (MLS) approximation [6] is used to approximate field variables and their derivatives. In a domainΩ, the MLS approximation $u^h(\mathbf{x})$ of the field variable $u(\mathbf{x})$ in the vicinity of a point $\bar{\mathbf{x}}$ is given as

$$u^h(\mathbf{x}) = \sum_{j=1}^{m} p_j(\mathbf{x})a_j(\bar{\mathbf{x}}) = \mathbf{p}^{\mathrm{T}}(\mathbf{x})\mathbf{a}(\bar{\mathbf{x}}) \tag{1}$$

where$p_j(\mathbf{x})$, $j = 1, 2, \ldots m$ is a complete basis function in the space coordinates bi x and $a_j(\bar{\mathbf{x}})$ are its coefficients. At each point $\bar{\mathbf{x}}$, $a_j(\bar{\mathbf{x}})$ is chosen so as to minimize the weighted residual $\mathrm{L}_2 -$ norm:

$$J = \sum_{I=1}^{N} w(\bar{\mathbf{x}} - \mathbf{x}_I)[\mathbf{p}^{\mathrm{T}}(\mathbf{x}_I)\mathbf{a}(\bar{\mathbf{x}}) - \hat{u}_I]^2 \tag{2}$$

where N is the number of nodes I in the neighborhood of $\bar{\mathbf{x}}$ for which the weight function $w(\bar{\mathbf{x}} - \mathbf{x}_I) \neq 0$, and $\hat{u}_I$ refers to the nodal parameter of u at $\mathbf{x} = \mathbf{x}_I$. The minimum of J with respect to $\mathbf{a}(\bar{\mathbf{x}})$ gives the standard form of MLS approximation as

$$u^h(\mathbf{x}) = \sum_{I=1}^{N} \phi_I(\mathbf{x})\hat{u}_I = \Phi(\mathbf{x})\hat{\mathbf{u}}, \tag{3}$$

3. FORMULATION

The following two-dimensional problem of solid mechanics in domain Ω bounded by Γ is considered:

$$\nabla \cdot \boldsymbol{\sigma} + b = 0, \boldsymbol{\sigma} \cdot \mathbf{n} = \bar{\mathbf{t}} \text{ On } \Gamma_t, \mathbf{u} = \bar{\mathbf{u}} \text{ On } \Gamma_u \tag{4}$$

The residual or error function is defined as:

$$\begin{aligned} R = {} & \sum_{\mathbf{x}_k \in \Omega} \sum_{j=1}^{2} (\sigma_{ij,i} + b_j)^2_{(\mathbf{x}_k)} + \sum_{\mathbf{x}_k \in \Gamma_{t)\bar{\sigma}_{11}}} (\sigma_{11} - \bar{\sigma}_{11})^2_{(\mathbf{x}_k)} + \sum_{\mathbf{x}_k \in \Gamma_{t)\bar{\sigma}_{22}}} (\sigma_{22} - \bar{\sigma}_{22})^2_{(\mathbf{x}_k)} \\ & + \sum_{\mathbf{x}_k \in \Gamma_{t)\bar{\sigma}_{12}}} (\sigma_{12} - \bar{\sigma}_{12})^2_{(\mathbf{x}_k)} + \sum_{\mathbf{x}_k \in \Gamma_{\bar{u}}} (u - \bar{u})^2_{(\mathbf{x}_k)} + \sum_{\mathbf{x}_k \in \Gamma_{\bar{v}}} (v - \bar{v})^2_{(\mathbf{x}_k)} \end{aligned} \tag{5}$$

where subscript $(\mathbf{x}_k)$ indicates the quantity at point $\mathbf{x}_k$, $\Gamma_{t)\bar{\sigma}_{ij}}$ refers to the boundary with prescribed traction $\bar{\sigma}_{ij}$, $\Gamma_{\bar{u}}$ and $\Gamma_{\bar{v}}$ refer to the boundaries with the prescribed displacement values $\bar{u}$and $\bar{v}$, respectively. In this equation, the internal stress tensor σ_{ij} has been assumed to be symmetric and so $\sigma_{12} = \sigma_{21}$. The MLS approximation is used to approximate displacements and its derivatives at each nodal point $\mathbf{x}_k$, for example:

$$u = \sum_{I=1}^{N} \phi_I u_I, \;\; u,_1 = \sum_{I=1}^{N} \phi_{I,1} u_I \tag{6}$$

Where N is the number of nodes in the domain of influence of point $\mathbf{x}_k$, u is the nodal displacement and $u,_1 = \partial u / \partial x$. In this work, the following technique is used to approximate the second derivative of field variables, for example:

$$u,_1 = \sum_{I=1}^{N} \phi_I u_{I,1}, \;\; u,_{11} = \sum_{I=1}^{N} \phi_{I,1} u_{I,1} \tag{7}$$

The main advantage of this technique is that when writing the second derivative of field variables, only the first derivative of MLS approximation functions is used. The system of algebraic equations is obtained by minimizing the error function (5) with respect to the unknown nodal displacement values. This satisfies the equilibrium equations at all nodal points and also the prescribed

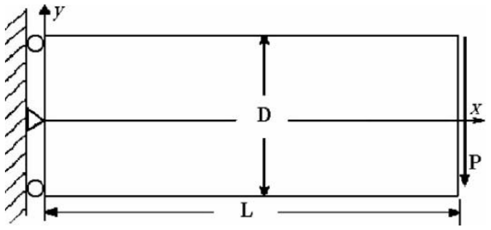


Figure 1. Cantilever beam.

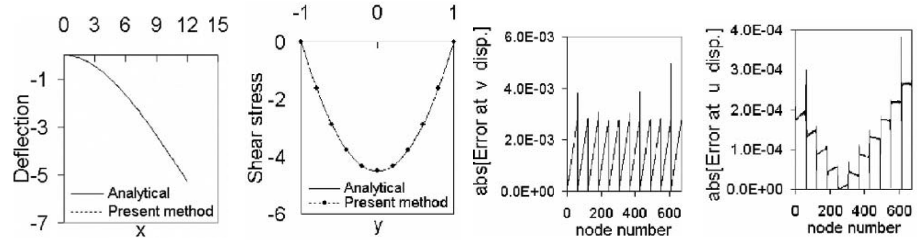
displacement and load boundary conditions:

$$\frac{\partial R}{\partial \mathbf{u}} = 0 \Rightarrow \mathbf{Cu} = \mathbf{d} \Rightarrow \mathbf{u} = \mathbf{C}^{-1}\mathbf{d}, \quad \mathbf{u}^{\mathrm{T}} = [u_1 \quad v_1 \quad u_2 \quad v_2 \quad \ldots \quad u_n \quad v_n] \tag{8}$$

n is the total number of nodes in domain Ω. The unknown nodal displacement parameters are obtained by Equation (8). The true value of the field variables should be retrieved using Equation (3).

4. NUMERICAL EXAMPLES

The first example is a cantilever beam subjected to a parabolic traction at the free end as shown in Figure 1. The exact solution is given by Timoshenko and Goodier [7]. The beam has a unit thickness and is solved with dimensionless parameters $E = 1000$, $\nu = 0.3$, $P = 6$, D = 2 and L = 12. A uniformly distributed nodes, 671 nodes (Nx = 61, Ny = 11), and a complete basis functions with m = 10 are used. Figure 2 shows a good agreement between the analytical and numerical results.



(a) **(b)** **(c)** **(d)**

Figure 2. (a) Deflection of the beam along y = 0; (b) Shear stress at the section x = L/2; (c) and (d) Comparison between the analytical displacements and those calculated by the present method.

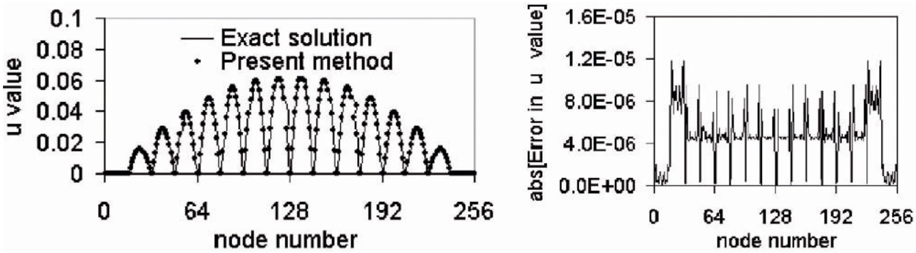


Figure 3. Comparison between the analytical results and those calculated by the present method.

As the second example a Poisson equation is considered as follows:

$$\nabla^2 u = -2(x + y - x^2 - y^2) \quad \text{in } \Omega = [0, 1] \times [0, 1] \text{ and } u = 0 \quad \text{on } \partial\Omega \tag{9}$$

whose analytical solution is given by $u(x, y) = (x - x^2)(y - y^2)$. A uniformly distributed nodes, 256 nodes (Nx = 16, Ny = 16), and a complete basis functions with m = 10 are used. Figure 3 shows an excellent agreement between the analytical and numerical results.

5. CONCLUSION

A truly mesh free method based on optimization techniques and the moving least-squares approximation has been proposed. The advantages of the new technique can be summarized as follows: (a) Prescribed boundary displacements and loads can be imposed directly and there is no need to use the Lagrange multiplier, penalty constant or coupling with the BEM or FEM methods, (b) The method does not involve integration, therefore there is no need for any background mesh. Since the method does not require a mesh, nodal points can be selected simply on the basis of the accuracy of solution that is required from place to place in solving the physical problem under consideration. Numerical examples, a cantilever beam and a Poisson equation, have been presented to evaluate the accuracy of the present method. The calculated results are in good agreement with the analytical solutions.

REFERENCES

1. B. Nayroles, G. Touzot and P. Villon (1992), Generalizing the finite element method: diffuse approximation and diffuse elements. *Computational Mechanics*, 10, pp. 307–318.

2. T. Belytschko, Y. Krongauz, D. Organ, M. Fleming and P. Krysl (1996), Meshless methods: an overview and recent developments. *Computer Methods in Applied Mechanics and Engineering*, 139, pp. 3–47.
3. G.R. Liu (2002), *Mesh Free Methods: Moving Beyond the Finite Element Method*. CRC Press, Boca Raton, USA.
4. G.R. Liu and K.Y. Yang (1998), A penalty method for enforcing essential boundary conditions in element free Galerkin method. In: *The Proceedings of the 3rd HPC Asia'98*, Singapore, pp. 715–721.
5. G.R. Liu and Y.T. Gu (2000), Coupling of element free Galerkin and hybrid boundary element methods using modified variational formulation. *Computational Mechanics*, 26, pp. 166–173.
6. P. Lancaster and K. Salkauskas (1981), Surfaces generated by moving least square methods. *Mathematics of Computation*, 37, pp. 141–158.
7. S.P. Timoshenko and J.N. Goodier (1970), *Theory of Elasticity*, 3rd ed. McGraw-Hill, New York.

FREE VIBRATION ANALYSIS OF TIMOSHENKO BEAMS BY RADIAL BASIS FUNCTIONS

A.J.M. Ferreira
Departamento de Engenharia Mecânica e Gestão Industrial, Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

Abstract A study of free vibration of Timoshenko beams is presented. The analysis is based on the radial basis function method. Clampled and simply supported boundary conditions are considered. Various approaches for computing eigenvalues using radial basis functions are presented and discussed. Numerical results are presented and discussed for various thickness-to-length ratios.

Keywords: free vibration, Timoshenko beams, radial basis functions, multiquadrics.

1. INTRODUCTION

Vibration of beams is an important subject in the design of mechanical, civil and aerospace applications. The thickness of most part of beams makes the transverse shear and the rotary inertia not negligible as in classical theories. Therefore the thick beam model of Timoshenko should be considered in general analysis. The analysis of free vibration of Timoshenko beams cannot be performed by analytical methods, with exception of few simple cases. Numerical techniques have been applied to this type of problems. The differential quadrature method [1–3], the boundary characteristic orthogonal polynomials [4] and the pseudo-spectral method [5] were used in recent years. The finite element method also proved to be very adequate for this type of problems. The number of references is too large to reference here.

The present paper addresses the free vibration analysis of Timoshenko beams by the radial basis function method. This is a truly meshless method where discretization of the equations of motion and boundary conditions (natural or essential) does not require a mesh. The method is simple to code and provides good solutions. In particular, the imposition of boundary conditions is

G. R. Liu et al. (eds.), Computational Methods, 1417–1421.

very simple. Although powerful and simple, the method has not been applied to structural mechanics with sufficient impact. The author used the multiquadric radial basis function method in the static analysis of composite beams and plates [6–8]. In the present work we illustrate the application of the radial basis function [9,10] to the eigenvalue analysis of Timoshenko beams.

2. TIMOSHENKO BEAMS

The equations of motion of a rectangular cross-section homogeneous beam based on the Timoshenko beam assumptions are

$$\frac{\partial M}{\partial x} + V = \rho I \frac{\partial^2 \Psi}{\partial t^2}, \frac{\partial V}{\partial x} = \rho h \frac{\partial^2 W}{\partial t^2} \tag{1}$$

where $W(x, t)$ and $\Psi(x, t)$ are the transverse displacement and the rotation of the normal, ρ is the mass density, h is the thickness of the beam, I is the second moment of area per unit width, and M and V are the moments and shear resultants defined as

$$M = EI \frac{\partial \Psi}{\partial x}, V = \alpha G h \left(\frac{\partial W}{\partial x} - \Psi \right) \tag{2}$$

where E is the modulus of elasticity, G is the shear modulus, and α is the shear coefficient. Assuming a sinusoidal motion in time

$$\Psi(x, t) = \theta(x) \cos \omega t, W(x, t) = w(x) \cos \omega t \tag{3}$$

The substitution of (2) into (1) produces the equations of motion in terms of generalized displacements

$$EI \frac{d^2\theta}{dx^2} + \alpha h G \left(\frac{dw}{dx} - \theta \right) = -\omega^2 \rho I \theta, -\alpha h G \frac{d\theta}{dx} + \alpha h G \frac{d^2 w}{dx^2} = -\omega^2 \rho h w \tag{4}$$

3. NUMERICAL RESULTS

We tested this method on the Timoshenko beam problem, described in (1)–(3). We computed eigenmodes in a linear beam of unit length. We used multiquadric RBFs $\phi(r) = \sqrt{c^2 + r^2}$, with $c = 2/\sqrt{N}$. We denote the jth eigenvalue by $lambda^{(j)}$, $0 < \lambda^{(1)} \leq \lambda^{(2)} \leq \ldots$ In Table 1, we run a convergence check of the eigenvalues of the Timoshenko beam with clamped–clamped boundary conditions for $h/L = 0.01$. The number of collocation points varies from 10 to 40. The results are compared with analytical (Euler–Bernoulli beam)

Table 1. Convergence test of the non-dimensionalized frequency parameter λ_i of the Timoshenko beam as the number of the collocation points N increases (clamped/clamped boundary condition, $\nu = 0.3, \alpha = 5/6, h/L = 0.01$).

Mode	Model	N = 10	N =15	N = 20	N = 25	N = 30	N = 35	N = 40
1	[5]	4.72840	4.72840	4.72840	4.72840	4.72840	4.72840	4.72840
	Present	4.8734	4.7551	4.7358	4.7308	4.7293	4.7287	4.7287
2	[5]	7.84691	7.84690	7.84690	7.84690	7.84690	7.84690	7.84690
	Present	8.1295	7.9063	7.8629	7.8522	7.8488	7.8478	7.8472
3	[5]	10.9827	10.9800	10.9800	10.9800	10.9800	10.9800	10,9800
	Present	11.4545	11.0600	11.0030	10.9880	10.9830	10.9820	10,9807
4	[5]	14.1132	14.1062	14.1062	14.1062	14.1062	14.1062	14.1062
	Present	14.8013	14.2180	14.1390	14.1180	14.1110	14.1080	14,1070
5	[5]	18.1397	17.2246	17.2246	17.2246	17.2246	17.2246	17,2246
	Present	19.3222	17.3170	17.2570	17.2380	17.2300	17.2290	17,2257
6	[5]	21.5723	20.3422	20.3338	20.3338	20.3338	20.3338	20,3338
	Present	31.0880	20.4180	20.3620	20.3490	20.3410	20.3370	20,3355
7	[5]	38.0443	23.4481	23,4325	23.4325	23.4325	23.4325	23,4325
	Present	44.0440	23.3960	23.4290	23.4440	23.4400	23.4390	23,4345
8	[5]	42.2513	27.1739	26.35192	26.5192	26.5192	26.5192	26,5192
	Present	60.5741	26.5670	26,4550	26.5250	26.5280	26.5250	26.5219
9	[5]		30.4163	29.6033	29.5926	29.5926	29.5926	29,5926
	Present		29.6940	29,4020	29.5780	29.6000	29.5990	29.5964
10	[5]		40.1211	32,6684	32.6515	32.6514	32.6514	32.6514
	Present		34.7720	32,2530	32.6030	32.6580	32.6610	32.6558
11	[5]		43.8761	36,2185	35.6947	35.6946	35.6946	35.6946
	Present		46.2410	35,0390	35.5720	35.6950	35.7000	35,7045
12	[5]		79.1294	39,3662	38.7331	38.7209	38.7209	38,7209
	Present		58.8870	37,7880	38.4770	38.7120	38.7420	38,7258
13	[5]		83.0320	46.4220	41.7474	41.7294	41.7293	41,7293
	Present		77.6700	40.8370	41.2290	41.6990	41.7370	41,7527
14	[5]			49,8430	45.1366	44.7191	44.7189	44.7189
	Present			44,0980	43.8660	44.6410	44.7830	44.7070
15	[5]			68.0861	48.1952	47.7008	47.6888	47.6888
	Present			48.9380	46.2550	47.5270	47.6740	47.6948

and pseudo-spectral results of Lee and Schultz [5]. This table shows clearly a rapid convergence of the radial basis function method. In the following tables we use $N = 35$, due to good results using such number of collocation points, but also to compare with [5]. Values in Tables 1 and 2 are non-dimensional frequency parameters λ_i, defined as

$$\lambda_i^2 = \omega_i L^2 \sqrt{\frac{m}{EI}} \tag{5}$$

where m is the mass per unit length of the beam and L is the length of the beam. In all results we fixed Poisson's ratio and the shear coefficient of the

Table 2. Non-dimensionalized frequency parameter λ_i of the Timoshenko beam (pinned/pinned boundary condition, $v = 0.3$, $\alpha = 5/6$, $N = 35$).

Mode	Classical theory	Model	h/L						
1	3.14159	[5]	3.14158	3.14153	3.14133	3.14053	3.13498	3.11568	3.04533
		Present	3.14080	3.14090	3.14120	3.14020	3.13510	3.11580	3.04540
2	6.28319	[5]	6.28310	6.28265	6.28106	6.27471	6.23136	6.09066	5.67155
		Present	6.28620	6.28500	6.28220	6.27520	6.23160	6.09080	5.67160
3	9.42478	[5]	9.42449	9.42298	9.41761	9.39632	9.25537	8.84052	7.83952
		Present	9.43070	9.42770	9.41790	9.39750	9.25560	8.84060	7.83960
4	12.5664	[5]	12.5657	12.5621	12.5494	12.4994	12.1813	11.3431	9.65709
		Present	12.5746	12.5688	12.5520	12.5003	12.1814	11.3430	9.65700
5	15.7080	[5]	15.7066	15.6997	15.6749	15.5784	14.9926	13.6132	11.2220
		Present	15.7136	15.7046	15.6740	15.5792	14.9924	13.6129	11.2219
6	18.8496	[5]	18.8473	18.8352	18.7926	18.6282	17.6810	15.6790	12.6022
		Present	18.8535	18.8396	18.7932	18.6278	17.6803	15.6785	12.6019
7	21.9911	[5]	21.9875	21.9684	21.9011	21.6443	20.2447	17.5705	13.0323
		Present	21.9862	21.9664	21.8977	21.6427	20.2435	17.5700	13.0323
8	25.1327	[5]	25.1273	25.0988	24.9988	24.6227	22.6862	19.3142	13.4443
		Present	25.1167	25.0902	24.9925	24.6193	22.6847	19.3140	13.4443
9	28.2743	[5]	28.2666	28.2261	28.0845	27.5599	25.0111	20.9325	13.8433
		Present	28.2406	28.2046	28.0759	27.5535	25.0098	20.9336	13.8432
10	31.4159	[5]	31.4053	31.3498	31.1568	30.4533	27.2263	22.4441	14.4378
		Present	31.3508	31.3089	31.1364	30.4466	27.2262	22.4481	14.4378
11	34.5575	[5]	34.5434	34.4697	34.2145	33.3006	29.3394	23.8639	14.9766
		Present	34.4613	34.4059	34.1873	33.2870	29.3430	23.8756	14.9769
12	37.6991	[5]	37.6807	37.5853	37.2565	36.1001	31.3581	25.2044	15.6676
		Present	37.5265	37.4733	37.2114	36.0971	31.3686	25.2248	15.6677
13	40.8407	[5]	40.8174	40.6962	40.2815	38.8507	33.2896	26.0647	16.0241
		Present	40.6156	40.5433	40.1637	38.8280	33.3168	26.0647	16.0269
14	43.9823	[5]	43.9531	43.8021	43.2886	41 5517	35.1410	26.2814	16.9584
		Present	43.5748	43.5349	43.2015	41.5892	35.1917	26.2814	16.9583
15	47.1239	[5]	47.0880	46.9027	46.2769	44.2026	36.9186	26.4758	17.0019
		Present	46.6156	46.5579	45.7626	44.1734	37.0548	26.5454	17.0096

beam as $\nu = 0.3$ and $\alpha = 5/6$, respectively. In Table 2 we compare eigenvalues with classical theory and [5] for pinned/pinned boundary conditions. The eigenvalues are calculated for different thickness-to-length ratios, from $h/L = 0.002$ to 0.2. Our results are in good agreement with the pseudo-spectral results of Lee and Schultz [5], using a regular node arrangement. Results show excellent agreement with classical solutions for $h/L < 0.01$. For larger h/L values Timoshenko theory produces eigenvalues with differences to the Euler-Bernoulli results. The convergence of the pseudo-spectral method is somewhat faster than the present method.

4. CONCLUSION

In this paper we used the multiquadric radial basis function method to analyse free vibrations of Timoshenko beams. The first-order shear deformation theory set of equations define a eigenproblem which can be solved by various algorithms. The results of eigenvalues are compared with a pseudo-spectral method and are found to be in excellent agreement with this recent method.

ACKNOWLEDGEMENTS

The financial support of the Materials Network for the Atlantic Area, of the Interreg IIIB programme of the European Comission is gratefully acknowledged. Also the financial support of Fundação para a Ciência e a Tecnologia under Financiamento Plurianual to Mecânica Experimental e Novos Materiais is gratefully acknowledged.

REFERENCES

1. P.A.A. Laura and R.H. Gutierrez (1993), Analysis of vibrating timoshenko beams using the method of differential quadrature. *Shock and Vibration*, 1, 1, pp. 89–93.
2. C.W. Bert and M. Malik (1996), Differential quadrature method in computational mechanics: a review. *Applied Mechanics Review*, 49, 1, pp. 1–28, 1996.
3. K.M. Liew, J.B. Han and Z.M. Xiao (1997), Vibration analysis of circular mindlin plates using differential quadrature method. *Journal of Sound and Vibration*, 205, 5, pp. 617–630.
4. K.M. Liew, K.C. Hung and M.K. Lim. (1995), Vibration of mindlin plates using boundary characteristic orthogonal polynomials. *Journal of Sound and Vibration*, 182, 1, pp. 77–90.
5. J. Lee and W.W. Schultz (2004), Eigenvalue analysis of timoshenko beams and axisymmetric mindlin plates by the pseudospectral method. *Journal of Sound and Vibration*, 269, pp. 609–621.
6. A.J.M. Ferreira (2003), A formulation of the multiquadric radial basis function method for the analysis of laminated composite plates. *Composite Structures*, 59, pp. 385–392.
7. A.J.M. Ferreira, C.M.C. Roque and P.A.L.S. Martins (2003), Analysis of composite plates using higher-order shear deformation theory and a finite point formulation based on the multiquadric radial basis function method. *Composites: Part B*, 34, pp. 627–636.
8. A.J.M. Ferreira (2003), Thick composite beam analysis using a global meshless approximation based on radial basis functions. *Mechanics of Advanced Materials and Structures*, 10, pp. 271–284.
9. R.L. Hardy (1971), Multiquadric equations of topography and other irregular surfaces. *Geophysical Research*, 176, pp. 1905–1915
10. E.J. Kansa (1990), Multiquadrics—a scattered data approximation scheme with applications to computational fluid dynamics. i: Surface approximations and partial derivative estimates. *Computers and Mathematics with Applications*, 19, 8/9, pp. 127–145.

MOVING LEAST SQUARE SPH USING FIXED KERNEL FOR LARGE DEFORMATION ELASTO-PLASTIC ANALYSIS

H. Minaki[1] and H. Noguchi[2]
[1]*School of Science for Open and Environmental Systems, Keio University, 3-14-1, Hiyoshi, Kohoku-ku, Yokohama 223-8522, Japan*
[2]*Department of System Design Engineering, Keio University, Yokohama 223-8522, Japan*

Abstract SPH using the moving least square (MLS) approximation and fixed kernel function is proposed and is applied to large deformation elasto-plastic analysis. SPH is one of the particle methods and is a powerful tool for large-scale hydro dynamics analysis. SPH is, however, not efficient when the kernel functions are calculated at every step and the spatial approximation is less accurate due to only the 0-th order consistency. In this study, Fixed Kernel method in which the kernel is calculated only at the original configuration is utilized in order to reduce numerical cost and MLS approximation is applied for the spatial approximation to improve accuracy. Furthermore, the Poisson's equation for pressure is solved to take account of incompressibility in plastic state. Numerical analysis of elastic problem and elasto-plastic problem are demonstrated to validate the present method.

Keywords: SPH, moving-least-square, fixed kernel, incompressibility.

1. INTRODUCTION

A particle method is one of the mesh free methods and is often utilized for dynamic analyses of large-scale problems. SPH is the most widely used particle method originally developed for astrodynamics problems [1] and has been applied to various engineering fields including elasto-plastic problems [2]. The conventional SPH is not efficient when the kernel functions are calculated at every step and the spatial approximation is less accurate due to only the 0-th order consistency. In this research, therefore, Fixed Kernel method in which the kernel is calculated only at the original configuration is

G. R. Liu et al. (eds.), Computational Methods, 1423–1428.

utilized in order to reduce numerical cost and MLS approximation is applied for the spatial approximation to improve accuracy. Furthermore, the Poisson's equation for pressure is solved to take account of incompressibility in plastic state.

2. IMPROVED SPH USING MLS AND FIXED KERNEL [3]

In SPH, the approximate value $f(\mathbf{x})$ of a function f at a point $\mathbf{x}$ is given by

$$\langle f(\boldsymbol{x})\rangle = \sum_{j=1}^{N} \frac{m_j}{\rho_j} f(\boldsymbol{x}_j) W(\mathbf{x}-\mathbf{x}_j) \tag{1}$$

where $W(\mathbf{x}-\mathbf{x}_j)$ is kernel function, m is mass, ρ is density.

2.1 Moving least square [4]

Although the Equation (1) is used for approximation in SPH, it is known that it has only the 0-th order accuracy. In order to improve this deficiency, in this research, the moving least squares (MLS) approximation is utilized for the approximate value $f(\mathbf{x})$ of a function f. In MLS, a function is approximated so as to minimize the following weighted functional.

$$J = \sum_{j=1}^{N} W(\mathbf{x}-\mathbf{x}_j)(u^h(\mathbf{x}_j) - u_j) \tag{2}$$

where u_j is a value in each particle. W is a weight function in an influence domain and u^h is a basis function and a polynomial function is usually used for u^h. In case a linear polynomial function is used, the obtained approximation function has at least the 1st order accuracy and can recover a linear function. As not only the function but its spatial derivative obtained by MLS becomes continuous, the accuracy of SPH may be improved. Finally, in MLS, a function u^h and its derivative can be expressed as follows. Where ϕ is shape function obtained by MLS.

$$u^h(\mathbf{x}_j) = \sum_{j=1}^{N} \phi_j u_j,\ \nabla u^h(\mathbf{x}_j) = \sum_{j=1}^{N} \nabla\phi_j u_j \tag{3}$$

2.2 Fixed kernel

As particle configuration changes with time, in SPH, the kernel is recomputed out using current configuration at $t_0 = t$ every time. In this research, therefore, the fixed kernel to the initial configuration at $t_0 = 0$ (Fixed Kernel) is used for computational efficiency. In Fixed Kernel method, the shape of influence domain deforms according to the change of particle configuration, however, the kernel estimate can be computed once at the initial configuration.

In order to use Fixed Kernel, the law of conservation of momentum expressed with initial configuration should be solved. The law of conservation of momentum in MLS–SPH is given by

$$\frac{d\mathbf{V}_i}{dt} = \frac{1}{\rho_0} \sum_{j=1}^{N} (\mathbf{\Pi}_i + \mathbf{\Pi}_j - Q_{ij}\mathbf{I}) \cdot \nabla\phi \tag{4}$$

where i is a particle at an evaluation position and j is a particle within an influence domain,$\mathbf{V}$ is velocity vector, ρ_0 is density at initial time, Π is the 1st Piola-Kirchhoff stress, Q_{ij} is artificial viscosity.

2.3 SMAC method

Even if metal solid is compressible in elastic state and small deformation, it becomes nearly incompressible in plastic state and large deformation. To assure stable solution in incompressible state, the pressure is modified according to SMAC scheme, which is occasionally used for analysis of incompressible flow, as shown in Equation (5). Equation (5) is the modified Poisson's equation with the effect of elastic deformation.

$$\nabla^2 P' - \frac{\rho}{\kappa \Delta t^2} P' = \frac{\rho}{\Delta t} \nabla \cdot \mathbf{V}' \tag{5}$$

where P' is pressure corrections ($P^{n+1} = P^n + P'$), Δt is time increment, κ is bulk modulus, $\mathbf{V}'$ is velocity estimated explicitly.

3. NUMERICAL EXAMPLES

3.1 Axial vibration of bar

A two dimensional free axial vibration analysis of an elastic bar is conducted. One end of the bar is fixed and impact velocity is applied to all particles. The Length of bar is 0.5 m, the Young's modulus is 200 GPa, the Poisson's ratio

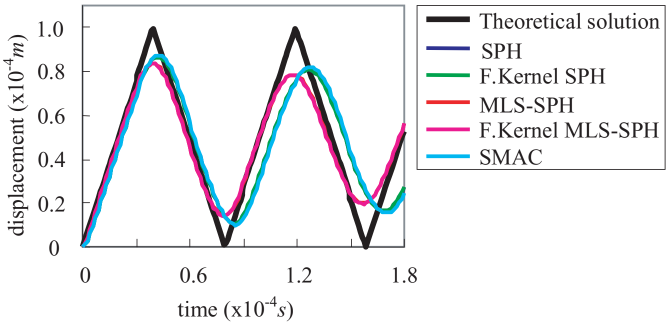


Figure 1. Displacement history.

is 0, the density is 7890 kg/m^3, the number of particles is 105, the number of steps is set to 6000. Impact velocity is 1.0 m/s.

The analysis results with theoretical solution of displacement history at $x = 0$ are shown in Figure 1. Normalized computational time is shown in Figure 2. Comparing the result by MLS–SPH with that by SPH, MLS–SPH gives better accuracy. It is also found that Fixed Kernel method gives identical results by the conventional SPH, and SMAC with the conventional SPH never deteriorate the accuracy. On the other hand, the computation time by MLS–SPH using Fixed Kernel is greatly reduced compared with that by MLS–SPH referring the current configuration.

3.2 Cylinder impact

In order to compare the conventional SPH and SPH extended by SMAC scheme, a two-dimensional large deformation elasto-plastic analysis is carried out. The analysis conditions are the Young's modulus is 200 GPa, the Poisson's ratio is 0.3, the yield stress is 500 MPa, the density is 7.89 g/cm^3, the length is

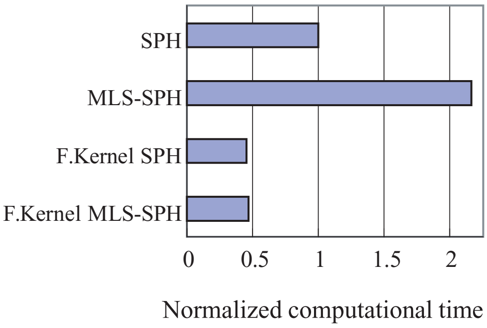


Figure 2. Computational time.

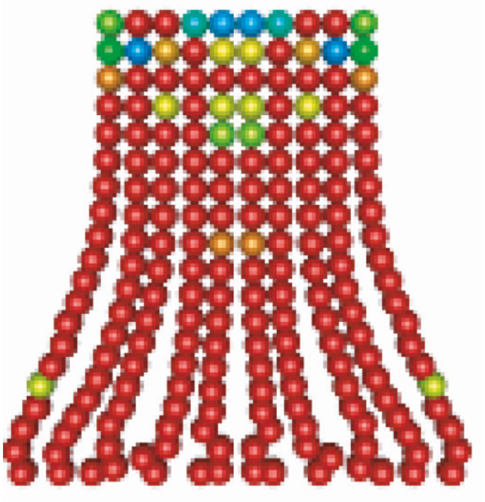


Figure 3. Conventional SPH.

0.012 m, the width is 0.006 m, the number of particles is 496, and the impact velocity is 221 m/s. Figures 3 and 4 show the deformation of cylinder and the distribution of equivalent stress at the time of 50 μs. As results, The stable deformation and stress distribution of cylinder are obtained by SPH with SMAC scheme.

4. CONCLUSION

SPH using the moving least square approximation and fixed kernel function was proposed and was applied to the linear elastic and the large deformation elasto-plastic analysis. It was demonstrated through the numerical examples that firstly the accuracy and the computational efficiency of the present method

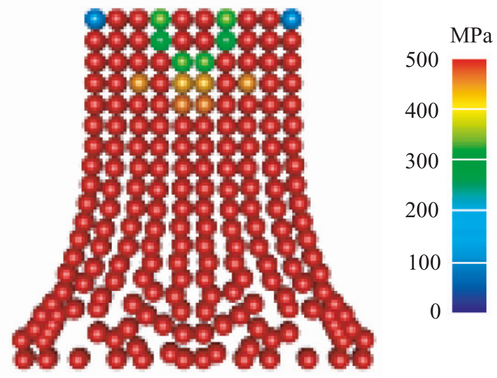


Figure 4. SPH with SMAC scheme.

were improved and secondly the modification of pressure field by solving the Poisson's equation to take account of incompressibility in plastic state was effective for the stabilization of the numerical solution.

REFERENCES

1. L.B. Lucy (1977), A numerical approach to the testing of the fission hypothesis. *Astronomy Journal*, 82, pp. 1013–1020.
2. L.D. Libersky, A.G. Petschek, T.C. Carney, J.R. Hipp and F.A. Allahdadi (1993), High strain lagrangian hydrodynamics a three-dimensional SPH code for dynamic material response. *Journal of Computational Physics*, 109, pp. 67–75.
3. P.W. Randles and D.L. Larry (1996), Smoothed particle hydrodynamics: some recent improvements and applications. *Computer Methods in Applied Mechanics and Engineering*, 139, pp. 375–408.
4. A.D. Gary (1999), Moving-least-squares-particle hydrodynamics. *Consistency and Stability. International Journal for Numerical Methods in Engineering*, 44, pp. 1115–1155.

A COUPLED MESHFREE/SCALED BOUNDARY METHOD

C.E. Augarde[1] and A.J. Deeks[2]
[1] *School of Engineering, University of Durham, South Road, Durham, DH1 3LE, UK*
[2] *School of Civil & Resource Engineering, The University of Western Australia, 35 Stirling Highway, Crawley 6009, Australia*

Abstract The scaled boundary method is an excellent way to model unbounded domains. However, it is limited to linear problems. Many soft-ground geotechnical problems require both non-linear constitutive behaviour for the soil, to capture pre-failure deformations, and the presence of an unbounded domain. Adaptive meshfree methods are ideally suited to such problems. This paper couples a meshless local Petrov–Galerkin method for the near field with a meshless scaled boundary method of similar type for the far field. The method presented is novel as the degrees of freedom of all nodes in the support of the interface nodes are coupled to the stiffness of the unbounded domain, rather than just the nodes on the interface.

Keywords: meshfree method, scaled boundary method.

1. INTRODUCTION

While meshless methods are increasingly seen as future replacements for the conventional finite element method, they retain the shortcomings of the conventional finite element method when modelling singularities or infinite boundaries. Both features can, however, be dealt with efficiently using the Scaled Boundary Method (SBM) although this method cannot incorporate non-linear material behaviour unlike meshless methods. This paper describes the coupling of a meshless method to the SBM to yield a numerical method ideal for problems in areas such as geomechanics where both non-linear constitutive behaviour and accurate modelling of infinite boundaries are required.

In this study we use the Meshless Local Petrov–Galerkin (MLPG) method [1]. This method is based on a moving least squares (MLS) approximation

G. R. Liu et al. (eds.), Computational Methods, 1429–1433.

for the displacement field, as are other popular meshless methods. The SBM is semi-analytical and was developed relatively recently by Wolf and Song [2, 3]. Understanding and interest in the SBM has since increased partly due to publication of a virtual work derivation of the method for elastostatics [4]. In its simplest form, a point in a domain is assigned a scaling centre, from which radiators are defined along which the displacement field is an analytical (i.e., closed-form) solution. In the circumferential direction the displacement field is approximated by conventional finite element type shape functions.

2. COUPLING THE METHODS

The coupled method is shown in Figure 1 for the case of a footing problem (later used to demonstrate the hybrid method). The MLPG method is used in the near field (i.e., close to an applied load or prescribed displacement condition). SBM elements are used along the material boundary of the meshless region, thus the remainder of the infinite domain is covered. Doherty and Deeks [5] have already combined the SBM with conventional finite elements. In that case coupling was relatively simple as the same shape functions are used in each method along the interface between the two zones. A first step in coupling the MLPG method and the SBM is to reformulate the SBM using the same shape functions in the circumferential direction as used in the MLPG method to produce a SBM without elements. This is also straightforward as

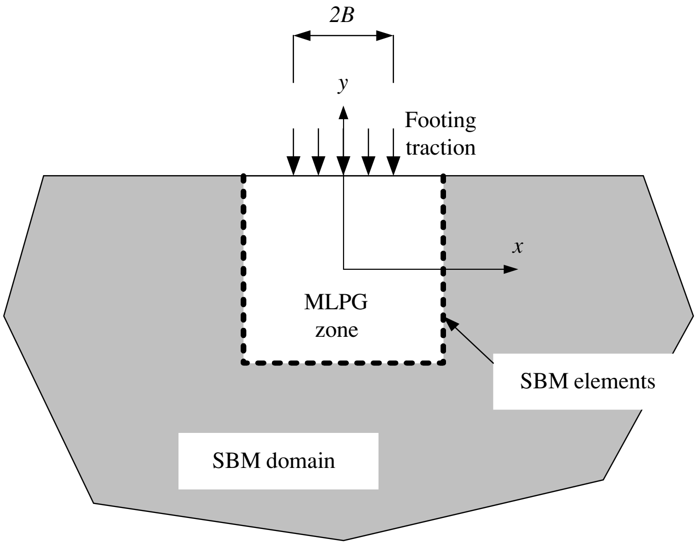


Figure 1. Hybrid method applied to the smooth footing problem.

the approximation is one-dimensional, along the boundary s. The procedure is described in detail in Deeks and Augarde [6].

Having produced an 'elementless' SBM it would seem to be straightforward to then couple MLPG with this method. However complications arise as the MLPG returns fictitious nodal values $\{\hat{u}_m\}$ to which the MLS approximation is fitted rather than actual nodal values, $\{u_u\}$ as returned by the SBM. This feature of meshless methods leads to difficulties in enforcement of essential boundary conditions, as highlighted elsewhere [7]. Along the SBM 'side' of the boundary the approximation is governed by nodes on the boundary. On the MLPG side, however, the approximation is influenced in addition by nodes inside the meshless domain. For these reasons, direct coupling is not possible and a penalty approach is used to couple the two methods. The SBM approximation is more restricted than the MLPG approximation to which it must be coupled, since at any point the latter is by definition based on fewer nodes. Therefore the MLPG approximation is restricted to the SBM values, rather than the other way round.

In brief, the coupling is implemented as follows. If Γ_i is the interface between the zone then equilibrium is satisfied in a weak sense in the meshless zone if

$$[K]\{\hat{u}_m\} - \int\limits_{\Gamma_i} \left[N^2(x,y)\right]^T \{t\} d\Gamma_i = \{f\} \tag{1}$$

where $[K]$ is the stiffness matrix for the meshless region, $[N^2(x,y)]$ is a matrix containing meshless test functions, $\{t\}$ are the tractions along the interface between the zones and $\{f\}$ are the externally applied forces (assumed not to occur along the interface). A similar equilibrium expression can be written for the SBM zone in terms of both $\{u_u\}$ and $\{\hat{u}_m\}$, recalling that along the interface the two displacements are kept separate. Two further equations are obtained from enforcing compatibility along the interface. Derivation for the meshless zone begins from

$$[N^1(s)]\{u_u\} = [N^1(x,y)]\{\hat{u}_m\} \quad \text{over } \Gamma_i \tag{2}$$

where $[N^1(s)]$ is a matrix of shape functions for the SBM (note the single coordinate, s) and $[N^1(x,y)]$ is the corresponding matrix for the meshless region. Combining the equilibrium expression for each zone with its corresponding compatibility condition (the latter weighted with a penalty parameter $\alpha \gg 0$) leads to a system of $2(n_u + n_m)$ linear equations in $2(n_u + n_m)$ unknown nodal values (where n_u is the number of nodes on the interface and n_m is the total number of nodes in the meshless region). Following solution for displacements, stresses can be recovered in both zones. A full derivation is given elsewhere [8].

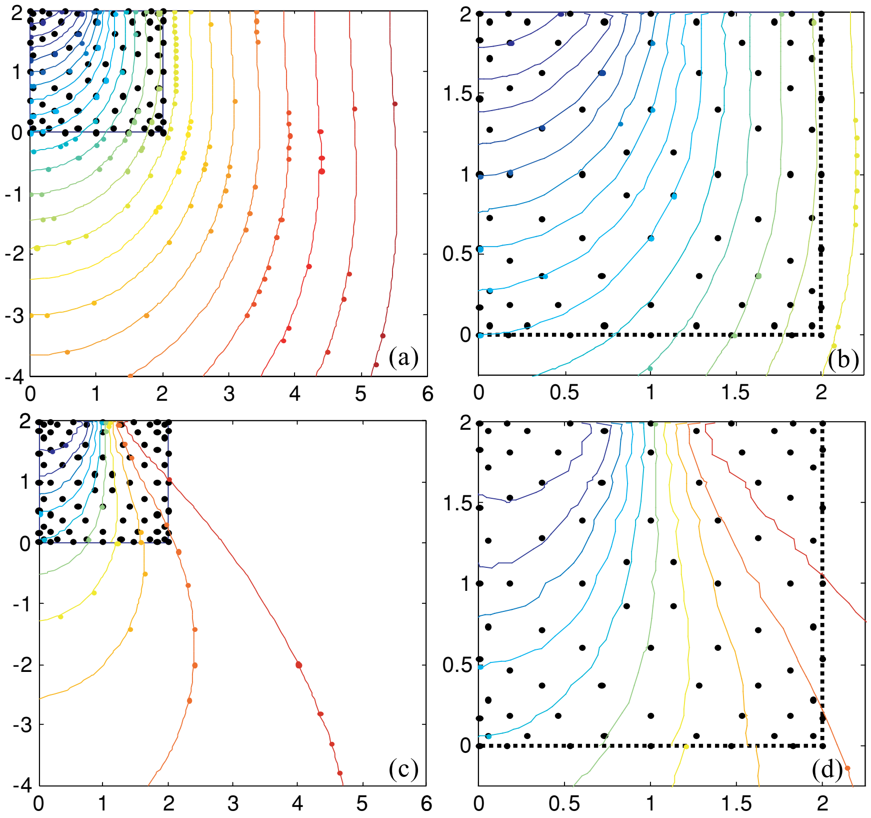


Figure 2. Results for the footing problem. Vertical displacements (a) and (b) in close up; vertical stresses (c) and (d) in close up.

3. AN EXAMPLE

The hybrid method outlined above is demonstrated on a simple elastostatics problem, namely a plane strain smooth footing of width 2 units on an elastic half-space (Figure 1). A uniform traction is applied in the negative y-direction over the footing width. Only one half of the problem is modelled due to symmetry; the symmetric boundary is the y-axis. A coarse irregular grid of 97 nodes covers the meshless region, which is square of size 2 units. (Grid spacing is based on Gauss–Lobatto intervals). Thirteen SBM nodes are used along the interface. Displacement and stress results are shown in Figure 2 for the meshless area (where the nodes are also shown) and in the infinite domain for $(0 \leq x \leq 6;\ -4 \leq y \leq 2)$.

The plots show that the penalty method is successful in enforcing the coupling between the methods on the boundary as displacements are smooth across the interface. The resulting stress field derived from processing the meshless displacements and the SBM displacements separately is also seen to produce an acceptable result, although there are minor discrepancies between the zones visible along the interface which could be reduced by refinement along the interface.

4. CONCLUSIONS

A hybrid method has been described and demonstrated that couples a meshless method with the Scaled Boundary method thereby potentially allowing nonlinear behaviour in the near-field with infinite boundaries, and the economies that produces. Further work is in progress to test and extend this hybrid method particularly for use in geomechanics problems.

REFERENCES

1. S.N. Atluri and T. Zhu (1998), A new Meshless Local Petrov–Galerkin (MLPG) approach in computational mechanics. *Computational Mechanics*, 22, pp. 117–127.
2. J.P. Wolf and Ch. Song (1996), *Finite-Element Modelling of Unbounded Medial*, John Wiley and Sons, Chichester.
3. J.P. Wolf and Ch. Song (2001), The scaled boundary finite-element method—a fundamental solutionless boundary-element method and Computer Methods in Applied Mechanics and Engineering, 190, pp. 5551–5568.
4. A.J. Deeks and J.P. Wolf (2002), A virtual work derivation of the scaled boundary finite-element method for elastostatics. *Computational Mechanics*, 28, 6, pp. 489–504.
5. J.P. Doherty and A.J. Deeks (2003), Scaled boundary finite-element analysis of a non-homogeneous half-space. International Journal for Numerical Methods of Engineering, 57, 7, pp. 955–973.
6. A.J. Deeks and C.E. Augarde (2004a), A meshfree scaled boundary method. Submitted to *Computational Mechanics*.
7. X. Zhang, X. Liu, MW Lu and Y. Chen (2001), Computer Methods in Applied Mechanics and Engineering, Imposition of essential boundary conditions by displacement constraint equations in meshless methods. *Communications in Numerical Methods in Engineering*, 17, 3, pp. 165–178.
8. A.J. Deeks and C.E. Augarde (2004b), A hybrid meshless local Petrov–Galerkin method for unbounded domains. In preparation.

BASIC DISCUSSION OF BOUNDARY CONDITION OF SMOOTHED PARTICLE HYDRODYNAMICS FOR ANALYSIS OF CEREBRAL CONTUSION

S. Hagihara and S. Motoda
Department of Mechanical Engineering, Saga University, 1 Honjyo, Saga, Japan 840-8502

Abstract Analysis of cerebral contusion using the smoothed particle hydrodynamics (SPH) method is presented. When a human head is subjected to external impacts in cases of crashing with cars in traffic accidents, a person falling down and etc., damage is often observed in the brain. It is called cerebral contusion which means damage of the brain of the human head. Although cerebral contusion which is called coup is usually observed at an impacted side of the head, it may be also observed at an opposite side, which is called contrecoup. There is no reasonable explanation of contrecoup.

The smoothed particle hydrodynamics method, which is one of the meshfree methods, can calculate impact analyses including fracture.

In the present paper, we analyse the damage of the head using the SPH method and obtain the stress distribution of the brain when the head collides with a glass of a car. We obtain the distribution of stress to cause cerebral contusion.

Keywords: smoothed particle hydrodynamics, cerebral contusion, impact, human head.

1. INTRODUCTION

When a cars crush and a person falling down occur, the human head crashes to the front window shield or a ground. When the human head is subjected to external impact, the bone of the head is sometimes broken and the damage of the brain is caused due to the impact. Damages are often observed in the brain. It is called cerebral contusion which means damage of the brain of the human head. Although cerebral contusion is usually observed at an impacted side damage which is called coup, it may be also observed at an opposite side damage which is called contrecoup [1]. There is no reasonable explanation of contrecoup.

G. R. Liu et al. (eds.), Computational Methods, 1435–1439.

Smoothed particle hydrodynamics (SPH) was developed to be applied to astrophysical and cosmological problem Lucy and Gingold and Monaghan. After that, SPH had been studied in the field of compressible fluid dynamics. Recently, SPH was extended to simulate the dynamic structural problems including failure, which was proposed by Libersky et al. [2]

SPH does not require a grid and a special mesh, which is truly meshfree method based on the Lagrangian calculation. When the large deformation occurs, Eulerian calculation is used in the analysis. The Lagrangian calculation is more accurate than the Eulerian one because there is not the convection term in the equations written in the moving Lagrangian frame.

It is difficult to simulate failure and fracture phenomena for the finite element method. Since SPH is assembly of particles which are smoothed by a kernel function, it can also treat failure of structural problems. Therefore, SPH is appropriate to apply the analysis of cerebral contusion with breaking a skull or a front glass shield of a car.

In the present paper, SPH is applied to an analysis of cerebral contusion to analyse the cause of the damage in the brain including skull and font window shield when cars crash in a traffic accident.

2. METHOD OF ANALYSIS

The conventional formulation of SPH is interpolated by physical quantities of discrete particles smoothed using a kernel function. An approximate function $\langle f(\mathbf{x})\rangle$ at a point $\mathbf{x}$ is represented by a kernel function W which has a influence domain radius h shown in Figure 1 as following equation.

$$\langle f(\mathbf{x})\rangle = \int f(\mathbf{x}')W(\mathbf{x}-\mathbf{x}', h)dx' \tag{1}$$

where $\mathbf{x}-\mathbf{x}'$ is a smoothing distance within a domain influence. The discrete

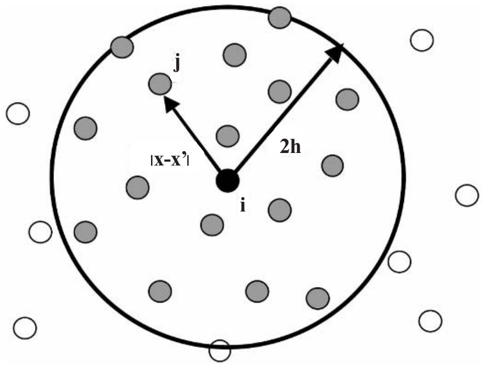


Figure 1. Typical concept of SPH.

form of the smoothed approximate function of Equation (1) is represented by the following equation.

$$\langle \nabla f(\mathbf{x}) \rangle = -\sum_{j=1}^{N} \frac{m_j}{\rho_j} f(\mathbf{x}_j) \nabla W(\mathbf{x} - \mathbf{x}', h) \tag{2}$$

The conservation equations of continuum are following set of discrete SPH equations.

$$\frac{d\rho_i}{dt} = \rho_i \sum_{j=1}^{n} \frac{m_j}{\rho_j} \left(U_i^{\beta} - U_j^{\beta} \right) W_{ij,\beta} \tag{3}$$

$$\frac{dU_i^{\alpha}}{dt} = -\sum_{j=1}^{n} m_j \left(\frac{\sigma_i^{\alpha\beta}}{\rho_i^2} + \frac{\sigma_j^{\alpha\beta}}{\rho_j^2} + \Pi_{ij} I \right) W_{ij,\beta} \tag{4}$$

$$\frac{dE_i}{dt} = -\sum_{j=1}^{n} m_j \left(U_i^{\alpha} - U_j^{\alpha} \right) \left(\frac{\sigma_i}{\rho_i^2} + \frac{1}{2} \Pi_{ij} I \right) W_{ij,\beta} \tag{5}$$

where U and σ are velocity and stress respectively. Π_{ij} is artificial viscosity shown as following equations. The conventional formulation of stress, pressure, time integration is used in the present analysis.

3. RESULTS AND DISCUSSION

A human head is subjected to impact when it collides with a front shield glass of a car in a traffic accident as shown in Figure 2. The figure is analysis model of collision between the human head and the glass. The human head collides with the glass in speed of 9 m/s (60 km/h). The material properties of the skull, the brain and the glass are summarized in Table 1.

The time dependence of distribution of Y-stress of the brain after the impact are shown in Figure 3. After the impact, the higher stress is observed due to

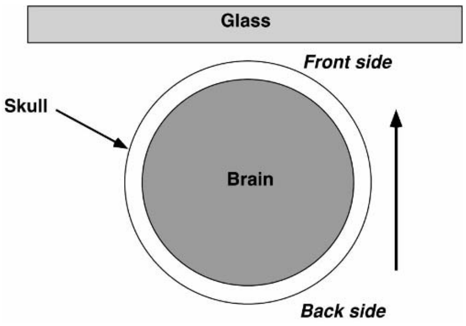
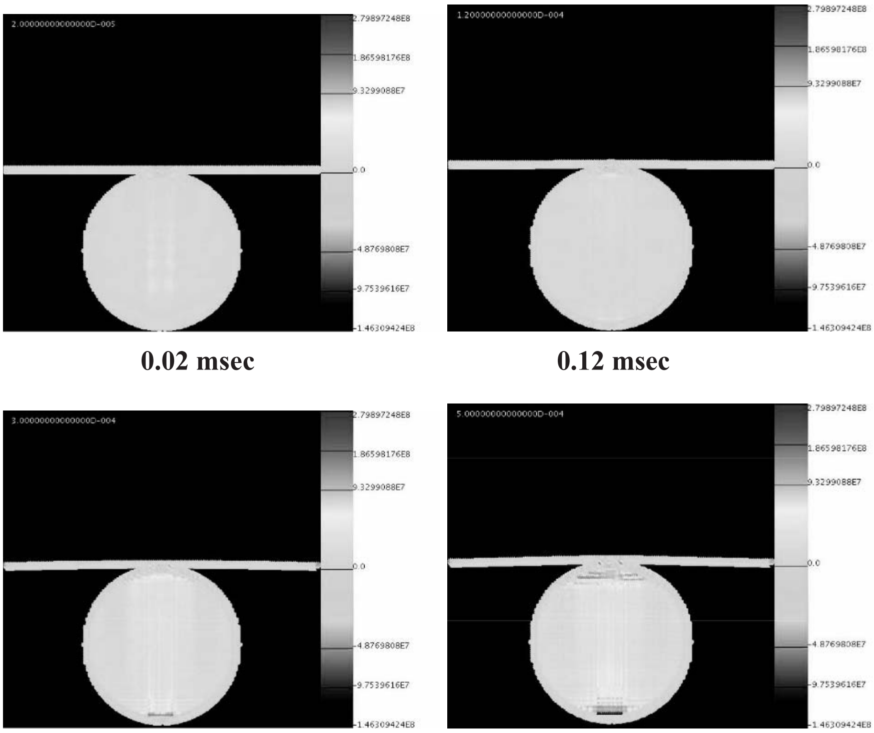


Figure 2. Schematic model of collision of head.

Table 1. Material properties.

	Brain	Skull	Glass
Density (kg/m^3)	1.456E + 03	1.456E + 03	2.5E + 03
Parameter C (m/s)	9.0E + 03	6.0E + 03	6.5E + 03
Parameter S (–)	1.5	1.5	1.5
Gruneisen parameter (–)	1.55	1.55	1.70
Shear modulus (Pa)	3.48E + 05	3.47E + 09	30.2E + 09
Young modulus (Pa)	11.14E + 05	8.35E + 09	74.1E + 09

the coup of the head. After the stress wave propagate to the opposite side of the head, the higher stress is caused by the contrecoup of the head in 0.3 ms. The cerebral contusion sometimes is observed due to not only coup but also contrecoup. It is consider the damage of the brain of the contrecoup can be simulated by the SPH method. The analysis performed by the SPH method will be able to take account of the failure criterion of glass and skull.



0.02 msec **0.12 msec**

0.30 msec **0.50 msec**

Figure 3. Time dependence of stress distribution of brain

4. CONCLUSIONS

The SPH method is applied to analysis for the damage of the brain when the human head is subjected to impact due to a traffic accident. The higher stress is observed in both the impacted side and opposite side. The SPH can be applied to the cerebral contusion in skull caused by coup and contrecoup. The SPH will be able to calculate the cerebral contusion including the failure of glass, skull and etc.

REFERENCES

1. S. Aomura, T. Ida, T. Ikoma and S. Fujiwara (2003), 3D finite element analysis of mechanism of cerebral contusion caused by external impact. In: *The 13th JSME Autumn Bioengineering Conference and Seminar*, pp. 133–134 (in Japanese).
2. L.D. Libersky, A.G. Petschek, T.C. Carney, J.R. Hipp and F.A. Allahdadi (1993), High strain Lagrangian hydrodynamics: a three-dimensional SPH code for dynamic material response. *Journal of Computational Physics*, 109, pp. 67–75.

DEPOSITION OF COLLOIDAL PARTICLES FROM PRESSURE DRIVEN MICROFLUIDIC FLOW—BROWNIAN DYNAMICS SIMULATION

H.N. Unni and C. Yang
School of Mechanical and Production Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798

Abstract This chapter investigated the irreversible adsorption of colloidal particles from pressure driven flow in a microchannel under the influence of colloidal (DLVO) interactions and external forces such as gravity. A theoretical model was developed on the basis of the stochastic Langevin equation, incorporating the Brownian motion of particles. Brownian dynamics simulation technique was employed to calculate the particle surface coverage. To validate the proposed theoretical model, experiments were carried out using the parallel plate flow cell technique, enabling direct Videomicroscopic observation of the deposition kinetics of polystyrene latex particles in NaCl electrolyte solution. The theoretical predictions were compared with experimental results and a good agreement was found.

Keywords: colloidal particle deposition, microchannel flow, Brownian dynamics simulation.

1. INTRODUCTION

Deposition of colloids, macromolecules, and bioparticles onto a solid surface is of great significance to many technological processes such as filtration, water treatment, biofouling of transplants and artificial organs etc. Hence it is vital to understand the underlying mechanisms in order to unravel and control these phenomena. There are two theoretical approaches of modeling deposition of colloidal particles onto the collector surface, namely Lagrangian method and Eulerian method. In the Eulerian method, the particle deposition onto the collector surface is governed by the convection diffusion equation, while the Lagrangian method determines the trajectory of the particle under the effect of colloidal and external forces. The governing equation is the stochastic Langevin equation, including the random Brownian motion of the particles. Adamczyk,

G. R. Liu et al. (eds.), Computational Methods, 1441–1445.

Siwek and Syzk [1] used the sequential Brownian dynamics simulation method to model the kinetics of particle adsorption in an impinging jet cell. Hutter [2] identified the coagulation time scales in colloidal suspensions using the Brownian dynamics simulation technique for various solid content ratios. In addition, Sholl et al. (2000) used Brownian dynamics simulations to numerically simulate TIRM (Total Internal Reflection Microscopy) experiments for the motion of a sphere in a viscous fluid near to a wall.

However, to the author's best knowledge, no attempt has been made yet to study the transport and deposition of colloidal particles from pressure driven flow in parallel plate microchannels, which possess interesting applications in BioMEMS (BioMicroelectromechanical systems). The goal of this chapter is therefore to investigate the kinetics of particle deposition in parallel plate microchannels using the Brownian Dynamics Simulation technique. The effect of electrolyte concentration and flow velocity on the particle surface coverage is studied.

2. THEORY

The colloidal particle trajectory is governed by the Langevin equation, given as

$$dr_i = \frac{\boldsymbol{D}_{ij}(t) \cdot \boldsymbol{F}_i(t)}{kT}\Delta t + \nabla \cdot \boldsymbol{D}_{ij}(t)\Delta t + \boldsymbol{U}(\boldsymbol{r}_i)\Delta t + (\Delta \boldsymbol{r})^B \tag{1}$$

where $\boldsymbol{D}_{ij}$ is the mutual diffusivity of the ith particle, $\boldsymbol{U}(\boldsymbol{r}_i)$ is the particle velocity, $\boldsymbol{F}_j$ is the total force acting on the particle, $(\Delta \boldsymbol{r})^B$ is the random Brownian displacement and kT is the thermal energy.

The mutual diffusivity of the particle is computed using the expression [3],

$$\boldsymbol{D}_{ij} = D_i I + kT \sum_{j=1}^{N} \boldsymbol{M}_{ij} \tag{2}$$

where D_i is the self diffusivity of the particle, I is the unit tensor and $\boldsymbol{M}_{ij}$ is the mobility tensor. N is the nearest neighbors to the ith particle. Tensor expressions for $\boldsymbol{M}_{ij}$ are given in [3]. The flow in the channel was assumed to be fully developed and the particle velocity was corrected from the fluid velocity by using hydrodynamic correction functions [3]. The total force on the particle, F_i is comprised of

$$\boldsymbol{F}_i = \boldsymbol{F}_{VDW} + \boldsymbol{F}_{EDL} + \boldsymbol{F}_G \tag{3}$$

where $\boldsymbol{F}_{VDW}$, $\boldsymbol{F}_{EDL}$ and $\boldsymbol{F}_G$ represent the DLVO (van der Waals, EDL) and gravity forces respectively. The expressions for the particle-wall and interparticle van der Waals and EDL forces were obtained from [3], and are given by

$$\boldsymbol{F}_{VDW} = \frac{-A\lambda a_p(\lambda + 22.22h)}{6h^2(\lambda + 11.11h)^2} + \sum_{j=1}^{N} \frac{-A_j\lambda a_p(\lambda + 22.22h_j)}{12h_j^2(\lambda + 11.11h_j^2)} \tag{4}$$

$$\boldsymbol{F}_{EDL} = 32\pi\varepsilon\kappa a_p \left(\frac{kT}{e}\right)^2 \tanh\left(\frac{ze\zeta_p}{4kT}\right) \tanh\left(\frac{ze\zeta_w}{4kT}\right)[2\exp(-\kappa h) + \exp(-\kappa h_j)] \tag{5}$$

A is the Hamaker constant, λ is the retardation wavelength, ε is the permittivity, κ is the Debye parameter (dependent on electrolyte concentration) and h and h_j are the particle—wall and interparticle separations, respectively. N is the number of nearest neighbors of a particle. ζ_p and ζ_w are particle and wall zeta potentials. These parameters for the simulation were selected from [4].

The Random Brownian motion of the particle was determined from a Gaussian distribution with zero mean and variance depending on the mutual diffusivity of the particle.

The particle transport was simulated in a cubic simulation cell containing 216 particles. Periodic boundary conditions [3] were applied on all the faces of the cell to eliminate the surface effects associated with the finite size of the cell. The particle positions were computed according to equation (1) using a modified Euler approach, described elsewhere [2]. Once the particle position has reached the primary energy minimum [1], it is assumed to be irreversibly adsorbed on the surface. The particle surface coverage is calculated using the expression,

$$\theta = \frac{\pi a_p^2 N_d}{\Delta S} \tag{6}$$

where N_d is the number of deposited particles, and ΔS is the surface area of the simulation cell. The dimensionless time (τ) is defined as the ratio,$\tau = t/t_s$, where t_s is the total simulation time.

3. EXPERIMENT

Videomicroscopic experiments were performed using monodisperse suspensions of latex particles (0.5 and 1 µm diameter, Duke scientific) in NaCl electrolyte solution flowing in a glass microchannel (300 µm depth). Particle images were captured by a digital camera (Leica DC300), which was connected to a PC. The image processing and particle counting were performed by software (Leica ImagePro). The particle surface coverage from experiments is

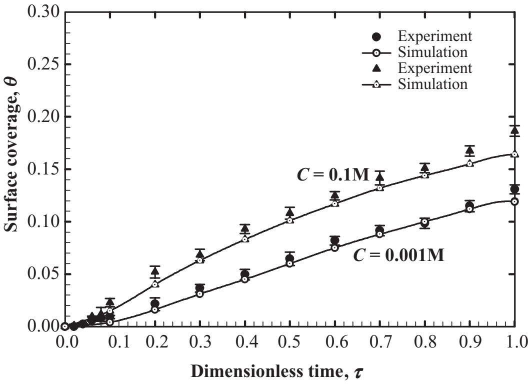


Figure 1. Surface coverage at $Re = 30$.

determined as

$$\theta = \frac{\pi a_p^2 N_c}{\Delta A} \tag{7}$$

N_c is the counted particle number and ΔA is the imaging area chosen in experiment.

4. RESULTS AND DISCUSSIONS

Figure 1 shows the effect of electrolyte concentration on particle deposition rate. As is evident from the figure, the surface coverage increases at higher electrolyte concentrations. This can be attributed to the repulsive double layer interactions between the similarly (negative) charged latex particles and glass

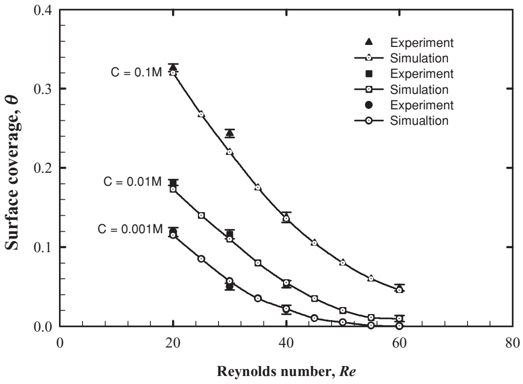


Figure 2. Surface coverage (vs) *Re plot*.

wall. The magnitude and range of EDL interaction decreases with increase in electrolyte concentration, resulting in higher deposition. The Simulation results are in good agreement with the experimental results, indicating the validity of the model approximation.

Figure 2 shows the variation of surface coverage with Reynolds number for three different electrolyte concentrations. The deposition rate decreases significantly with increase in Reynolds number. The difference in the surface coverage values at the three different concentrations is non-uniform at lower flow intensities, due to the reduced EDL interactions. At higher Reynolds numbers, the gap between the curves tend to decrease as convection predominates and most of the particles are transported without contacting the surface.

REFERENCES

1. Z. Adamczyk, B. Siwek and L. Szyk (1994), Flow induced surface blocking effects in adsorption of colloidal particles. *Journal of Colloid and Interface Science*, 174, pp. 130–141.
2. M. Hutter (1999), Brownian dynamics simulation of stable and coagulating colloids in aqueous suspension. *Ph.D dissertation*, Swiss Federal Institute of Technology, Zurich.
3. M. Elimelech, J. Gregoy, X. Jia and R.A. Williams (1995), *Particle Deposition and Aggregation—Measurement, Modeling and Simulation,* Butterworth-Heinemann, USA.
4. R.J. Hunter (2000), *Foundations of Colloid Science,* Oxford University press, UK.
5. D.S. Sholl, M.K. Fenwick, E. Atman and D.C. Prieve (2000), Brownian dynamics simulation of the motion of a rigid sphere in a viscous fluid very near a wall. *The Journal of Chemical Physics*, 113, pp. 9268–9278.

S SHAPE PARAMETERS OF MULTIQUADRICS IN THE HEAVISIDE WEIGHTED MLPG METHOD

J.R. Xiao[1], B.A. Gama1[1], J.W. Gillespie Jr.[1] and E.J. Kansa[2]
[1]*Center for Composite Materials, University of Delaware, Newark DE 19716, USA*
[2]*Department of Physical Sciences, Embry-Riddle Aeronautical University,*
7700 Edgewater Dr., Oakland, CA 94621

Abstract Radial basis functions (RBFs) have shown excellent interpolation properties and great promise in meshless methods for solving partial differential equations. However, the accuracy of the RBF meshless method is found to depend on the shape parameters, c, of RBFs; too large a value of c leads to severe ill-conditioning. In this paper, the selection of shape parameters of Multiquadrics (MQ) used in the Heaviside weighted MLPG meshless method has been investigated and a relationship between the parameter c and the nodal distance is proposed for solving the stress analysis of two-dimensional solids.

Keywords: MLPG meshless, multiquadrics, shape parameter, stress analysis.

1. INTRODUCTION

RBFs have been popularly employed in solving partial differential equations [1] and meshless methods [2–4]. The shape functions based on RBFs satisfy the delta function property. In the previous research [4], the extended multiquadrics $g(r) = (r^2 + c^2)^\beta$ for constructing trial functions was applied to the local Heaviside weighted MLPG method for the analysis of two-dimensional solids. There are many advantages in the so developed method: no domain integration is needed, no element matrix assembly is required and no special treatment is needed to impose the essential boundary conditions. However, we faced the problem of how to select a good value for the shape parameters to prevent severe illconditioning. Xiao and McCarthy [4] have conducted a preliminary study on selection of shape parameters of MQ in the Heaviside weighted

G. R. Liu et al. (eds.), Computational Methods, 1447–1451.

MLPG method. They found for the problems considered that both 1.03 and 1.99 are optimal values for the parameter β and the parameter c is sensitive to nodal distance. The present paper aims to provide further exploration on the selection of shape parameter.

2. LOCAL HEAVISIDE WEIGHTED MLPG FOR 2D SOLIDS

The two-dimensional solid mechanics problem is considered here:

$$\sigma_{ij,j} + b_i = 0, \qquad t_i = \sigma_{ij} n_j = \bar{t}_i \quad \text{on } \Gamma_t, u_i = \bar{u}_i \text{ on } \Gamma_u \tag{1}$$

A local weak form of Equation (1), over a local subdomain Ω_s bounded by Γ_s, can be obtained using the weighted residual method with a weighted function v_i:

$$\int_{\Omega_s} \sigma_{ij} v_{i,j}\, d\Omega - \int_{\Gamma_{si}} t_i v_i\, d\Gamma - \int_{\Gamma_{su}} t_i v_i\, d\Gamma = \int_{\Gamma_{st}} \bar{t}_i v_i\, d\Gamma + \int_{\Omega_s} b_i v_i\, d\Omega \tag{2}$$

where Γ_{st} is the intersection of Γ_t and the boundary Γ_s, and Γ_{su} is the intersection of Γ_u and the boundary Γ_s. When a Heaviside step function is used as the test function, the local weak form (2) can be rewritten as:

$$\int_{\Gamma_{si}} t_i\, d\Gamma - \int_{\Gamma_{su}} t_i\, d\Gamma = \int_{\Gamma_{st}} \bar{t}_i\, d\Gamma + \int_{\Omega_s} b_i\, d\Omega \tag{3}$$

It can be seen that only the regular boundary integral along the sub-domains boundaries is involved in the weak form (3).

3. DISCRETIZATION USING RADIAL BASIS FUNCTIONS

An interpolation of a continuous function $u(\mathbf{x})$ from the neighbouring nodes of a point $\mathbf{x}_Q$ within the domain Ω ,using RBFs with a polynomial basis can be given as:

$$u(\mathbf{x}) = \sum_{i=1}^{n} g_i(\mathbf{x}) \xi_i(\mathbf{x}_{\mathrm{Q}}) + \sum_{j=1}^{m} p_j(\mathbf{x}) \zeta_j(\mathbf{x}_{\mathrm{Q}}) = G^T \xi + P^T \zeta \tag{4}$$

with the constraint condition $\sum\limits_{i=1}^{n} p_j(x_i, y_i)\xi_i = 0$, $j = 1, 2, \ldots m$, where $g_i(\mathbf{x})$ is the radial basis function, $p_j(\mathbf{x})$ is a monomial.

The constructed shape function $\phi_k(\mathbf{x})$ and their derivatives using the extended *Multiquadrics* (MQ) can be found in Xiao and McCarthy [4]. Substitution of interpolation into the weak form (3) yields the following discrete equation for each node:

$$\sum_{j=1}^{n} K_{ij} u^j = f_i \tag{5}$$

4. NUMERICAL EXAMPLES AND SELECTION OF SHAPE PARAMETER *C*

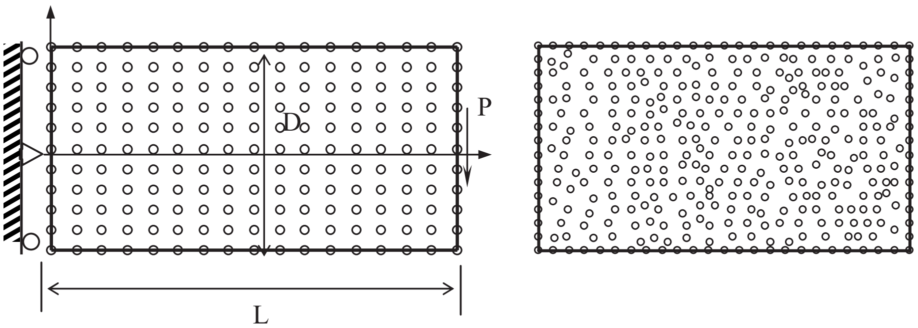
For the purposes of error estimation, the displacement norm and the energy norm are defined as follows:

$$\text{displacement norm} = \left\{ \int_\Omega (u^{Num} - u^{Exact})^T (u^{Num} - u^{Exact})\, d\Omega \right\}^{1/2} \tag{6}$$

$$\text{energy norm} = \left\{ \frac{1}{2} \int_\Omega (\varepsilon^{Num} - \varepsilon^{Exact})^T (\sigma^{Num} - \sigma^{Exact})\, d\Omega \right\}^{1/2} \tag{7}$$

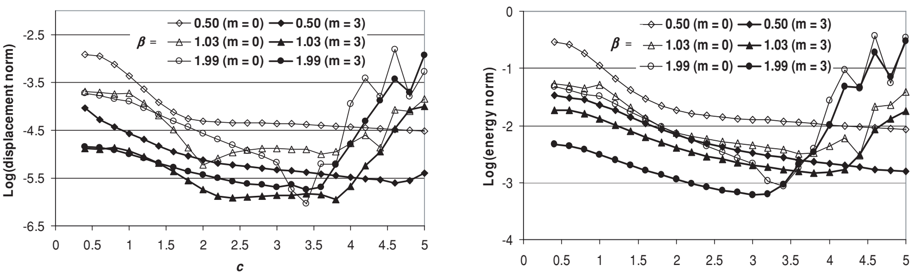
In this section, effects of shape parameters of MQ on the performance of the present method and the selection of the parameter c are investigated with the cantilever beam problem shown in Figure 1(a). The analytical solution is available in Timoshenko and Goodier [5]. The beam parameters are taken as $E = 30.0 \times 10^6$, $\nu = 0.3$, $D = 5$, $L = 8$ and $P = 1000$.

The previous study [4] shows that the sensitivity of c to nodal distance can be demonstrated by a linear relationship between c and nodal distance: $c = \lambda d$, where λ is a coefficient to be determined and d is the characteristic



(a) Regular 187 nodes (b) Irregular 353 nodes

Figure 1. Cantilever beam with a parabolic-shear end load and different nodal arrangements.



(a) Error in displacement norm (b) Error in energy norm

Figure 2. Effect of shape parameter c in cantilever beam (187 nodes).

nodal distance. Figure 2 shows the effects of varying c on displacement norm and energy norm with 187 regularly spaced nodes and the characteristic nodal distance d is 0.5 here. The results of relevance here are the $m = 0$ values given by the open symbols (no polynomial terms) and $m = 3$ for the effects of (only linear) polynomial terms. From Figure 2, we find that values of λ ranging from 4 to 7 result in a displacement norm of less than 0.0000315 [i.e. Log(displacement norm) less than -4.5] and an energy norm of less than 0.01 [i.e. Log(energy norm) $= -2$]. It is also seen that polynomial terms do not change optimal values for parameter c but increase the accuracy for all β values used. Further study (not shown here) with d reduced to 0.25 resulting in refined nodal pattern gives similar result that better accuracy is achieved when λ ranges from 4 to 7. The beam problem with an irregular node distribution as shown in Figure 1(b) was also analysed. After several calculations using different irregular nodal patterns, we found that the accuracy is more sensitive to the minimum nodal distance d_{min}. In this case, $d_{min} = 0.167$ is used as the characteristic nodal distance, one has good results for λ ranging from 4.1 to 7.6 which is remarkably consistent with the values extracted from the cases with regular nodes.

5. CONCLUSION

The selection of shape parameters of MQ in the Heaviside weighted MLPG method has been presented. For any given value of β, the higher the parameter c, the greater the condition number. However, if the condition number becomes too high it may result in unstable solutions. The preconditioning developed by Ling and Kansa [6] may provide better solution of the present method, which is under investigation. The effect of polynomial terms has also been studied and it has been found that the accuracy can be increased substantially, and there is minor effect of polynomial terms on the selection of the parameter c. In this research we assume the problem domain is represented by properly scattered nodes, i.e. the nodal pattern and the ratio between the maximum nodal distance and the minimum nodal distance are reasonable. Favorable accuracies were found for c values in the range of $5d$–$7d$, and $c = 6d$ is recommended for two-dimensional solids.

REFERENCES

1. E.J. Kansa (1990), Multiquadrics–A scattered data approximation scheme with application to computational fluid-dynamics–I and II. *Computer and Mathematics in Application*, 19, pp. 127–161.
2. J.G. Wang and G.R. Liu (2002), On the optimal shape parameters of radial basis functions used for 2-D meshless methods. *Computer Methods in Applied Mechanics and Engineering*, 191, pp. 2611–2630.
3. S.N. Atluri and S. Shen (2002), The Meshless Local Petrov-Galerkin (MLPG) approach for solving problem up elastic-statics, *Computational Machanics*, 25, pp. 169–179.
4. J.R. Xiao and M.A. McCarthy (2003), A local Heaviside weighted meshless method for two-dimensional solids using radial basis functions. *Computational Mechanics*, 31, pp. 301–315.
5. S.P. Timoshenko and J.N. Goodier (1970), *Theory of Elasticity*, 3rd ed. McGraw-Hill, New York.
6. L. Ling and E.J. Kansa (2005), A least-squares preconditioner for radial basis functions collocation methods. *Advances in Computational Mathematics*, 23, pp. 31–54.

APPLICATION OF HIGH ORDER BASIS FUNCTIONS IN SOLID MECHANICS BY ELEMENT FREE GALERKIN (EFG) METHOD

A. Kiasat[1], M. Moradi[1] and H. Hashemolhosseini[2]
[1]*Department of Mechanical Engineering, Isfahan University of Technology, Isfahan, 84156, Iran*
[2]*Department of Mining Engineering, Isfahan University of Technology, Isfahan, 84156, Iran*

Abstract The aim of this research is to investigate the effect of basis functions order, on the error norm, in EFG method. For this purpose, a cantilever elastic beam is considered. To get rid of the singularity, the essential boundary conditions, based on analytical solution, are applied by using both Lagrange Approach and Penalty Function, separately. In this research, the effect of some parameters like radius of support, order of basis functions, number of points within the domain of influence and density of background mesh on the computational error has been investigated. The sensitivity of the accuracy criteria, i.e. energy norm and displacements, corresponding to the different values of penalty factor and also the order of basis functions has been examined.

Keywords: EFG method, basis functions order, essential B.C., Penalty Function.

1. INTRODUCTION

Finite Element Method (FEM) is robust and has been thoroughly developed for many practical engineering problems related to solids and structures, as well as fluid flows. However, there are some difficulties with this method. A close consideration of these problems reveals that the root of these difficulties is in using elements, which are the building block of FEM. Therefore, the idea of eliminating the elements and hence the mesh has been proposed, in which the domain of the problem is represented by a set of scattered nodes (regular or irregular). Element Free Galerkin method (EFG) is one of the most important methods, developed on Diffuse Element Method (DEM) by Belytschko [1]. DEM was originated by Nayroles [2]. In EFG method, MLS approximation is

G. R. Liu et al. (eds.), Computational Methods, 1453–1457.

employed for constructing the shape functions [3]. Galerkin weak form is used to develop the discretized system equations. In using EFG method, essential boundary conditions cannot be enforced directly and easily. For overcoming this problem, some techniques such as Lagrange Multiplier Approach (LMA) or Penalty Function are used. In most published papers, for example [2–5], the use of LMA and also first order basis functions are emphasized. In this work, the main purpose is to investigate the effect of increasing the order of basis functions on errors. As a numerical example, a cantilever elastic beam is considered. To impose essential boundary conditions, Penalty Function and LMA are implemented separately and alternately for comparing the results to each other. To evaluate errors, both the deflection of beam at the free end and Error in Energy Norm criteria have been used for obtaining a better view on the total error within the domain of problem.

2. MOVING LEAST SQUARE APPROXIMATION (MLS)

In MLS approximation, a set of points in the vicinity of a selected node is used to approximate field variable(s) at that point. For this purpose, a set of linear independent functions is used to make the local approximation around the central point. These series of functions, called basis functions, are usually selected from complete order polynomials. Selecting proper weight function and also basis functions are two essential items in MLS approximation. Weight function is selected so that at a limited domain around the nodes, called the domain of influence, to be non-zero. The domain of influence can be either circular or rectangular. In this paper, circular support along Cubic Spline weight functions are used to produce the shape functions. In MLS technique, the unknown coefficients of the basis functions are computed by minimizing the difference between the approximated values of the field function and the nodal parameters [3, 4].

3. ENFORCING BOUNDARY CONDITIONS IN EFG METHOD

As mentioned above, in EFG method MLS approximation is used to form shape functions. These shape functions do not satisfy Kronecker Delta function properties. This problem arises some difficulties in directly enforcing the essential boundary conditions. For removing this problem the following solutions are usually suggested:

A. Applying the Constrained Galerkin Weak Form by using Lagrange Multiplier Approach [1, 4].

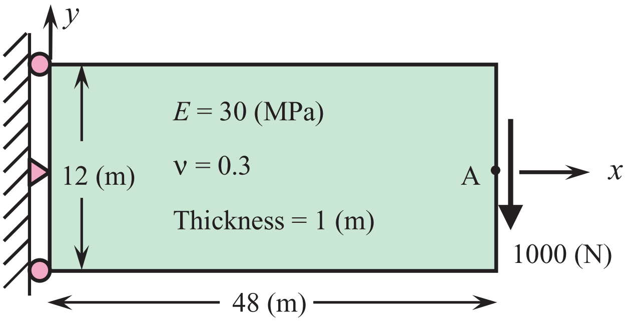


Figure 1. Cantilever beam. V_A = Vertical displacment at end (A).

B. Applying the Constrained Galerkin Weak Form by using Penalty Function [5, 6].

Each of above techniques can be used to result in the discretized system equations. EFG method, like the other Mesh Free methods, uses a set of nodes scattered within the problem domain as well as sets of nodes scattered on the boundaries of the domain to represent the problem domain and its boundaries. For this aim, the shape functions, obtained by MLS approximation, are substituted in to the Constrained Galerkin Weak Form. By using EFG method, therefore partial differential governing equations can be approximated by a set of algebraic equations for each node. The system of algebraic equations for the whole problem domain can be formed by assembling sets of algebraic equations for all integral points [4, 5].

4. NUMERICAL EXAMPLE

In this paper, a shear force with parabolic distribution is enforced at the free end of a cantilever beam. All used parameters are shown in Figure 1. The analytical solution of this problem has been given in [4]. To apply the EFG method for this problem, a uniform rectangular background mesh has been used while a monotonic sample point distribution is adapted to it. Actually, increasing the number of nodes causes the density of background mesh, to be increased as well.

5. RESULTS AND DISCUSSION

In this paper, the penalty factor (P.F.) in X and Y direction is considered to have the same value. To obtain a dimensionless value of Error in Energy Norm

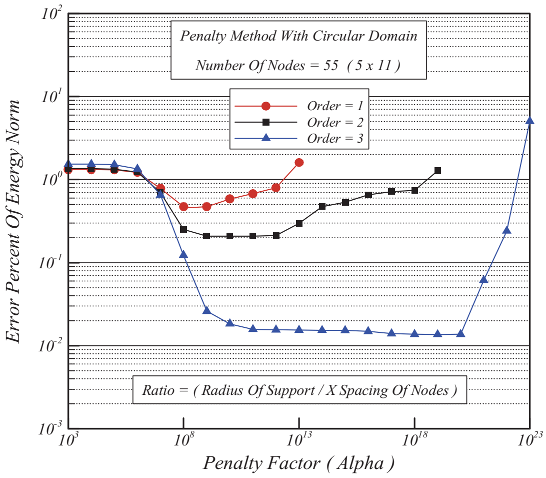


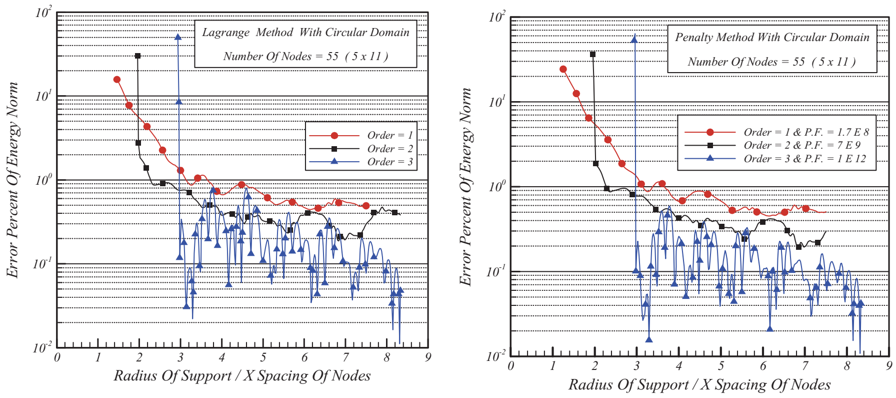
Figure 2. Error percent of energy norm v.s. penalty factor for different orders.

criterion, the following formula is used:

$$\textit{Error Percent of Energy Norm} = \left\{ \int_{\Omega} \frac{(\boldsymbol{\varepsilon}^{Num} - \boldsymbol{\varepsilon}^{Exact})^{\mathrm{T}} \mathbf{D} (\boldsymbol{\varepsilon}^{Num} - \boldsymbol{\varepsilon}^{Exact})}{(\boldsymbol{\varepsilon}^{Exact})^{\mathrm{T}} \mathbf{D} \boldsymbol{\varepsilon}^{Exact}} \mathrm{d}\,\Omega \right\}^{1/2} \times 100 \tag{1}$$

According to the obtained results, the following cases are concluded:

1. As shown in Figure 2, increasing the order of the basis function decreases the sensitivity of responses relative to changes in penalty factor and also increases the accuracy of solution.



(3-a) **(3-b)**

Figure 3. Error percent of energy norm v.s. Ratio for different orders.

Table 1. Effect of basis functions order on error percent of V_A ($3 < \text{Ratio} < 6$).

	Order	1	2	3
55 nodes (5 × 11)	Minimum Error Percent	5.5E-1	5.2E-4	6.8E-5
	Average Error Percent	8.3E-1	6.5E-2	2.3E-2
85 nodes (5 × 17)	Minimum Error Percent	8.6E-1	1.4E-3	6.5E-5
	Average Error Percent	9.8E-1	3.2E-2	2.3E-3
175 nodes (7 × 25)	Minimum Error Percent	1.0E0	7.3E-4	1.1E-6
	Average Error Percent	1.0E0	1.7E-2	1.4E-3

2. According to Figure 3a and b, by increasing the order of basis function, it is possible to obtain the maximum available accuracy. This result is achieved in a less support radius ratio. With selecting a proper value for penalty factor, the accuracy of solution is increased and converges to the ones corresponding to the Lagrange Multiplier Approach.
3. As shown in Table 1, by increasing the order of the basis function, it is possible to obtain a high accuracy with rather less nodes. For a better understanding, the Table 2 in [4] can be compared with the Table 1.

REFERENCES

1. T. Belytschko, Y.Y. Lu and L. Gu (1994), Element free Galerkin methods. *International Journal for Numerical Methods in Engineering*, 37, pp. 229–256.
2. B. Nayroles, G. Touzot and P. Villon (1992), Generalizing the finite element method: diffuse approximation and diffuse elements. *Computational Mechanics*, 10, pp. 307–318.
3. P. Lancaster and K. Salkauskas (1981), Surfaces generated by moving least squares methods. *Mathematics of Computation*, 37, pp. 141–158.
4. J. Dolbow and T. Belytschko (1998), An introduction to programming the meshless element-free Galerkin method. *Archives of Computational Mechanics in Engineering*, 5, pp. 207–241.
5. G.R. Liu (2003), *Mesh Free Methods: Moving Beyond the Finite Element Method*, CRC Press, Boca Raton. USA.
6. J.G. Wang and G.R. Liu, (2001), A point interpolation method for simulating dissipation process of consolidation. *Computer Methods in Applied Mechanics and Engineering,* 190, pp. 5907–5922.

AN ADAPTIVE MESHFREE COLLOCATION METHOD FOR STATIC AND DYNAMIC NONLINEAR PROBLEMS

G.R. Liu, D.B.P. Huynh and Y.T. Gu
Department of Mechanical Engineering, National University of Singapore, 9 Engineering Drive 1, Singapore 117576

Abstract In this paper, an adaptive algorithm using meshfree strong-form collocation methods for nonlinear partial differential equations is proposed. The meshfree method uses polynomial point interpolation method and radial point interpolation method to form shape functions, and applicable for both static and dynamic problems. The adaptivity of the scheme relies on rules of both refinement and coarsening of scattered nodes. An error estimation based on solution interpolation is used for static and time-dependent partial differential equations. The nodal refinement scheme in each adaptive step is performed using Delaunay triangulation and Voronoi diagram. The numerical examples confirm the good performance of the present adaptive meshfree collocation method.

Keywords: meshfree methods, adaptive methods, radial basis functions.

1. INTRODUCTION

Meshfree methods have been proposed and achieved remarkable progress in recent years [1, 2]. As its name implies, meshfree methods evaluate field variables entirely based on a group of discrete nodes and require no predefined nodal connectivity. In meshfree methods, nodes in the problem domain are unstructured. They can be inserted, moved or deleted conveniently—this advantage makes these methods extremely attractive for adaptive analysis.

There are some adaptive meshfree methods have been developed so far. For example, Liu [1] has proposed an adaptive procedure based on weak-form meshfree methods for static problems. Behrens et al. [3, 4] have developed

G. R. Liu et al. (eds.), Computational Methods, 1459–1464.

an adaptive meshfree method based on strong form and characteristic line approach for dynamic problems. In this paper, we present a general adaptive meshfree method based on strong-form formulation for both static and dynamic problems. The algorithm uses a meshfree point collocation method and radial basis functions created using nodes in local support domains. Error estimation is performed based on solution interpolation at each step to adaptively modify nodes at each step. Numerical examples are presented to demonstrate the present procedure.

2. RADIAL POINT INTERPOLATION METHOD (RPIM)

Consider a field function $u(\mathbf{x})$ defined in a domain Ω, which is represented by a set of N nodes at $\mathbf{x}_i (i = 1, 2, \dots, N)$. It is assumed that only the surrounding nodes in the local domain around the point $\mathbf{x}_Q$ have effects on $u(\mathbf{x})$. This local domain is called the support domain of $\mathbf{x}_Q$. RPIM interpolates $u(\mathbf{x})$ using its values at nodes in it local support domain and radial basis functions augmented with polynomial reproduction in the following form

$$u^h(\mathbf{x}, \mathbf{x}_Q) = \sum_{i=1}^{n} r_i(\mathbf{x}) a_i + \sum_{j=1}^{m} p_j(\mathbf{x}) b_j = \mathbf{r}^T(\mathbf{x})\mathbf{a} + \mathbf{p}^T(\mathbf{x})\mathbf{b} \tag{1}$$

where $r_i(\mathbf{x})$ and $p_j(\mathbf{x})$ are, respectively, radial and monomial basis functions of coordinates $\mathbf{x}$; n is the number of nodes in the local support domain of $\mathbf{x}_Q$; m is the number of selected monomial terms. $m = 3$ is used to construct shape function with linear polynomial reproduction. There are a number of radial basis functions. In this paper, we use two radial basis functions:

Thin plate spline (TPS):	$r_i(x, y) = r^\eta$	(2)
Logarithm thin plate spline (LTPS):	$r_i(x, y) = r^\eta \log r$	(3)

Follow the standard procedure described in Section 5.7 in the meshfree method book by Liu [1], we arrive at

$$u^h(\mathbf{x}) = \mathbf{\Phi}(\mathbf{x})\mathbf{U}_s \tag{4}$$

where $\mathbf{\Phi}(\mathbf{x})$ is the matrix of shape functions, and $\mathbf{U}_s$ is the nodal values for all the nodes in the support domain.

Performing simple collocation at all nodes in the problem domain using Equation (4), we finally obtain

$$\mathbf{KU} = \mathbf{F} \tag{5}$$

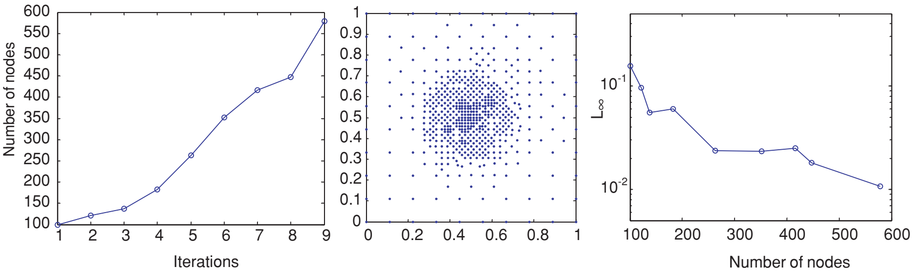


Figure 1. Number of nodes at each iteration (left), final nodal distribution (middle) and the error L_∞ at each iteration (right).

where **K** is the global stiffness matrix, $\mathbf{U} = \{u_1, u_2, \ldots, u_N\}^T$ is the nodal values for all the nodes in the entire problem domain, and **F** is the nodal 'force' vector. Since local support domains are used, the matrix K is sparse and banded.

Solving Equation (5), one can obtain the values of the field variable at all the nodes.

3. ADAPTION RULES

The adaptive algorithm uses the following error estimate based on solution interpolation to perform nodes coarsening and refinement at each step [4].

$$\eta_i = \left| u_i^{\bar{N}} - u_i \right| \tag{6}$$

where u_i is the value of a field variable $\mathbf{x}_i$; $u_i^{\bar{N}}$ is the value of the field variable obtained by an interpolation using nodal values at the nodes in the domain

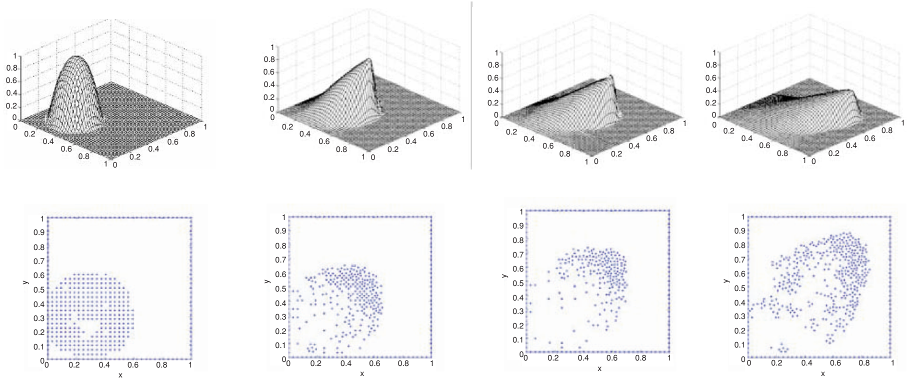


Figure 2. Solution and nodal distribution of Burger equation at $t = 2\tau, 40\tau, 80\tau, 160\tau$.

excluding node i, i.e., $\bar{N} \equiv N \backslash \{\mathbf{x}_i\}$. We note that, if $\bar{N} \equiv N$, then $\eta_i \equiv 0$. In addition, the value η will also vanish whenever u is linear around $\mathbf{x}$.

In order to balance the accuracy of the solution and the computational cost, we add new nodes into regions where value of η is large (refinement), and remove nodes from regions where value of η is small (coarsening). In our refinement procedure, new nodes are either added directly on Voronoi points or on the middle of edges or centre of Delaunay triangles.

4. NUMERICAL EXAMPLES

Example 1. (*'High gradient' 2D problem*)

We consider the following Poisson's equation

$$\nabla^2 u(x, y) = \left[-400 + (200x - 100)^2 + (200y - 100)^2\right] e^{-100(x-0.5)^2 - 100(y-0.5)^2} \tag{7}$$

defined in the square domain $\Omega \equiv [0, 1] \times [0, 1]$, with boundary condition of $u(x, y) = 0$ on all boundaries.

We use TPS radial basis function in this problem. The initial nodes are at 10×10 regularly distributed. At each refinement, we consider to refine up to 5% the number of available nodes. No coarsening rule was applied and the refinement rule is based on Voronoi diagram. The stopping criteria is set at $L_\infty \leq 1 \times 10^{-2}$, which normally can be obtained by a 729 nodes regularly distributed using the same meshfree method. The final solution is achieved after 10 iterations and $L_\infty = 1 \times 10^{-2}$. The final mesh contains only 579 irregular distributed nodes.

Example 2. (*Burger's equation and Buckley–Leverett equations*)

Burger's equation and Buckley–Leverett equation have the form

$$\frac{\partial u}{\partial t} + \nabla f(u) = 0 \tag{8}$$

the flux term $f(u)$ is defined as $f(u) = \frac{1}{2}u^2 \cdot r$ and $f(u) = \frac{u^2}{u^2 + \mu(1-u)^2} \cdot r$ (Burger equation and Buckley–Leverett equation, respectively), and r is the direction of the flux. These problems have also been studied by Behrens et al. [4]. For these problems, we prefer to work with an alternative form, so that our solution will be entropy-satisfied.

$$\frac{\partial u}{\partial t} + \nabla f(u) = \varepsilon \cdot \Delta u \tag{9}$$

where $\varepsilon > 0$ is the artificial viscosity coefficient. We prefer to work with $\varepsilon = 8 \times 10^{-3}$ for Burger equation and $\varepsilon = 4 \times 10^{-3}$ for Buckley–Leverett equation.

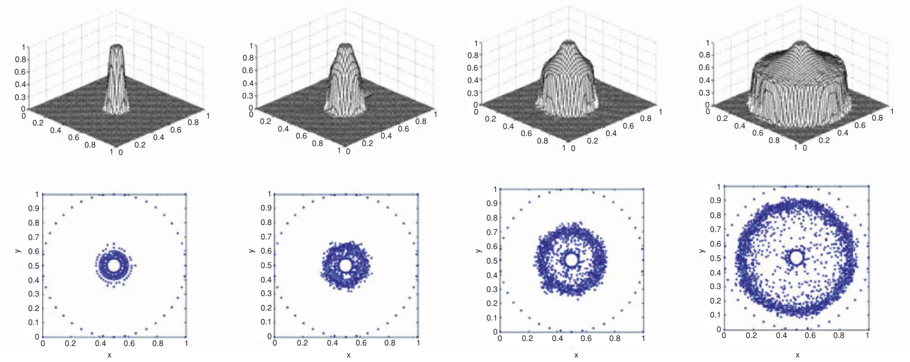


Figure 3. Solutions and nodal distribution of Buckley–Leverett equation at $t = 15\tau, 66\tau, 132\tau, 264\tau$.

The LTPS radial basis function is used for these problems. We note that more nodes were added in the model in order to capture the moving shock. The result is compared with reference solutions from Behrens et al. [4] and good agreed is observed.

Due to the appearance of the shock wave, regular nodal distribution has failed to give acceptable solutions for both above examples. The solution is unstable after a first few steps after the shock wave has been formed. We can also use a very finer model for these problems, but the cost of computation for such model is very expensive.

5. CONCLUSIONS

The developed adaptive procedure using strong-form meshfree collocation methods gave very good solutions for high-gradient problem (static problems) and moving-shock problems (time-dependent problems). Numerical examples confirm the good performance of the present procedure.

REFERENCES

1. G.R. Liu (2002), *Meshfree Methods: Moving Beyond the Finite Element Method*. CRC Press.
2. G.R. Liu and M.B. Liu (2003), *Smooth Particle Hydrodynamics: A Meshfree Particle Method*. World Scientific Publishing, Singapore.
3. J. Behrens, A. Iske and S. Pöhn (2001), Effective node adaption for grid-free semi-Lagrangian advection. In: *Discrete Modelling and Discrete Algorithms in Continuum Mechanics*, T. Sonar and I. Thomas (eds.), Logos, Berlin, pp. 110–119.

4. J. Behrens, A. Iske and M. Käser (2002), Adaptive meshfree method of backward characteristics for nonlinear transport equations. In: *Meshfree Methods for Partial Differential Equations*, M. Gribel and M.A Schweitzer (eds.), Springer-Verlag, Heidelberg, pp. 21–36.
5. G.R. Liu and Z.H. Tu (2002), An adaptive procedure based on background cells for meshless methods. *Computer Methods in applied Mechanics and Engineering*, 191, pp. 1923–1943.

MESHLESS NATURAL NEIGHBOUR METHOD AND ITS APPLICATION IN ELASTO-PLASTIC PROBLEMS

H.H. Zhu, Y.B. Miao and Y.C. Cai
Department of Geotechnical Engineering, School of Civil Engineering, Tongji University, 200092,P. R. China

Abstract The meshless natural neighbour method (MNNM) is a truly meshless method, which does not need the Delaunay tessellation of the whole domain to construct the Laplace interpolation. At the same time, some difficulties in other meshless methods, such as the imposition of essential boundary conditions, the treatment of material discontinuities and the choice of weight functions are avoided. The governing equations of elasto-plastic for MNNM are obtained to apply the MNNM to the analysis of two-dimensional elasto-plastic problems. The numerical results indicate that the theory and programmes are accurate and effective.

Keywords: meshless, natural neighbour, natural element, elasto-plastic analysis.

1. INTRODUCTION

Meshless methods are newly developed numerical methods in the last decade. As the approximation/interpolation functions are constructed in terms of scattered nodes, they do not need any mesh structures in the formulation. Because of the mesh-free property, some problems brought by mesh generation, mesh distortion and mesh motion are avoided. More and more engineers and computational mechanics researchers have paid their attentions to this kind of methods.

According to the definition of the shape functions, meshless methods can be mainly grouped into three classes: meshless methods based on kernel function approximation [1–4,], meshless methods based on least square approximation [5–13], and meshless methods based on natural neighbour interpolation [14–20]. In addition, there are some more general meshless methods such as

G. R. Liu et al. (eds.), Computational Methods, 1465–1475.

meshless method based on collocation with radial basis function developed by Zhang *et al.* [21]. They are not ranged into the classification, but the aforementioned three types of meshless methods can be deduced from them.

The representative method of the first class of meshless methods is the Smooth Particle Hydrodynamics (SPH) [1], which was proposed by Lucy and Gingold in 1977. This method has high computational efficiency, but as a result of collocation method used in the discretization of equilibrium equation, the solution of SPH is instable. To improve the stability of SPH, Liu *et al.* developed a modified kernel function method, the Reproducing Kernel Particle Method (RPKM) [2].

The family of meshless methods based on least square approximation has the most members, which is represented by the Element-free Galerkin Method (EFG) developed by Belytschko *et al.* [5]. The family also includes the Finite Point Method (FPM) [9], Hp Clouds method [10], and Local Boundary Integral Equation method (LBIE) [11] etc. These methods have better compatibility and stability. But due to more time consuming, this kind of methods are not so preferable as SPH-like methods in nonlinear large deformation analysis field.

The above-mentioned two classes of methods are the most popular meshless methods at present. Compared to conventional numerical methods (such as FEM), they have many advantages. Unfortunately, many difficulties arise in the computational time, the enforcement of boundary conditions and the treatment of discontinuities, which limit the research within theoretic scope.

The third type of meshless methods are based on the natural neighbour concept to define the shape function. The representative one is Natural Element Method (NEM) proposed by Braun and Sukumar [13, 14]. By constructed in meshless way, the natural neighbour shape function shares several advantanges of the FEM, thus avoids the drawbacks of the previous meshless methods. However, the Delauney tessellation of the distinct points in the whole analysis region is needed to construct the natural neighbour shape function, NEM can hardly be recognized as a meshless method. There are fewer research papers about NEM, and it is not as well known as some other meshless methods, such as EFG and SPH.

Recently, a truly meshless method, Meshless Natural Neighbour Method (MNNM) was proposed by our research group. This method adopts the means similar to EFG to seek the natural neighbour points of the integral points and the Delaunay tessellation of the whole region is avoided consequently. Therefore, its shape function takes full advantages of natural neighbour shape function and almost all the advantages of EFG. The objective of the present work is to apply the MNNM to the analysis of two-dimensional elasto-plastic problems. The governing equations of elasto-plastic for MNNM are described, and results of several standard examples are compared with FEM or analytical solutions to verify this.

2. THEORY OF MESHLESS NATURAL NEIGHBOUR METHOD

2.1 Search for natural neighbours

In order to solve the differential equations of boundary value problems in MNNM, a set of distinct nodes $N = \{n_1, n_2, \cdots, n_m\}$ should first be set up at the arbitrary geometry shape of domain Ω (see Figure 1).

Suppose that sample point $p(\mathbf{x})$ is an arbitrary numerical integral point of domain Ω, the algorithm for the neighbour-search in MNNM is based on the locally Delaunay triangles. Let the initial influence nodes $M = \{m_1, m_2, \ldots\}$ of the point p be confined within the dashed lines of the square as shown in Figure 2.

The next step is to directly determine the natural neighbours of the point p by using the empty circumcircle criterion—if $\mathrm{DT}(n_I, n_J, n_K)$ is any Delaunay triangle of the nodal set N, then the circumcircle of DT contains no other nodes of N. Find the node 1 which is the nearest to the sample point p from the nodal set M. Starting with edge $p - 1$ and using the empty circumcircle criterion, we form a set of locally defined triangles $\{p, 1, 2\}$, $\{p, 2, 3\}$, $\{p, 3, 4\}$, $\{p, 4, 5\}$, and $\{p, 5, 6\}$, where the nodes 1–6 are selected from the nodal set M in Figure 2. Now, the nodes 1–6 are just the natural neighbours of the point p..

The natural neighbours of the given point p are unique after the nodal set N has been set up at the domain Ω. The size of the square edges $2r$ in Figure 1 will not and cannot influence the definition of the natural neighbours of point p. The purpose of the restriction of the influence nodes to the square region is to reduce the time for searching for the natural neighbours. Hence, the size of

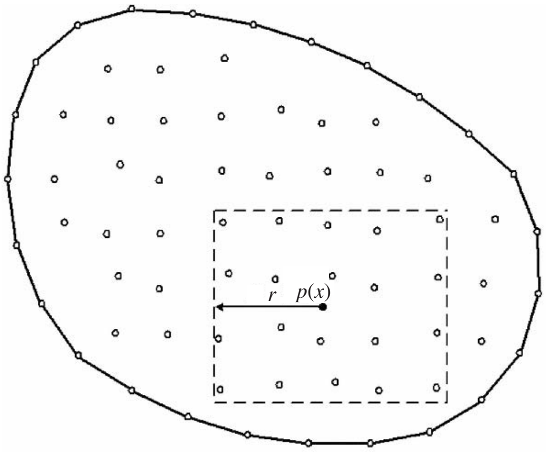


Figure 1. Discrete model of region Ω and its arbitrary integrate point p(x).

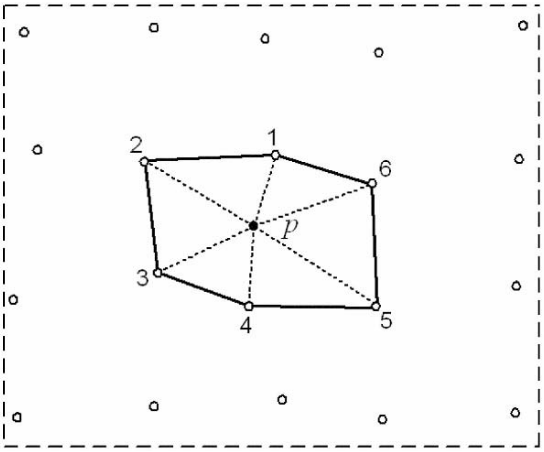


Figure 2. Natural neighbour of the point p.

the length r must be large enough to contain all the natural neighbours of point p, and should be small enough to save the time consume of the neighbour-search. Of course, the nodal set N of domain Ω can be regarded as the initial influence nodal set M of the point p, but the neighbour-search will be much more expensive and can not be afforded.

2.2 Laplace interpolation

Natural neighbours provide a means to define a robust approximation for scattered nodes in Ω according to the relative spatial density and position of nodes. Silbson and Laplace (non-silbson) natural neighbour interpolations are differently used in NEM and NNM. In 2D, the Laplace shape function defined by the ration of length measures [16], whereas the Silbson shape function is based on the ration of areas [14]. The compuational cost and algorithm are more favorable in Laplace interpolant than in Silbson interpolant. In this paper, we choose Laplace interpolant to develop the MNNM.

In Figure 3,C_{12p}, C_{23p}, C_{34p}, C_{45p}, and C_{61p} are the circumcentres of the triangles $\{p, 1, 2\}$, $\{p, 2, 3\}$, $\{p, 3, 4\}$, $\{p, 4, 5\}$, and $\{p, 5, 6\}$, respectively. Connecting the points $\left\{C_{12p}, C_{23p}, C_{34p}, C_{45p}, C_{61p}\right\}$ by sequence, we can form the voronoi cell of the point p. If the point $p(\mathbf{x})$ has n natural neighbours (6 neighbours in Figure 3), then the Laplace shape function for node i is defined as:

$$\phi_i(\mathbf{x}) = \frac{\alpha_i(\mathbf{x})}{\sum_{j=1}^{n} \alpha_j(\mathbf{x})}, \alpha_j(\mathbf{x}) = \frac{s_j(\mathbf{x})}{h_j(\mathbf{x})}, \ \mathbf{x} \in R^2 \tag{1}$$

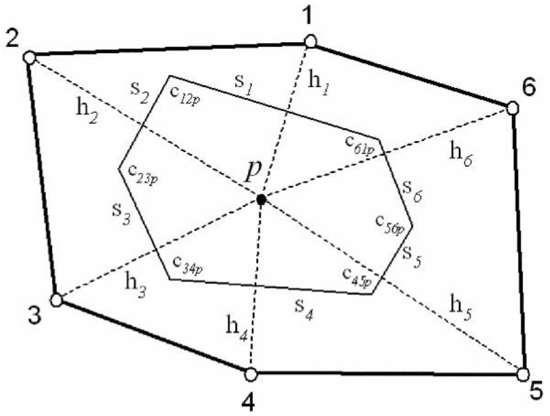


Figure 3. Computation of Laplace interpolation of the natural neighbours.

Where $\alpha_j(\mathbf{x})$ is the Laplace weight function, $s_j(\mathbf{x})$ is the length of the voronoi edge associated with point p and node i, and $h_j(\mathbf{x})$ is the Euclidean distance between point p and node i (Figure 3).

The derivatives of the coordinates are obtained by differentiating Equation (1):

$$\Phi_{i,j}(\mathbf{x}) = \frac{\alpha_{i,j}(\mathbf{x}) - \Phi_i(\mathbf{x})\alpha_{,j}(\mathbf{x})}{\alpha(\mathbf{x})},$$

$$\alpha(\mathbf{x}) = \sum_{k=1}^{n} \alpha_k(\mathbf{x}), \alpha_{,j}(\mathbf{x}) = \sum_{k=1}^{n} \alpha_{k,j}(\mathbf{x}) \tag{2}$$

The global forms of displacement approximations $\mathbf{u}^h(\mathbf{x})$ of point $p(\mathbf{x})$ can be written as

$$\mathbf{u}^h(\mathbf{x}) = \sum_{i=1}^{n} \phi_i(\mathbf{x})\mathbf{u}_i \tag{3}$$

where $\mathbf{u}_i$ $(i = 1, \cdots, n)$ are the vectors of nodal displacements at the n natural neighbours of point p, and $\Phi_i(\mathbf{x})$ are the shape functions associated with each node.

By defining of the shape function given in Equation (1), the following properties are self-evident.

$$\sum_{i=1}^{n} \Phi_i(\mathbf{x}) = 1 \quad \mathbf{x} \in \Omega \tag{4}$$

$$\begin{cases} 0 \le \Phi_i(\mathbf{x}) \le 1 \\ \Phi_i(\mathbf{x}_j) = \delta_{ij} \end{cases} \quad \mathbf{x} \in \Omega \tag{5}$$

From Equation (5), it can be seen that the Laplace interpolation passes through the nodal values, which is in contrast to most meshless approximations, where the nodal parameters $\mathbf{u}_i$ are not nodal displacements. Also, the Laplace shape function is C^0 at nodal locations as well as on the boundary of the support. These properties make the Laplace interpolant the only meshless data interpolation that will exactly satisfy the (linear) essential boundary conditions. A more detailed discussion of the Laplace interpolation and its application to PDEs can be found in Sukumar *et al.* [16] and the references therein.

It is noted that if the position of the point p is the same as an arbitrary node i, the algorithm in Equation (1) will fail because of the Euclidean distance $h_j(\mathbf{x}) = 0$ and the Laplace weight function $\alpha_i(\mathbf{x}) = \pm\infty$. This situation never arises when the triangular finite elements are used as integral cell so that all integration points are interior to the triangles. But when the regular cells similar to the EFG method are used to the integral scheme, this computational difficulty is encountered. A generally applicable method to overcome this numerical difficulty is to randomly move nodes by a small distance (e.g.1e–10) before computing.

3. THE INCREMENT METHOD FOR ELASTO-PLASTIC ANALYSIS

When the elasto-plastic material has reached the plastic state, the stress–strain relationship is,

$$\{d\sigma\} = [D]_{ep}\{d\varepsilon\} \tag{6}$$

This relationship is nonlinear. The step-up loading method can be chosen to linearize the nonlinear problem. At a certain stress and strain level, another loading will produce stress increment $\{\Delta\sigma\}$ and strain increment $\{\Delta\varepsilon\}$. As long as the incremental load is small enough, the Equation (6) can be expressed approximately as follows,

$$\{\Delta\sigma\} = [D]_{ep}\{\Delta\varepsilon\} \tag{7}$$

in which $[D]_{ep}$ is elasto-plastic matrix, e.g., $[D]_{ep} = [D]_e - [D]_p$, which is only dependent on the stress level at the beginning of loading.

The common solution methods of elasto-plastic problem include incremental tangent stiffness method, incremental initial stress method and incremental initial strain method. As the time consume is longer than conventional finite element method, the incremental initial stress method is used in this paper to avoid generating and decomposing of the stiff matrix in every iterate step. In this way, the stiff matrix which is generated and decomposed at the beginning of solution, can maintain unchanged in the subsequent increments.

4. DISCRETE EQUATIONS

We consider the two-dimensional problem with small displacements on the domain Ω bounded by Γ. The equilibrium equation and boundary conditions are given as follows:

$$\nabla \cdot \boldsymbol{\sigma} + \mathbf{b} = 0, \quad \text{in } \Omega \tag{8a}$$

$$\boldsymbol{\sigma} \cdot \mathbf{n} = \bar{\mathbf{t}}, \quad \text{on } \Gamma_t \tag{8b}$$

$$\mathbf{u} = \bar{\mathbf{u}}, \quad \text{on } \Gamma_u \tag{8c}$$

where $\boldsymbol{\sigma}$ is the stress tensor which corresponds to the displacement filed $\mathbf{u}$; $\mathbf{b}$ is the body force vector; the superposed bar in Equations (8b) and (8c) denotes prescribed boundary values, and $\mathbf{n}$ is the unit normal to the domain.

The variational form of Equation (8) is posed as follows

$$\int_{\Gamma_t} \delta u \cdot t\mathbf{d}\Gamma - \int_{\Omega} \delta\varepsilon \cdot \sigma d\Omega + \int_{\Omega} \delta u \cdot b\mathbf{d}\Omega = 0 \ \forall \delta u \in H_0^1 \tag{9}$$

The discretized system can be obtained by substituting Equation (1) into (9)

$$K \cdot D = f \tag{10}$$

where

$$K_{\mathbf{IJ}} = \int_{\Omega} B_{\mathbf{I}}^{\mathbf{T}} \cdot DE \cdot B_{\mathbf{J}} \mathbf{d}\Omega \int \tag{11}$$

$$\mathbf{f_I} = \int_{\Gamma_t} \Phi_{\mathbf{I}} \cdot \mathbf{t} d\Gamma + \int_{\Omega} \Phi_{\mathbf{I}} \cdot \mathbf{b} d\Omega \tag{12}$$

where DE is the elasticity matrix, $\mathbf{B}_I$ is the strain matrix.

From the above deducation, we can know that the numerical results obtained by NNM [16] and the proposed MNNM are the same because the algorithm for the neighbour-search is the only difference in NNM and MNNM.

5. NUMERICAL EXAMPLES

The MNNM is coded in standard C++. Cases are run in order to examine the MNNM in two-dimensional elastostatics.

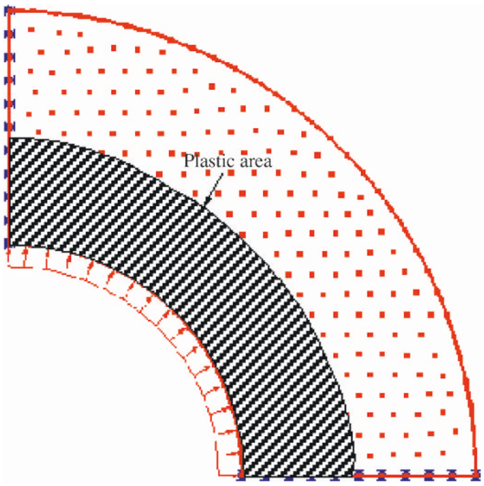


Figure 4. Computational model and plastic area.

5.1 Thick cylinder

Consider an axially restrained thick cylinder of inside diameter a = 10 mm, and outside diameter b = 20 mm, which is subjected to internal pressure p = 12 kPa. The material is perfectly elasto-plastic with Young's modulus $E = 85570$ kPa and Poisson's ratio $\mu = 0.3$. The Von-Mises yield criterion is adopted and the tensile yield limit σ_s is equal to 10 kPa. The problem statement is given in Figure 4.

The comparison in Figures 5 and 6 show that, with the same nodes distribution, the results obtained by MNNM are in good agreement with the analytical solutions and much better than those obtained by triangular FEM.

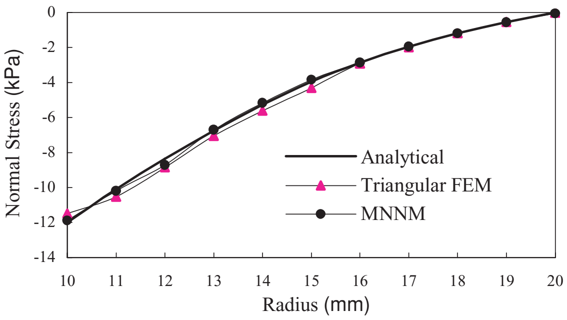


Figure 5. Comparison of normal stress.

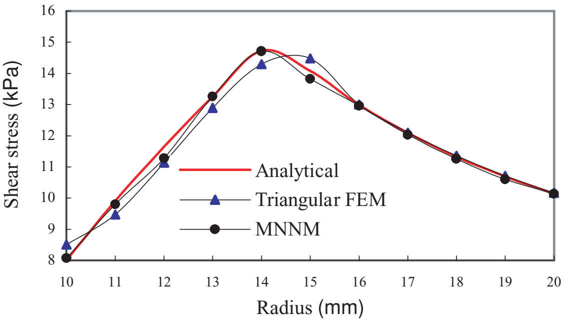


Figure 6. Comparison of shear stress.

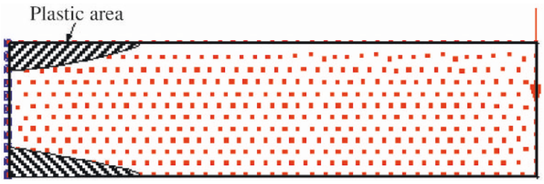


Figure 7. Computational model and plastic area for cantilever beam problem.

5.2 Cantilever beam

The second example considered here is a rectangle cantilever beam of length $l = 40\,m$ and height $h = 10\,m$ as shown in Figure 7. The beam is fixed at one end and subjected to concentrated load $P = -11kN$ at the free end. The material is perfectly elasto-plastic and obeys the Von-Mises yield criterion. The material parameters are $E = 86670$ kPa, $\mu = 0.3$ and $\sigma_s = 10$ kPa. With the same nodes distribution, the comparison of σ_x, the normal stress in x direction, is shown in Figure 8.

6. CONCLUSION

The Meshless Natural Neighbour Method, which is a truly meshless method similar to EFG, can treat material discontinuities conveniently and keep some advantages of FEM. In this paper, by combining MNNM and incremental initial stress method, a general C++ programme has been worked out and applied into elasto-plastic analysis, in which perfectly elasto-plastic model and Von-Mises yield criterion is adopted. Although only several simple examples were presented in this chapter, it is expected that the proposed method can be used to solve some more complicated problems in geotechnical engineering,

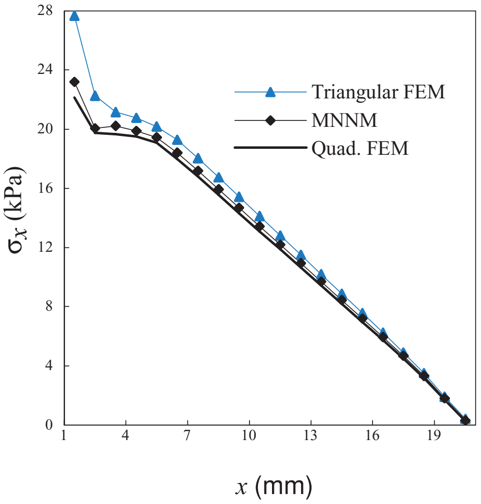


Figure 8. Comparison of σ_x

as well as problems with large deformation such as landsliding and pile penetration.

REFERENCES

1. R.A. Gingold and J.J. Moraghan (1977), Smoothed particle hydrodynamics: theory and applications to non-spherical stars. *Monthly Notices of the Royal Astronomical Society*, 181, 2, pp. 375–389.
2. W.K. Liu, S. Jun and Y.F. Zhang (1995), Reproducing kernel particle methods. *International Journal for Numerical Methods in Fluids*, 20, pp. 1081–1106.
3. W.K. Liu and Y. Chen (1995), Wavelet and multiple scale reproducing kernel method.*International Journal for Numerical Methods in Fluids*, 21, pp. 901–931.
4. R.R. Ohs and N.R. Aluru (2001), Meshless analysis of piezoelectirc devices. *Computational Mechanics*, 27, pp. 23–36.
5. T. Belytschko, Y.Y. Lu and L. Gu (1994), Element-free Galerkin method.*International Journal for Numerical Methods in Engineering*, 37, pp. 229–256.
6. B. Nayroles, G. Touzot and P. Villon (1992), Generalizing the finite element method: diffuse approximation and diffuse elements. *Computational Mechanics*, 10, pp. 307–318.
7. L.W. Cordes and B. Moran (1996), Treatment of material discontinuity in the Element-free Galerkin method. *Computer Methods in Applied Mechanics and Engineering*, 139, pp. 75–89.
8. Y. Krongauz and T. Belytschko (1998), EFG approximation with discontinuous derivatives. *International Journal for Numerical Methods in Engineering*, 41, pp. 1215–1233.
9. E. Oñate, S.R. Idelsohn, O.C. Zienkiewicz *et al.* (1996), A finite point method in computational mechanics: applications to convective transport and fluid flow.*International Journal for Numerical Methods in Engineering*, 39, pp. 3839–3866.

10. C.A. Duarte and J.T. Oden (1996), Hp clouds: a h-p meshless method. *Numerical Methods for Partical Differential Equations*, 12, pp. 673–705.
11. T. Zhu, J. Zhang and S.N. Atluri (1998), A local boundary integral equation (LBIE) method in computational mechanics, and a meshless discretization approach. *Computational Mechanics*, 21, pp. 223–235.
12. S.N. Atluri and T. Zhu (1998), A new meshless local Petrov-Galerkin (MLPG) approach in computational mechanics. *Computational Mechanics*, 22, pp. 117–127.
13. J. Braun and M. Sambridge (1995), A numerical method for solving partial differential equations on highly irregular evolving grids. *Nature*, 376, pp. 655–660.
14. N. Sukumar, B. Moran and T. Belytschko (1998), The nature element method in solid mechanics. *International Journal for Numerical Methods in Engineering*, 43, pp. 839–887.
15. E. Cueto, M. Doblare and L. Gracia (2000), Imposing essential boundary conditions in the natural element method by means of density-scaled α-shapes.*International Journal for Numerical Methods in Engineering*, 49, pp. 519–546.
16. N. Sukumar, B. Moran and Y. Semenov (2001), Natural neighbour Galerkin method. *International Journal for Numerical Methods in Engineering*, 50, pp. 1–27.
17. N. Sukumar (2003), Voronoi cell finite difference method for the diffusion operator on arbitrary unstructured grids.*International Journal for Numerical Methods in Engineering*, 57, pp. 1–34.
18. S.R. Idelsohn, E. Oñate, N. Calvo and F.D. Pin (2003), The meshless finite element method. *International Journal for Numerical Methods in Engineering*, 58, pp. 893–912.
19. E. Cueto, N. Sukumar, B. Calvo, M.A. Martínez, J. Cegoñino and M. Doblaré (2003), Overview and recent advances in natural neighbor Galerkin methods. *Archives of Computational Methods in Engineering*, 10, 4, pp. 307–384.
20. Y.C. Cai and H.H. Zhu (2004), A meshless local natural neighbour interpolation method for stress analysis of solids. *Engineering Analysis with Boundray Elements*, 28, 6, pp. 607–613.
21. X. Zhang, K.Z. Song, M.W. Lu *et al.* (2000), Meshless methods based on collocation with radial basis function. *Computational Mechanics*, 26, 4, pp. 333–343.

A MESHLESS LOCAL PETROV–GALERKIN METHOD FOR ELASTO-PLASTIC PROBLEMS

Y.B. Xiong[1], S.Y. Long[1], K.Y. Liu[1], and G.Y. Li[2]
[1]*Department of Engineering Mechanics, Hunan University, Changsha, Hunan, China 410082*
[2]*College of Mechanical and Automotive Engineering, Hunan University, Changsha, Hunan, China 410082*

Abstract A meshless local Petrov–Galerkin method (MLPG) is presented for solving the elasto-plasticity problem in the paper. It is a truly meshless method using the moving least square (MLS) approximation as a trial function and the MLS weighted function as a test function in the weighted residual method. The incremental tangent stiffness method is applied in computation. Numerical examples show that the local Petrov–Galerkin method is applicable and effective for solving the elasto-plasticity problem.

Keywords: elasto-plastic problem, MLPG, MLS, increment tangent stiffness method.

1. INTRODUCTION

Atluri and Zhu [1] originally proposed a MLPG method which is one of meshless methods. The MLPG method requires only nodal information and no element connectivity is needed, which leads to a simple and convenient preprocess. The MLPG method has many advantages over traditional numerical methods such as FEM and BEM. It is attracting many researches' attention. Atluri and Zhu [1] solved Laplace equations and Poisson equations, and Long [2, 4] solved the elasticity problems and thin plates by the MLPG method. In this paper, the MLPG method is extended to solve elasto-plastic problems.

G. R. Liu et al. (eds.), Computational Methods, 1477–1482.

2. MLPG FORMULATIONS FOR ELASTO-PLASTIC PROBLEMS

The meshless local Petrov–Galerkin integral equation is given as follows for 2-D geometrically elasto-plastic problems

$$\int_{\Omega_s} \dot{\sigma}_{ij} v_{i,j} d\Omega - \int_{\Gamma_{su}} \dot{t}_i v_i d\Gamma + \alpha \int_{\Gamma_{su}} \dot{u}_i v_i d\Gamma$$

$$= \int_{\Gamma_{st}} \bar{\dot{t}}_i v_i d\Gamma + \alpha \int_{\Gamma_{su}} \bar{\dot{u}}_i v_i d\Gamma + \int_{\Omega_s} \dot{b}_i v_i d\Omega \tag{1}$$

where $\dot{\sigma}_{ij}$ is the rate of stress tensor and v_i the test function. The quantity α is a penalty factor used to impose essential boundary conditions, whose value is usually suggested to be 10^5–10^8 times of Young's modulus. $\dot{\boldsymbol{t}}_i$ is the rate of surface traction and $\dot{\boldsymbol{u}}_i$ the velocity. The subdomain Ω_s is a circle of radius r_0 for a two-dimensional problem. The local boundary $\partial\Omega_s$ is further divided into three parts and $\partial\Omega_s = L_s \cup \Gamma_{su} \cup \Gamma_{st}$, where no boundary conditions are specified on L_s, tractions are prescribed on Γ_{st} and displacements are prescribed on Γ_{su}.

Using MLS interpolation function [5] to approximate the velocity $\dot{\mathbf{u}}^h(\mathbf{x}) = \sum_{i=1}^{n} \Phi_i(\mathbf{x})\hat{\dot{\mathbf{u}}}_i$, the discretization equation can be written in matrix form

$$\sum_{j=1}^{n} \boldsymbol{K}_{Ij} \hat{\dot{\boldsymbol{u}}}_j = \boldsymbol{f}_I \tag{2}$$

where

$$\boldsymbol{K}_{Ij} = \int_{\Omega_s} \boldsymbol{G}_{\boldsymbol{I}}^{\boldsymbol{T}} \boldsymbol{S}_j \, d\boldsymbol{\Omega} + \alpha \int_{\Gamma_{su}} \boldsymbol{W}_{\boldsymbol{I}} \varphi_j \, d\Gamma - \int_{\Gamma_{st}} \boldsymbol{W}_{\boldsymbol{I}} \, \psi_j \, d\Gamma \tag{3}$$

$$\boldsymbol{f}_{\boldsymbol{I}} = \int_{\Omega_s} \boldsymbol{W}_{\boldsymbol{I}} \dot{b} \, d\Omega + \int_{\Gamma_{st}} \boldsymbol{W}_{\boldsymbol{I}} \bar{\dot{t}} \, d\Gamma + \alpha \int_{\Gamma_{su}} \boldsymbol{W}_{\boldsymbol{I}} \bar{\dot{\boldsymbol{u}}} \, d\Gamma \tag{4}$$

where,

$$\boldsymbol{G}_{\boldsymbol{I}} = \begin{bmatrix} w_{I,x} & 0 \\ 0 & w_{I,y} \\ w_{I,y} & w_{I,x} \end{bmatrix}, \quad \boldsymbol{W}_{\boldsymbol{I}} = \begin{bmatrix} w_{\boldsymbol{I}} & 0 \\ 0 & w_{\boldsymbol{I}} \end{bmatrix}$$

$$\Phi_i(x) = \begin{bmatrix} \varphi_i(x) & 0 \\ 0 & \varphi_i(x) \end{bmatrix}, \quad \boldsymbol{S}_i = \boldsymbol{D}\boldsymbol{B}_i = \frac{E}{1-v^2} \begin{bmatrix} \varphi_{i,x} & v\varphi_{i,y} \\ v\varphi_{i,x} & \varphi_{i,y} \\ \frac{1-v}{2}\varphi_{i,y} & \frac{1-v}{2}\varphi_{i,x} \end{bmatrix}$$

$$\psi_i = \boldsymbol{n}\boldsymbol{S}_i = \frac{E}{1-v^2}\begin{bmatrix} n_1\varphi_{i,x} + \dfrac{1-v}{2}n_2\varphi_{i,y} & n_1 v\varphi_{i,y} + \dfrac{1-v}{2}n_2\varphi_{i,x} \\ n_2\varphi_{i,x} + \dfrac{1-v}{2}n_1\varphi_{i,y} & n_2\varphi_{i,y} + \dfrac{1-v}{2}n_1\varphi_{i,x} \end{bmatrix} \tag{5}$$

During the solving process, the incremental tangent stiffness method [6] is adopted.

3. NUMERICAL EXAMPLES

Example 1. A cantilever beam is now studied with a parabolic-shear end load, as shown in Figure 1. The geometric and material parameters are given as: $l = 8$ m, $h = 1$ m, one unit thickness, $P = 1$ N, $E = 10^5$ Pa, $\mu = 0.25$, $\sigma_s = 25$ Pa. The linear hardening model is used here and the material behaviour obeys the V. Mises yield condition with $E_t = 0.2E$ and $H' = EE_t/(E - E_t)$. 102(17 × 6) nodes in the global domain are regularly distributed. In MLS approximation the cubic basis function and the weighted function recommended by the authors [3] are employed, and the radius of a support domain is equal to 4.0. The quartic spline weighted function is used as the test function in the weighted residual method. The radius of a local subdomain is equal to 0.1 m. 10 gauss points and 36(6 × 6) gauss points are used on local boundary $\partial\Omega_s$ and in local subdomain Ω_s, respectively. The MLPG results for part of nodal displacements are listed in Table 1 and compared with the finite element method

Table 1. Vertical displacements of a cantilever beam subjected to a concentrated load at the free end.

Coordinate (m)	FEM (mm)	MLPG (mm)
(2.0, 0.0)	2.947	2.906
(3.5, 0.0)	7.807	7.791
(5.0, 0.0)	13.926	13.824
(6.5, 0.0)	20.854	20.546
(8.0, 0.0)	28.183	28.227

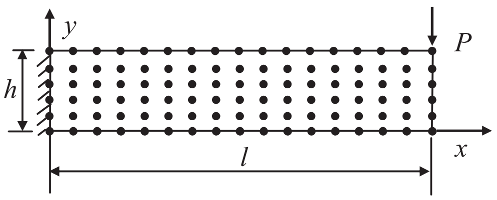


Figure 1. A Cantilever beam subjected to a concentrated force at the free end.

Table 2. Vertical displacements of a simply supported beam subjected to a uniformly distributed load.

Coordinate (m)	FEM (mm)	MLPG (mm)
(1.0, 0.0)	3.911	3.882
(2.5, 0.0)	8.657	8.568
(4.0, 0.0)	10.601	10.547
(5.5, 0.0)	8.657	8.568
(7.0, 0.0)	3.911	3.882

(FEM) solution computed with the same nodes. Table 1 shows that the MLPG and the FEM results virtually coincide. Figures 3 and 4 display respectively the contours of effective stress at the concentrated load $P = 1$ N computed by MLPG method and FEM. The numbers in the contours of the figures are the value of effective stress. The figures indicate that the difference between the MLPG and the FEM results are small. The relative error of effective stress is less than 2.73%. The situation can be improved by increasing the number of nodes. Example 2: A simply supported cantilever beam is discussed that under a uniformly distributed load $q = 1$ N/m, as shown in Figure 2. All data are the same as Example 1. The MLPG results for part of nodal displacements listed in Table 2 are compared with the FEM solutions computed with the same node distribution. Table 2 again shows that the MLPG and the FEM results agree well. Figures 5 and 6 display, respectively, the contours of effective stress

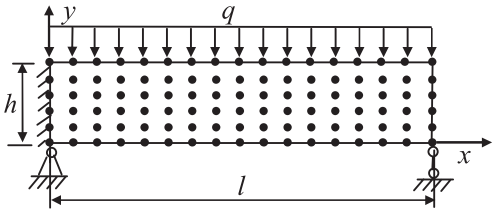


Figure 2. A simply supported beam subjected to a uniformly distributed load.

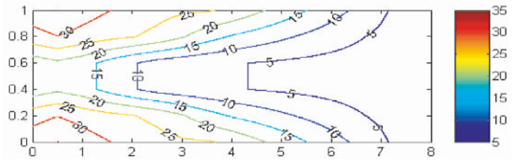


Figure 3. The contour of the effective stress by MLPG method at the concentrated load $P = 1$ N.

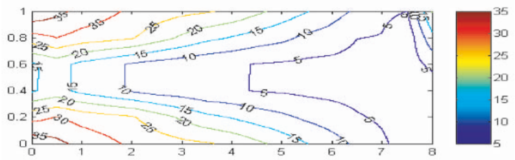


Figure 4. The contour of the effective stress by FEM at the concentrated load $P = 1$ N.

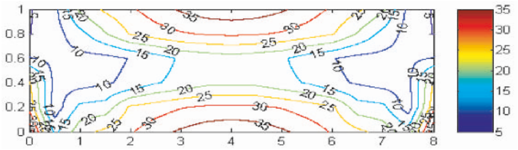


Figure 5. The contour of the effective stress by MLPG method at the uniformly distributed load $q = 1$ N/m.

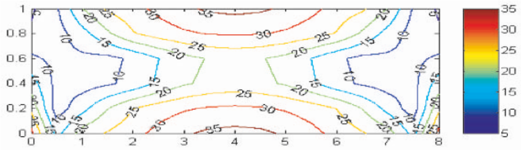


Figure 6. The contour of the effective stress by FEM at the uniformly distributed load $q = 1$ N/m.

obtained by MLPG method and FEM. The relative error of effective stress is less than 2.19%. Similarly the situation can be also improved by increasing the number of nodes.

4. CONCLUSIONS

In this paper, MLPG is applied to solve elasto-plastic problems. The numerical experiments show that MLPG method is applicable and effective in solving the elasto-plastic problem and it is easily extended to solve more complex non-linear problems and nonhomogeneous mediums problems.

ACKNOWLEDGEMENTS

The support of this work by Natural Science Foundation of China (no. 10372030), National 863 project of China (no. 2003AA411230) and Foundation of the Ministry of Education of China (no. 20020532021, [2002] 350) is greatly acknowledged.

REFERENCES

1. S.N. Atluri and T.L. Zhu (1998), A new local meshless local Petrov–Galerkin (MLPG) approach in computational mechanics. *Computational Mechanics*, 22(2), pp. 117–127.
2. S. Long. A local Petrov–Galerkin method for the elasticity problem. *Acta Mechanica Sinica*, 2001, 33(4), pp. 508–518 (in Chinese).
3. W. Zhou and X. Kou (1998), Element free method and its application in engineering. *Acta Mechanica Sinica*, 30(2), pp. 193–202 (in Chinese).
4. Y. Xiong and S. Long (2004) A meshless local Petrov–Galerkin method for a thin plate. *Applied Mathematics and Mechanics*, 28(2), pp. 189–196 (in Chinese).
5. T. Belytschko, Y.Y. Lu, L. Gu (1994), Element-free Galerkin method. *International Journal of Numerical Methods in Engineering*, 37, 229–256.
6. S. Long and S. Chen (2003), Element-free Galerkin method for elasto-plastic problems. *Engineering Mechanics*, 20(2), pp. 66–70(in Chinese).

PARTICLE-PARTITION OF UNITY METHODS IN ELASTICITY

Michael Griebel and Marc Alexander Schweitzer
Institut für Numerische Simulation, Universität Bonn, Wegelerstrasse 6, D-53115 Bonn, Germany

Abstract We consider discretizations of problems in elasticity using the particle–partition of unity method (PUM). We focus on discretization issues and fast solution techniques. Numerical results for applications in two and three dimensions also for obstacle problems are presented.

Keywords: meshfree method, partition of unity, multi-level solvers, Nitsche's method.

1. INTRODUCTION

The particle–partition of unity method (PUM) [1–6] is a meshfree Galerkin method for the numerical treatment of partial differential equations (PDE). In essence, it is a generalized finite element method (GFEM) which employs piecewise rational shape functions rather than piecewise polynomial functions. The PUM shape functions, however, make up a basis of the discrete function space unlike other GFEM approaches which allows us to construct fast multilevel solvers in a similar fashion as in the finite element method (FEM).

The paper is organized as follows: In Section 2 we shortly review the construction of PUM spaces, the Galerkin discretization of a linear elliptic PDE using our PUM as well as the fast multi-level solution of the arising linear system. Then we present some numerical results with respect to approximation as well as fast solution techniques in two and three space dimensions obtained with our PUM for the numerical solution of the Navier–Lamé equations in Section 2.4. The discretization of constrained minimization problems like the obstacle problem is the subject of Section 3. Then, some numerical results for the obstacle problem in two space dimensions are given in Section 3.2. Finally, we conclude with some remarks.

G. R. Liu et al. (eds.), Computational Methods, 1483–1493.

2. PARTITION OF UNITY METHOD

In the following, we shortly review the construction partition of unity spaces and the meshfree Galerkin discretization of an elliptic PDE, see [1, 2] for details. Furthermore, we give a summary of the efficient multi-level solution of the arising linear block-system, see [3] for details.

2.1 Construction of Partition of Unity Spaces

In a partition of unity method, we define a global approximation u^{PU} simply as a weighted sum of local approximations u_i,

$$u^{\mathrm{PU}}(x) := \sum_{i=1}^{N} \varphi_i(x) u_i(x). \tag{1}$$

These local approximations u_i are completely independent of each other, i.e., the local supports $w_i := \mathrm{supp}(u_i)$, the local basis $\{\psi_i^n\}$ and the order of approximation p_i for every single $u_i := \sum u_i^n \psi_i^n \in V_i^{pi}$ can be chosen independently of all other u_j. Here, the functions φ_i form a partition of unity (PU). They are used to splice the local approximations u_i together in such a way that the global approximation u^{PU} benefits from the local approximation orders p_i yet it still fulfils global regularity conditions. Hence, the global approximation space on Ω is defined as

$$V^{\mathrm{PU}} := \sum_i \varphi_i V_i^{pi} = \sum_i \varphi_i \ \mathrm{span}\left\langle \{\psi_i^n\} \right\rangle = \ \mathrm{span}\left\langle \{\varphi_i \psi_i^n\} \right\rangle. \tag{2}$$

The starting point for any meshfree method is a collection of N independent points $P := \{x_i \in \mathbb{R}^d \mid x_i \in \overline{\Omega}, i = 1, \dots, N\}$. In the PU approach we need to construct a partition of unity $\{\varphi_i\}$ on the domain of interest Ω to define an approximate solution (1) where the union of the supports $\mathrm{supp}(\varphi_i) = \overline{\omega}_i$ covers the domain $\overline{\Omega} \subset \cup_{i=1}^{N} \omega_i$ and $u_i \in V_i^{pi}(\omega_i)$ is some locally defined approximation of order p_i to u on w_i. Thus, the first (and most crucial) step in a PUM is the efficient construction of an appropriate cover $C_\Omega := \{w_i\}$. Throughout this paper we use a tree-based construction algorithm for d-rectangular covers C_Ω presented in [2, 6]. Here, the cover patches w_i are products of intervals $(x_i^l - h_i^l, x_i^l + h_i^l)$ for $l = 1, \dots, d$. With the help of weight functions W_k defined on these cover patches w_k we can easily generate a partition of unity by *Shepard's method*, i.e., we define

$$\varphi_i x = \frac{W_i(x)}{\sum_{wk \in C_\Omega^i} W_k(x)}, \tag{3}$$

where $C_i := \{\omega_j \in C_\Omega \mid \omega_i \cap \omega_j \neq \emptyset\}$ is the set of all geometric neighbors of a cover patch ω_i. Due to the use of d-rectangular patches ω_i, the most natural choice for a weight function W_i is a product of one-dimensional functions, i.e., $W_i(x) = \prod_{l=1}^{d} W_i^l(x^l) = \prod_{l=1}^{d} \mathcal{W}(\frac{x-x_i^l+h_i^l}{2h_i^l})$ with supp $(\mathcal{W}) = [0, 1]$ such that supp $(W_i) = \overline{\omega}_i$. It is sufficient for this construction to choose a one-dimensional weight function $\mathcal{W}$ with the desired regularity which is non-negative. The partition of unity functions φ_i inherit the regularity of the generating weight function $\mathcal{W}$. Throughout this paper we use a linear B-spline as the generating weight function $\mathcal{W}$.

In general, a partition of unity $\{\varphi_i\}$ can only recover the constant function on the domain Ω. Hence, we need to improve the approximation quality to use the method for the discretization of a PDE. To this end, we multiply the partition of unity functions φ_i locally with polynomials ψ_i^n. Since we use d-rectangular patches ω_i only, a local tensor product space is the most natural choice. Here, we use products of univariate Legendre polynomials as local approximation spaces V_i^{pi}, i.e., we choose

$$V_i^{pi} = \operatorname{span}\left\langle \left\{ \psi_i^n \mid \psi_i^n = \prod_{l=1}^{d} \mathcal{L}_i^{\hat{n}_l}, \|\hat{n}\|_1 = \sum_{l=1}^{d} \hat{n}_l \leq p_i \right\} \right\rangle,$$

where $\hat{n}$ is the multi-index of the polynomial degrees $\hat{n}_l$ of the univariate Legendre polynomials $\mathcal{L}_i^{\hat{n}_l} : [x_i^l - h_i^l, x_i^l + h_i^l] \to \mathbb{R}$, and n is the index associated with the product function $\psi_i^n = \Pi_{l=1}^{d} \mathcal{L}_i^{\hat{n}l}$.

For the approximation of vector-fields we employ vector-valued shape functions $\varphi_i \psi_i^n$; i.e., we simply change the definition of our local approximation spaces $V_i^{pi} = \operatorname{span}\langle \vec{\psi}_i^n \rangle$ but keep the partition of unity functions φ. To this end, we choose the local vector-valued basis functions

$$\vec{\psi}_i^n := \vec{\psi}_i^{\tilde{n},l} := \psi_i^{\tilde{n}} \vec{e}_l$$

where we simply multiply the scalar functions $\psi_i^{\tilde{n}}$ with an appropriate unit vector $\vec{e}_l$. In the following we will drop the explicit vector notation and use the symbol ψ_i^n also for vector-valued functions.

In summary, we can view the construction given above as follows:

$$\begin{pmatrix} \{x_i\} \\ \mathcal{W} \\ \{p_i\} \end{pmatrix} \to \begin{pmatrix} \{\omega_i\} \\ \{W_i\} \\ \{V_i^{pi} = \operatorname{span}\langle \psi_i^n \rangle\} \end{pmatrix} \to \begin{pmatrix} \{\varphi_i\} \\ \{V_i^{p_i}\} \end{pmatrix} \to V^{\mathrm{PU}} = \sum \varphi_i V_i^{p_i},$$

where the set of points $P = \{x_i\}$, the generating weight function $\mathcal{W}$ and the local approximation orders p_i are assumed to be given.

2.2 Galerkin Discretization

Consider the elliptic boundary value problem

$$Lu = f \quad \text{in } \Omega \subset \mathbb{R}^d, \qquad Bu = g \quad \text{on } \partial\Omega', \tag{4}$$

where L is a symmetric partial differential operator of second order and B expresses suitable boundary conditions. The imposition of essential boundary conditions within meshfree methods is more involved than in the FEM for a number of reasons and many different approaches have been proposed. We use Nitsche's method [7] to enforce Dirichlet boundary conditions. The main advantages of this approach are that it does not require a second function (or multiplier) space and that it leads to a positive definite linear system, see [5, 6] for a more detailed discussion of Nitsche's method in the PUM context. Here, we just state resulting weak formulation $a(u, v) = l(v)$ of the simple Poisson problem

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega \subset \mathbb{R}^d, \\ u &= g_D \quad \text{on } \Gamma_D \subset \partial\Omega, \\ u_n &= g_N \quad \text{on } \Gamma_N = \partial\Omega \backslash \Gamma_D, \end{aligned}$$

with mixed boundary conditions which reads as

$$\int_\Omega \nabla u \nabla v + \int_{\Gamma_D} u(\beta v - v_n) - u_n v = \int_\Omega f v + \int_{\Gamma_D} g_D(\beta v - v_n) + \int_{\Gamma_N} g_N v, \tag{5}$$

where the subscript n denotes the normal derivative and β is the Nitsche regularization parameter which depends on the employed PUM space but can be pre-computed without much additional cost. Finally, for the Galerkin discretization of (5) we have to compute the stiffness matrix

$$A = (A_{(i,n),(j,m)}), \quad \text{with } A_{(i,n),(j,m)} = a(\varphi_j \psi_j^m, \varphi_i \psi_i^n) \in \mathbb{R}$$

and the right-hand side vector

$$\hat{f} = (f_{(i,n)}), \quad \text{with } f_{(i,n)} = \langle f, \varphi_i \psi_i^n \rangle L^2 = \int_\Omega f \varphi_i \psi_i^n \in \mathbb{R}.$$

The stable approximation of these integrals is somewhat more involved than in the finite element method (FEM). Due to the meshfree construction given above the shape functions $\varphi_i \varphi_i^n$ are piecewise rational functions only so that the respective integrands have a number of jumps within the integration domain which need to the resolved. For the stable numerical integration of the weak form we use a tree-based decomposition scheme together with efficient sparse grid integration rules, see [2, 8].

2.3 Multi-level Solution of Resulting Linear System

The product structure of the shape functions $\varphi_i \psi_i^n$ implies two natural block-partitions of the resulting linear system $A\tilde{u} = \hat{f}$, where $\tilde{u}$ denotes a coefficient vector and $\hat{f}$ denotes a moment vector.

1. The stiffness matrix A can be arranged in *spatial blocks*. A spatial block A_{nm} corresponds to a discretization of the PDE on the complete domain Ω using the trial functions $\varphi_j \psi_j^m$ and the test function $\varphi_i \psi_i^n$ with fixed n and m. Here, all blocks A_{nm} are sparse matrices and have the same row and column dimensions which corresponds to the number of partition of unity functions φ_i.
2. The stiffness matrix A may also be arranged in *polynomial blocks*. Here, a single block A_{ij} corresponds to a local discretization of the PDE on the domain $\omega_i \cap \omega_j \cap \Omega$. The polynomial blocks A_{ij} are dense matrices and may have different dimensions corresponding to the dimensions of the local approximation spaces V_j^{pj} and V_i^{pi}.

This separation of the degrees of freedom into local approximation functions ψ_i^n and partition of unity functions φ_i can be used to define two different multilevel concepts [3]. Throughout this paper we assume that the stiffness matrix is given in polynomial block-form and we use the corresponding spatial multilevel solver developed in [3] for the fast and efficient solution of the resulting large sparse linear block-system $\mathrm{A}\tilde{u} = \hat{f}$, where $\tilde{u}$ denotes a coefficient block-vector and $\hat{f}$ a moment block-vector.

In a multi-level method we need a sequence of discretization spaces V_k with $k = 0, \ldots, J$ where J denotes the finest level. To this end we construct a sequence of PUM spaces V_k^{PU} with the help of a tree-based algorithm developed in [2, 3]. As a first step we generate a sequence of point sets P_k and covers C_Ω^k from a given initial point set $\tilde{P}$ with this algorithm, see Figure 1. Following the construction given in Section 2.1 we can then define an associated sequence of PUM spaces V_k^{PU}. Note that these spaces are nonnested, i.e., $V_{k-1}^{\mathrm{PU}} \not\subset V_k^{\mathrm{PU}}$, and that the shape functions $\varphi_{i,k}\psi_{i,k}^n$ are non-interpolatory. Thus, we need to construct appropriate transfer operators $I_{k-1}^k : V_{k-1}^{\mathrm{PU}} \to V_k^{\mathrm{PU}}$ and $I_k^{k-1} : V_k^{\mathrm{PU}} \to V_{k-1}^{\mathrm{PU}}$. With such transfer operators I_{k-1}^k, I_k^{k-1} and the stiffness matrices A_k coming from the Galerkin discretization on each level k we can then set up a standard multiplicative multi-level iteration to solve the linear system $A_J\tilde{u}_J = \hat{f}_J$. Our multi-level solver utilizes special localized L^2-projections for the interlevel transfers and a block-smoother to treat all local degrees of freedom ψ_i^n within a patch ω_i simultaneously. Namely, we use the so-called local-to-local L^2-projections as prolongation operators I_k^{k-1} for scalar as well as vector-valued problems. For further details see [3, 6].

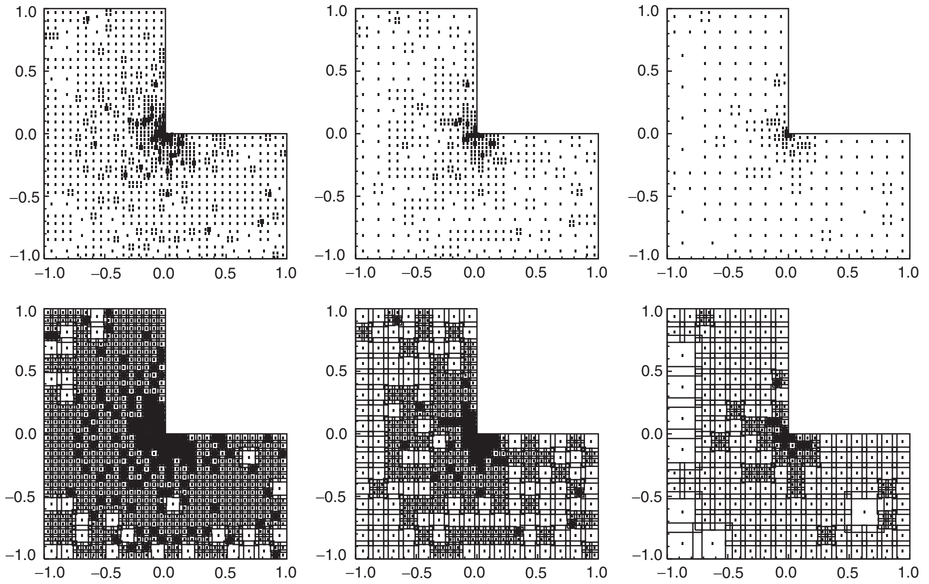


Figure 1. Point sets P_k and covers C_Ω^k for $k = 10, \dots, 8$ generated for an initial graded Halton (2,3) point set $\tilde{P}$ with $\tilde{N} = 678$ points. The number N of generated points on the finest level $J = 10$ is $N = 1293$.

2.4 Numerical Results for Linear Elasticity

In the following we consider the numerical solution of the Navier–Lamé equations

$$-\mu \Delta u - (\lambda + \mu) \nabla (\nabla \cdot u) = f \quad \text{in} \quad \Omega \subset \mathbb{R}^d, \quad d = 2, 3$$

together with suitable boundary conditions $u_D = g_D$ on $\Gamma_D \subset \partial\Omega$ and $\boldsymbol{\sigma}(u).n = g_N$ on $\Gamma_N = \partial\Omega / \Gamma_D$ where $\boldsymbol{\sigma}(u) := \lambda \nabla \cdot u I + 2\mu\epsilon(u)$ denotes the symmetric stress tensor and $\epsilon(u) := \frac{1}{2}(\partial_i U_j + \partial_j u_i)$ the strain tensor associated with the displacement field $u = (u_i), i = 1, \dots, d$. The parameters λ and μ are the so-called Lamé parameters. They are related to the Poisson ratio ν and the Young modulus E of the material via $\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}$ and $\mu = \frac{E}{2(1+\nu)}$. The associated bilinear form arising from Nitsche's approach is given by

$$\begin{aligned} a(u, v) = \int_\Omega \boldsymbol{\sigma}(u) : \epsilon(v) + \int_{\Gamma_D} & 2\mu\beta_\epsilon u \cdot v + \lambda\beta_{\text{div}}(u \cdot n)(v \cdot n) \\ & - ((\boldsymbol{\sigma}(u) \cdot n) \cdot v + u \cdot (\boldsymbol{\sigma}(v) \cdot n)) \end{aligned}$$

Table 1. Convergence rates $\rho_1^{v,v}$ for the V^{v,v}-cycle and convergence rates $\rho_2^{v,v}$ for the W^{v,v}-cycle with $v = 2, 3, 4$ in two and three dimensions.

N	J	p	D$_p$	$\rho_1^{2,2}$	$\rho_1^{3,3}$	$\rho_1^{4,4}$	$\rho_2^{2,2}$	$\rho_2^{3,3}$	$\rho_2^{4,4}$
16,384	7	1	3	0.502	0.438	0.388	0.240	0.182	0.150
65,536	8	1	3	0.503	0.437	0.389	0.229	0.170	0.141
262,144	9	1	3	0.504	0.437	0.389	0.221	0.162	0.136
1,048,576	10	1	3	0.505	0.437	0.389	0.217	0.158	0.134
512	3	1	4	0.769	0.712	0.678	0.549	0.415	0.398
4,096	4	1	4	0.741	0.714	0.677	0.367	0.268	0.222
32,768	5	1	4	0.735	0.706	0.675	0.285	0.221	0.195
262,144	6	1	4	0.726	0.694	0.671	0.261	0.202	0.175

and the linear form on the right-hand side by

$$\begin{aligned} l(v) = \int_\Omega f \cdot v + \int_{\Gamma_N} g_N \cdot v + \int_{\Gamma_D} & 2\mu\beta_\epsilon g_D \cdot v + \lambda\beta_{\text{div}}(g_D \cdot n)(v \cdot n) \\ & - g_D \cdot (\boldsymbol{\sigma}(v) \cdot n). \end{aligned}$$

We measure the convergence of our multi-level solver via a simple test problem, where we use a sequence of uniform point sets in the domain $\Omega = [-1, 1]^d$ for $d = 2, 3$. In this example we use the parameters $E = 1$ and $\nu = \frac{1}{3}$ for the material and the boundary conditions $u = 0$ on $\Gamma_D := \{x \in \partial\Omega : x_0 = -1\}$, $\boldsymbol{\sigma}(u) \cdot n = 0$ on $\Gamma_{N1} := \{x \in \partial\Omega : x_1 = -1 \text{ or } x_1 = 1\}$, and $\sigma(u) \cdot n = (0, -1)^T$. In two dimensions the finest discretization is based on $N = 1{,}048{,}576$ points and employs $\mathsf{dof} = 6{,}291{,}456$ degrees of freedom, in three dimensions the finest discretization uses $N = 262{,}144$ and $\mathsf{dof} = 1{,}048{,}576$.

In Table 1 we give the measured convergence rates ρ for our multi-level solver using the V- and W-cycle (ρ^1 and ρ^2 respectively) with 2, 3 and 4 smoothing steps, the number of points N on the finest level J, the polynomial degree p and the dimension D_p of the local approximation spaces V_i^p. These rates $\rho := ||u_r||_{L2}^{1/r}$ are determined using a vanishing right-hand side and the stopping criteria $||ur||_{L2} < 10^{-12}$ or $r = 50$ where u_r denotes the current iterate. From these numbers we can observe an optimal level-independent convergence of our solver. The slight fluctuations in the measured rates are due to the parallelization of the block-Gauss-Seidel smoother, see [4, 6] for details on the parallelization.

For the approximation of more complicated geometries we reuse the tree-based cover construction to define so-called domain integration cells, see Figures 2 and 3. Note, however, that the resolution of the domain is not directly coupled to the cover construction.

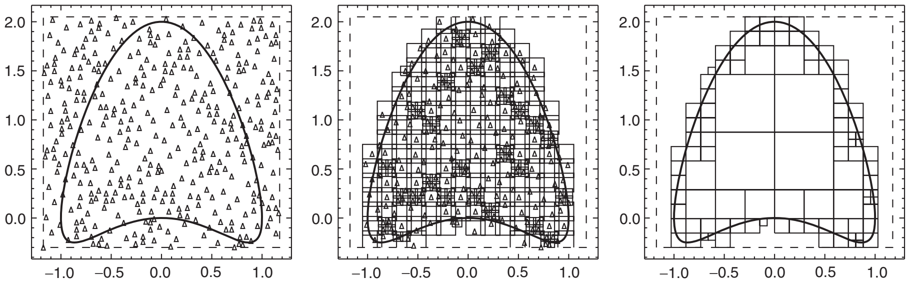


Figure 2. Domain approximation using the cover tree. Left: Domain and all sampling points. Center: Domain and constructed cover. Right: Domain and respective integration cells.

Currently, the approximation of the boundary of the domain is given by the boundary of the respective tree-cells. However, a higher order reconstruction of the boundary is straight-forward, see e.g., [8].

In summary, these results show that the PUM can be used effectively for the numerical solution of linear elliptic PDE, i.e., for unconstrained minimization problems. Let us now focus on the approximation of constrained minimization problems.

3. CONSTRAINED MINIMIZATION PROBLEMS

A classical example for such a minimization problem with constraints is the Poisson–Obstacle problem

$$\begin{aligned} -\Delta u \leq\ & f \quad \text{on } \Omega, \\ u =\ & 0 \quad \text{on } \partial\Omega, \\ u \leq\ & o \quad \text{on } \Omega, \\ (-\Delta u - f)(u - o) =\ & 0 \quad \text{on } \Omega, \end{aligned} \tag{6}$$

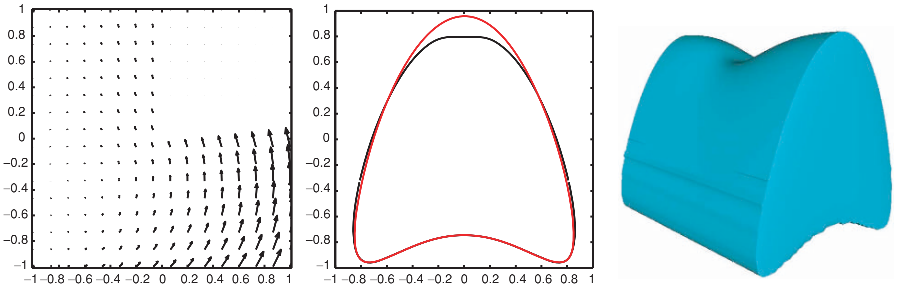


Figure 3. Displacement field on a L-shaped domain (left), comparison of deformed geometry (red) and original geometry (black) in two dimensions (center), and deformed geometry in three dimensions.

or the more involved Poisson–Signorini problem

$$\begin{aligned} -\Delta u &= f \quad \text{on } \Omega, \\ u &= g_D \quad \text{on } \Gamma_D \subset \partial\Omega, \\ \frac{\partial u}{\partial n} &= g_N \quad \text{on } \Gamma_N \subset \partial\Omega, \\ \frac{\partial u}{\partial n} \geq 0,\ u\frac{\partial u}{\partial n} \geq 0,\ u \geq 0 & \quad \text{on } \Gamma_C \subset \partial\Omega. \end{aligned} \tag{7}$$

In the obstacle problem (6) the constraints are enforced throughout the entire domain Ω, whereas in the more involved Signorini problem the constraints are enforced on a certain part of the boundary Γ_C, the contact boundary, only.

Let us consider the respective weak formulation of such a constrained problem. To this end, we define the classical energy function

$$\mathcal{J}(u) := \frac{1}{2}\alpha(u, u) - \langle f, u\rangle_{L^2}$$

associated with the underlying PDE problem, i.e., for (6) $a(u, v) = ||\nabla_u||^2_{L2}$ and restrict its minimization to the closed cone

$$K := \{v \in H^1 \mid v(x) \leq o(x) \text{ a.e. in } \Omega\}.$$

That is we are now looking for the minimum in a *convex* subset $K \subset H^1$ only, i.e., we try to find $u \in K$ such that $\mathcal{J}(u) \leq \mathcal{J}(v)$ for all $v \in K$ holds. This is unlike in the unconstrained minimization case where we are looking for the minimum in the linear space H^1. The discretization of this cone K of valid functions is the main issue in the numerical treatment of problems like (6) and (7). For instance, within the FEM pointwise conditions like

$$v(x) \leq o(x) \quad \text{for almost all } x \in \Omega$$

on the functions $v \in H^1$ are approximated in the vertices x_i of the mesh. Since the linear FEM shape functions are interpolatory, this approximation directly translates into a simple comparison of the coefficient vectors $\tilde{v}_h = (v_i) \in \mathbb{R}^n$ and $\tilde{o_h} = (o_i) \in \mathbb{R}^n$ associated with the discrete function $v_h \in V_h$ and the obstacle $o_h \in V_h$. Hence, in the FEM the cone K is usually discretized as

$$K_h := \{v_h \in V_h \mid v_h(x_i) \leq o_h(x_i)\} = \{\tilde{v}_h \in \mathbb{R}^n \mid v_i \leq o_i\}.$$

3.1 Partition of Unity Discretization

In the PUM, however, this approach is not valid since the shape functions of the PUM are non-interpolatory. Hence, we cannot directly compare the discrete coefficients of the current solution and the obstacle to determine whether the

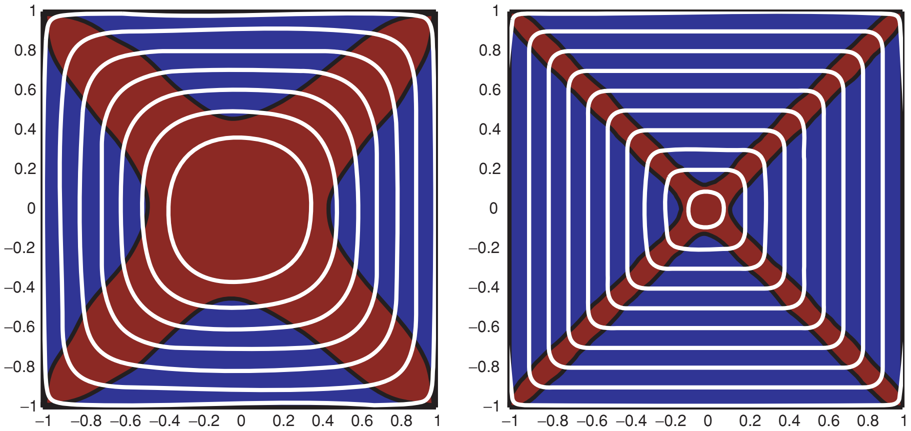


Figure 4. Active sets for right-hand side $f = 2.5$ (left) and $f = 10$ (right).

solution is valid. However, if we employ just linear local spaces $V_i^{pi=1}$ then we can easily compute the *minimum* and *maximal* function value of the difference $u_i - o_i$ locally on the patch ω_i. Hence, we discretize the closed cone K_{PUM} of valid functions within our PUM via

$$K_{\text{PUM}} := \{v \in V^{\text{PUM}} | \max_{\omega_i}(v_i - o_i) \leq 0 \quad \text{for all } i\}.$$

Here, we exploit the PU property of the functions φ_i to localize the pointwise conditions $v = \sum \varphi_i v_i \leq \sum_i \varphi_i o_i = o$ on the global shape functions $\varphi_i \psi_i^n$ to the local shape functions ψ_i^n patches ω_i.

3.2 Numerical Results for Obstacle Problem

Let us now present some numerical results obtained with our PUM for the obstacle problem (6) with $o = \text{dist}_\Omega$ on the unit square $[-1, 1]^2$. From the isoline plots depicted in Figure 4 we can observe that we capture the active set, i.e., the part of the domain where the solution actually coincides with the obstacle, very well.

4. CONCLUDING REMARKS

In this paper we presented the PUM and its application to unconstrained as well as constrained minimization problems in elasticity. The presented numerical results clearly indicate the applicability of the PUM in this context. An open question at this time is the extension of our multi-level solver to the constrained minimization case, where certain monotony properties are essential. The implementation of these properties within our PUM, however, is not trivial due to the fact that the sequence of PUM function spaces are non-nested and the bilinear form involves level-dependent regularization parameters due to the Nitsche approach for essential boundary conditions.

REFERENCES

1. M. Griebel and M.A. Schweitzer (2000), A particle-partition of unity method for the solution of elliptic, parabolic and hyperbolic PDE. *SIAM Journal on Scientific Computing*, 22, pp. 853–890.
2. M. Griebel and M.A. Schweitzer (2002), A particle-partition of unity method—Part II: efficient cover construction and reliable integration. *SIAM Journal on Scientific Computing*, 23, pp. 1655–1682.
3. M. Griebel and M.A. Schweitzer, (2002), A Particle-Partition of Unity Method—Part III: A Multilevel Solver. *SIAM Journal on scientific computing*, 24, pp. 377–409.
4. M. Griebel and M.A. Schweitzer (2002), A particle-partition of unity method—part IV: parallelization. In: *Meshfree Methods for Partial Differential Equations*, M. Griebel and M. A. Schweitzer (eds.), Vol. 26 of *Lecture Notes in Computational Science and Engineering*, Springer, pp. 161–192.
5. M. Griebel and M.A. Schweitzer (2002), A particle-partition of unity method—part V: boundary conditions. In: *Geometric Analysis and Nonlinear Partial Differential Equations*, S. Hildebrandt and H. Karcher (eds.), Springer, pp. 517–540.
6. M. A. Schweitzer (2003), A Parallel Multilevel Partition of Unity Method for Elliptic Partial Differential Equations, Vol. 29 of *Lecture Notes in Computational Science and Engineering*, Springer.
7. J. Nitsche (1970–1971), Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 36, pp. 9–15.
8. O. Klaas and M.S. Shepard (2000), Automatic Generation of Octree-based Three-Dimensional Discretizations for Partition of Unity Methods, *Computational Mechanics*, 25, pp. 296–304.

A MESHFREE APPROXIMATION WITH ALLMAN'S ROTATIONAL DOFS

R. Tian and G. Yagawa
Collaborative Research Center of Frontier Simulation Software for Industrial Science,
University of Tokyo, 4-6-1 Komaba, Meguroku, Tokyo, Japan 1538505

Abstract The Allman's rotational dofs are often used to develop advanced elements in modeling of shells as an assembly of flat elements in FEM. Yet the technique does not extend to the area of meshfree methods. In this paper we first present an extremely simple expression of the Allman's rotational dofs of 1984 and then develop a meshfree approximation containing the rotational dofs.

Keywords: meshfree, Allman's rotational dofs, moving least square.

1. A SIMPLIFIED EXPRESSION OF THE ALLMAN'S ROTATIONAL DOFS OF 1984

Allman in 1984 designed a triangle with rotational dofs (or alternatively called drilling freedom in plane elements) with successes not previously attained. The triangle has two translational dofs and one rotational dof per node and is defined as

$$\begin{aligned} u_e = \sum_{i=1}^{3} u_i \xi_i &+ \frac{1}{2} l_{12} \cos\gamma_{12}(\omega_2 - \omega_1)\xi_1\xi_2 \\ &+ \frac{1}{2} l_{23} \cos\gamma_{23}(\omega_3 - \omega_2)\xi_2\xi_3 + \frac{1}{2} l_{31} \cos\gamma_{31}(\omega_1 - \omega_3)\xi_3\xi_1 \end{aligned} \quad (1)$$

$$\begin{aligned} v_e = \sum_{i=1}^{3} v_i \xi_i &+ \frac{1}{2} l_{12} \sin\gamma_{12}(\omega_2 - \omega_1)\xi_1\xi_2 \\ &+ \frac{1}{2} l_{23} \sin\gamma_{23}(\omega_3 - \omega_2)\xi_2\xi_3 + \frac{1}{2} l_{31} \sin\gamma_{31}(\omega_1 - \omega_3)\xi_3\xi_1 \end{aligned} \quad (2)$$

G. R. Liu et al. (eds.), Computational Methods, 1495–1501.

where $l_{ij}(i, j = 1, 2, 3)$ are the length of the edge of the triangle, $\gamma_{ij}(i, j = 1, 2, 3)$ are the angles between the outward normal of the edge ij and the x-axis, and $\omega_i(i = 1, 2, 3)$ are the commonly called Allman's rotational dofs.

Equations (1) and (2) are simplified as follows. First it is easy to have the relation

$$l_{ij}\,\cos\gamma_{ij} = y_j - y_i, \quad (i, j = 1, 2, 3). \tag{3}$$

Substituting Equation (3) into Equation (1), we have

$$\begin{aligned} u_e = \sum_{i=1}^{3} u_i\xi_i &+ \frac{1}{2}((y_2 - y_1)(\omega_2 - \omega_1)\,\xi_1\xi_2 + (y_3 - y_2)(\omega_3 - \omega_2)\,\xi_2\xi_3 \\ &+ (y_1 - y_3)(\omega_1 - \omega_3)\,\xi_3\xi_1). \end{aligned} \tag{4}$$

After simple mathematical transformation, Equation (4) becomes

$$\begin{aligned} u_e &= \sum_{i=1}^{3} u_i\xi_i - \frac{1}{2}\omega_1\xi_1((y_2 - y_1)\xi_2 + (y_3 - y_1)\xi_3) \\ &\qquad - \frac{1}{2}\omega_2\xi_2((y_1 - y_2)\xi_1 + (y_3 - y_2)\xi_3) \\ &\qquad - \frac{1}{2}\omega_3\xi_3((y_1 - y_3)\xi_1 + (y_2 - y_3)\xi_2) \\ &= \sum_{i=1}^{3} u_i\xi_i - \frac{1}{2}\omega_1\xi_1\sum_{i=1}^{3}(y_i - y_1)\xi_i - \frac{1}{2}\omega_2\xi_2\sum_{i=1}^{3}(y_i - y_2)\xi_i \\ &\qquad - \frac{1}{2}\omega_3\xi_3\sum_{i=1}^{3}(y_i - y_3)\xi_i. \end{aligned} \tag{5}$$

Since

$$\sum_{j=1}^{3}(y_j - y_i)\xi_j = y - y_i = \hat{y}_i \quad (i = 1, 2, 3), \tag{6}$$

the Allman's original formulation (1) is finally simplified as

$$\begin{aligned} u_e &= \sum_{i=1}^{3} u_i\xi_i - \frac{1}{2}\omega_1\xi_1\hat{y}_1 - \frac{1}{2}\omega_2\xi_2\hat{y}_2 - \frac{1}{2}\omega_3\xi_3\hat{y}_3 \\ &= \sum_{i=1}^{3} \xi_i\left(u_i - \frac{1}{2}\hat{y}_i\omega_i\right). \end{aligned} \tag{7}$$

Similarly the Allman's formulation (2) is simplified as

$$v_e = \sum_{i=1}^{3} \xi_i \left(v_i + \frac{1}{2} \hat{x}_i \omega_i \right) \tag{8}$$

where $\hat{x}_i = x - x_i$. In a matrix form we finally express the triangle with the Allman's rotational dofs as

$$\mathbf{u}_e = \sum_{i=1}^{3} \xi_i \left[\begin{array}{cc|c} 1 & 0 & -\hat{y}_i/2 \\ 0 & 1 & \hat{x}_i/2 \end{array} \right] \begin{pmatrix} u_i \\ v_i \\ \omega_i \end{pmatrix}. \tag{9}$$

Compared with the usual nodal displacements,

$$\mathbf{u}_i = \begin{pmatrix} u_i \\ v_i \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} u_i \\ v_i \end{pmatrix}, \tag{10}$$

we can define a generalized form of the nodal displacements $\mathbf{u}_i$ as

$$\mathbf{u}_i = \left[\begin{array}{cc|c} 1 & 0 & -\hat{y}_i/2 \\ 0 & 1 & \hat{x}_i/2 \end{array} \right] \begin{pmatrix} u_i \\ v_i \\ \omega_i \end{pmatrix} = \mathbf{R}_i \mathbf{D}_i \tag{11}$$

in which the Allman's rotational dof is contained. Now in a different but much easier way one can obtain the same element by just replacing the conventional nodal displacement (10) with the generalized node (11). Equation (11) is termed a generalized node that can be used as a substitute of the original Allman's formulae [2, 3]. Reader may also be referred to Tian, Matsukara and Yagawa [4] for an extended discussion in the context of the tetrahedrons with Allman's rotational dofs.

2. A MESHFREE APPROXIMATION WITH THE ALLMAN'S ROTATIONAL DOFS

Denote $\Omega \subset \mathrm{R}^d$, $d = 1, 2, 3$, be the domain of interest, $\{\mathbf{x}_i\}$ be nodes generated on the domain, $\mathrm{D}_r(\mathbf{x}_i)$ be a d-dimension open ball of radius r_i and center $\mathbf{x}_i$, $D_r(\mathbf{x}_i) = \{\mathbf{x} : ||\mathbf{x}{-}\mathbf{x}_i|| < r_i\}$. $\{r_i\}$ are set so that any point $\mathbf{x}\in \Omega$ has only finitely many i of $\mathbf{x}\in\mathrm{D}_r(\mathbf{x}_i)$. Define unknowns at the node i be (u_i, v_i, ω_i), i.e., two translational dofs (u_i, v_i) and a Allman's rotational dof ω_i.

A linear approximation of the field function $\mathbf{u(x)}$ is assumed,

$$\mathbf{u}^h(\mathbf{x}) = \sum_{i=1}^{m} p_i(\mathbf{x}) a_i(\mathbf{x}) = \mathbf{p}^{\mathrm{T}}(\mathbf{x})\mathbf{a}(\mathbf{x}), \quad \mathbf{p}^{\mathrm{T}} = [1, x, y], \quad \mathbf{x} \in \Omega, \tag{12}$$

where **a** are the unknown parameters. At an arbitrary point **x**, a Moving Least Square (MLS) fitting is performed locally with respect to the nodes $\{\mathbf{x} | \mathbf{x} \in D_r(\mathbf{x}_i)\}$ to determine the unknowns. A discrete L_2 error norm of the fitting is defined by,

$$J(\mathbf{x}) = \sum_{I=1}^{n} w_I(\mathbf{x})(\mathbf{P}^{\mathrm{T}}(\mathbf{x}_I)\mathbf{a}(\mathbf{x}) - \mathbf{u}_I)^2, \quad \mathbf{x} \in D_r(\mathbf{x}_I), \tag{13}$$

where n is the dimension of the node set $\{\mathbf{x}_i | \mathbf{x} \in D_r(\mathbf{x}_I)\}$, $w_I(\mathbf{x})$ is the weight function. The MLS fitting among the generalized nodes becomes accordingly,

$$J(\mathbf{x}) = \sum_{I=1}^{n} w_I(\mathbf{x})(\mathbf{P}^{\mathrm{T}}(\mathbf{x}_I)\mathbf{a}(\mathbf{x}) - \mathbf{R}_I\mathbf{D}_I)^2, \quad \mathbf{x} \in D_r(\mathbf{x}_I). \tag{14}$$

Minimizing the L_2 norm directly leads to the solution to the unknown **a** like,

$$\mathbf{a} = \mathbf{A}^{-\mathbf{1}}\mathbf{B}\mathbf{R}_I\mathbf{D}_I, \tag{15}$$

$$\mathbf{A} = \sum_{i=1}^{n} w_i(\mathbf{x})\mathbf{p}(\mathbf{x}_i)\mathbf{p}^{\mathrm{T}}(\mathbf{x}_i), \mathbf{B} = [w_1(\mathbf{x})\mathbf{p}(\mathbf{x}_1), w_2(\mathbf{x})\mathbf{p}(\mathbf{x}_2), \ldots, w_n(\mathbf{x})\mathbf{p}(\mathbf{x}_n)]. \tag{16}$$

Back substitute **a** into Equation (12), a meshfree approximation with the Allman's rotational dofs is constructed

$$\mathbf{u}^h(\mathbf{x}) = \sum_{I=1}^{n} \phi_I(\mathbf{x})\mathbf{R}_I\mathbf{D}_I, \quad \phi_I(\mathbf{x}) = \mathbf{p}\mathbf{A}^{-1}\mathbf{B}. \tag{17}$$

Instead the new meshfree approximation (17) can also be directly obtained by replacing the conventional nodal displacement with the generalized node (11) as follows:

$$\mathbf{u}^h(\mathbf{x}) = \sum_{I=1}^{n} \phi_I(\mathbf{x})\mathbf{u}_I = \sum_{I=1}^{n} \phi_I(\mathbf{x})\mathbf{R}_I\mathbf{D}_I, \tag{18}$$

where $\mathbf{u}_I = (u_i, v_i)^{\mathrm{T}}$ denotes the conventional node. The rotational dof is also applicable to the approximations based on other meshfree techniques in the similar way.

3. GOVERNING EQUATIONS

Consider a two-dimensional, small strain, linear elastic problem on the domain Ω,

$$\nabla \cdot \sigma + \mathbf{b} = 0 \quad \text{in } \Omega \tag{19}$$

$$\varepsilon = \nabla_s \mathbf{u}, \ \sigma = \mathbf{C} : \varepsilon \tag{20}$$

$$\mathbf{u} = \bar{\mathbf{u}} \quad \text{on } \Gamma_u \tag{21}$$

$$\sigma \cdot \mathbf{n} = \bar{\mathbf{t}} \quad \text{on } \Gamma_\text{t} \tag{22}$$

where σ is the Cauchy stress tensor for a displacement field **u, b** the body force, **n** is the unit outward normal vector to traction boundary Γ_t, ∇_s is the symmetric gradient operator, **C** is the elastic module tensor. The superposed bar denotes the prescribed values at the boundaries. The modified variational principle with penalty treatments of the essential boundary condition (21) is established as the following:

$$\int_\Omega \nabla_s \delta\mathbf{u} : \sigma d\Omega - \int_\Omega \delta\mathbf{u} \cdot \mathbf{b}\, d\Omega - \int_{\Gamma_\text{t}} \delta\mathbf{u} \cdot \bar{t} d\Gamma + \mathbf{KT} \int_{\Gamma_u} \delta\mathbf{u} \cdot (\mathbf{u} - \bar{\mathbf{u}})\, d\Gamma = 0 \tag{23}$$

where **k** and **T** are defined as

$$\mathbf{k} = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{or} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{24}$$

where **T** is used to customize the penalty matrix **k** according to boundary constraint. Discretizing Equation (24) we obtain the system of equations to be solved

$$\mathbf{KU} = \mathbf{F} \tag{25}$$

where

$$\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)^\text{T}, \mathbf{u}_I = (u_I, v_I, \omega_I)^\text{T}, \tag{26}$$

$$\mathbf{K} = \int_\Omega \mathbf{B}^\text{T}\mathbf{DB}\, \text{d}\Omega + \int_{\Gamma_u} \mathbf{N}^\text{T}\mathbf{kTN}\, \text{d}\Omega,$$

$$\mathbf{F} = \int_\Omega \mathbf{N}^\text{T}\mathbf{b}\, d\Omega + \int_{\Gamma_\text{t}} \mathbf{N}^\text{T}\bar{\mathbf{t}}\, d\Gamma + \int_{\Gamma_u} \mathbf{N}^\text{T}\mathbf{kT}\bar{\mathbf{u}}\, d\Gamma, \tag{27}$$

$$\mathbf{N} = [\mathbf{N}_1\mathbf{R}_1, \mathbf{N}_2\mathbf{R}_2, \ldots, \mathbf{N}_N\mathbf{R}_N], \ \mathbf{N}_I = \begin{bmatrix} \phi_I & 0 \\ 0 & \phi_I \end{bmatrix},$$

$$\mathbf{R}_I = \left[\begin{array}{cc|c} 1 & 0 & -\hat{y}_I/2 \\ 0 & 1 & \hat{x}_I/2 \end{array}\right] \tag{28}$$

$$\mathbf{B} = \nabla_s\mathbf{N}, \nabla_s(\mathbf{N}_I\mathbf{R}_I) = \begin{bmatrix} \phi_{I,x} & 0 \\ 0 & \phi_{I,y} \\ \phi_{I,y} & \phi_{I,x} \end{bmatrix} \mathbf{R}_I + 0 = \mathbf{B}_I^0\mathbf{R}_I, \tag{29}$$

D is the material matrix, N is the number of the nodes on the whole domain.

It may be seen from Equation (29) that the new stiffness matrix can be obtained by the simply following matrix transformation to that without the rotational dofs,

$$\mathbf{B}^\text{T}\mathbf{DB} = \mathbf{R}^\text{T}\mathbf{B}_0^\text{T}\mathbf{DB}_0\mathbf{R}, \ \mathbf{B}_0 = [\mathbf{B}_1^0, \mathbf{B}_2^0, \ldots, \mathbf{B}_N^0],$$

$$\mathbf{R} = [\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_N]^\text{T}. \tag{30}$$

3.1 Zero Energy Modes

By setting all the rotational dofs to the same non-zero constant while letting all the translational dofs vanish, one zero energy mode can be identified as follows,

$$\omega_1 = \omega_2 = \cdots = \omega_n = \omega \neq 0, \tag{31}$$

$$\mathbf{u}^h(\mathbf{x}) = \sum_{I=1}^{n} \phi_I(\mathbf{x}) \left[\begin{array}{cc|c} 1 & 0 & -\hat{y}_I/2 \\ 0 & 1 & \hat{x}_I/2 \end{array}\right] \begin{pmatrix} 0 \\ 0 \\ \omega \end{pmatrix} = \mathbf{0}. \tag{32}$$

It is also shown in Equation (32) that the zero-energy mode does not induce any deformation and is also a zero displacement mode. It thus can be suppressed by prescribing an arbitrary rotational dof to an arbitrary constant say simply zero.

4. NUMERICAL EXAMPLES

Two numerical examples are presented in the conference separately due to the space limitation here, Eigenvalue examination of the stiffness matrix of the meshfree approximation with the rotational dof detects one zero spurious mode that has been identified in Section 3.1. The mode, however, can be suppressed easily as discussed in Section 3.1. The other example, the standard cantilever benchmark, shows an evident improvement in both accuracy and convergence of the meshfree approximation by equipping the rotational dof. We remark that the performance of the meshfree approximation developed should lie between the standard MLS approximations of the linear and completely second bases.

5. CONCLUSIONS

(1) An extremely simple expression of the Allman's rotational dofs has been found.
(2) A meshfree approximation containing the Allman's rotations has been presented.

REFERENCES

1. D.J. Allman (1984), A compatible triangular element including vertex rotations for plane elasticity analyses. *Computers and Structures*, 19, pp. 1–8.

2. R. Tian and G. Yagawa (2005), Generalized nodes and high performance elements. *International Journal for Numerical Methods in Engineering*, 64(15), pp. 2093–2071.
3. R. Tian and G. Yagawa (2005), Allman's triangle and partition of unity. *International Journal for Numerical Methods in Engineering* (submitted for publication).
4. R. Tian, H. Matsubara, and G. Yagawa (2005), Advanced 4-node tetrahedrons. *International Journal for Numerical Methods in Engineering*, (accepted).

THE VORTEX METHOD APPLIED TO SIMULATION OF HOMOGENEOUS ISOTROPIC TURBULENCE

Y. Totsuka and S. Obi
Department of Mechanical Engineering, Keio University, Yokohama 223-8522, Japan

Abstract The present study examines two representative viscous diffusion models in combination with the vortex method. It is shown that the MPS Laplacian model is superior to the conventional core spreading model in terms of calculating the decay rate of enstrophy and energy spectra. Computational time is remarkably reduced by using fast multi-pole method while retaining accuracy.

Keywords: turbulent flow, vortex method, viscous diffusion, dissipation, MPS.

1. INTRODUCTION

Vortex methods have a lot of advantages in engineering application because of the grid-free nature. Problems comprizing complex geometry, moving and/or deforming boundary are typical applications where the vortex method exhibits its advantage over Eulerian type of approach. Because only the motion of vortex elements is considered, computational efforts are remarkably reduced in, e.g., aerodynamics where inviscid flow occupies major part of the domain. Flow around a blade of wind turbine is a good example of the unsteady flow with moving boundary [1].

Apart from the engineering application, it is of great importance for any kind of computational method that it satisfies the law of conservation. In turbulent flow computations, it is the kinetic energy conservation that should always be kept in mind. Up to now, there are only few investigations that consider the appropriateness of the representation of energy dissipation due to viscosity. A study on the interference of a pair of vortex ring has demonstrated the ability of the vortex method to express the inertial sub-range, which is of a great importance in dynamics of turbulence [2], while the so-called Vortex in Cell method

G. R. Liu et al. (eds.), Computational Methods, 1503–1508.

[3], is reported to be useful for simulation of homogeneous isotropic turbulence. Nevertheless, systematic investigations on the performance of commonly used diffusion model are necessary.

In the present study, the two-dimensional homogeneous isotropic turbulence is considered, where the energy cascade due to vortex stretching does not occur. The energy decay directly reflects the viscous dissipation; hence pure assessment of the viscous diffusion model is possible.

2. OUTLINE OF THE VORTEX METHOD

2.1 Basic Equation

The vortex method handles with a system of ordinary differential equations derived from vorticity transport equation. The location and intensity of individual vortex elements are calculated from the following equations for individual vortex elements:

$$\frac{dx_i}{dt} = \mathbf{u}_i, \tag{1}$$

$$\frac{d\omega_i^s}{dt} = \left(\nu \nabla^2 \omega^s\right)_i \tag{2}$$

The right hand side of Equation (2) expresses the diffusion of vortex intensity ω^s due to molecular viscosity, with ν being the kinematic viscosity. The approximation of this term is considered in the present study; i.e., the appropriateness of the diffusion model in representing the dissipation of kinetic energy.

2.2 Core Spreading Model

The core spreading model [4] is most popular in vortex method and has been adopted for various engineering applications. The model stems from the expression of individual vortex element by core-function with core radius ϵ_i. The use of this model replaces Equation (2) by the equation below for the rate-of-change of ϵ_i:

$$\frac{d\epsilon_i}{dt} = \frac{4\nu}{2\epsilon_i}. \tag{3}$$

2.3 MPS Laplacian Model

The Laplacian model developed by Koshizuka and Oka [5] for the purpose Moving Particle Semi-implicit (MPS) scheme may be applied for the vortex method simulation. The MPS mosel approximates the right hand side of Equation (2) as follows:

$$\left(\nu\nabla^2\omega^s\right)_i = \frac{4\nu}{\lambda_i\bar{n}}\sum_{j\neq i}\left[\left(\omega_j^s - \omega_i^s\right) w\left(r_{ij}\right)\right], \tag{4}$$

where λ_i is a function controlling the number density of vortex elements. $\bar{n}$ represents the average number density of elements. The weighting function $w(r_{ij})$ is evaluated by

$$w(r_{ij}) = \begin{cases} \dfrac{r_e}{r_{ij}} - 1 & (r_{ij} \leq r_e) \\ 0 & (r_{ij} \leq r_e) \end{cases}, \tag{5}$$

with r_{ij} being the position vector connecting i^{th} and j^{th} elements. The parameter r_e indicates the limit to distance that is to be determined according to the number density of elements.

Because the MPS Laplacian model becomes erroneous in case of heterogeneous distribution of elements, we controlled the number density at every time step and kept certain level of homogeneous distribution by adding/subtracting the elements when necessary.

3. COMPUTATION OF THE 2D HOMOGENEOUS ISOTROPIC TURBULENCE

A energy spectrum distribution is specified as the initial condition for the calculation according to Tatsumi and Yanase [6]:

$$E(k) = \frac{\overline{u_0^2}k^3}{k_0^4}\exp\left(-\frac{k^2}{k_0^2}\right), \tag{6}$$

where u_0 is fluctuating velocity and k_0 stands for the wave number at which the energy spectrum reaches its maximum. The present study specified $k_0 = 4$. The given energy spectrum was expressed by the vortex elements which were first distributed on an equidistant grid on the $x{-}y$ plane. The intensity of individual elements was corrected so that the total energy becomes equal to that of the initial distribution.

The computations were performed in a rectangle domain with each face equal to 2π at two different Reynolds numbers: $Re = 1300$ and 260, with Re being the Reynolds number based on the length of the computation domain and the RMS of the initial velocity distribution. To accomplish the infinitely large field for homogeneous turbulence, a finite number of domains were used and the results are obtained from the domain located in the center. Preliminary computations by changing the number of domains indicated that 81 × 81 domains are sufficient to achieve the periodicity in every direction; the estimated error with respect to the values obtained by 2561 × 2561 domains was less than $O(10^{-6})$, based on the norm of second order.

4. RESULTS AND DISCUSSION

4.1 Computational Details

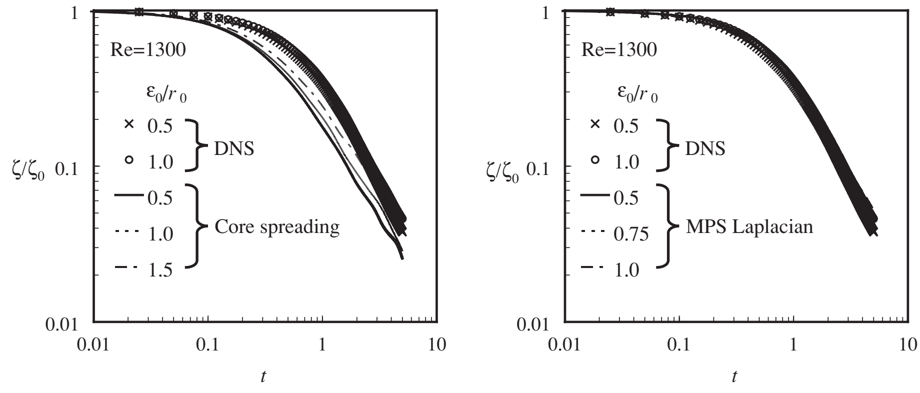
Equations (1) and (2) were solved by means of second-order Adams method. The diffusion term was also treated in explicit manner. The time step was kept constant to $\Delta t = 10^{-3}$, and computations were performed for 5000 steps which corresponded to approximately 10 times the eddy turnover time. For the MPS Laplacian model, r_e was set equal to three times the initial distance between regularly distributed elements.

4.2 Enstrophy

The energy decay is represented by the enstrophy ζ evaluated by integration of whole computation domain as

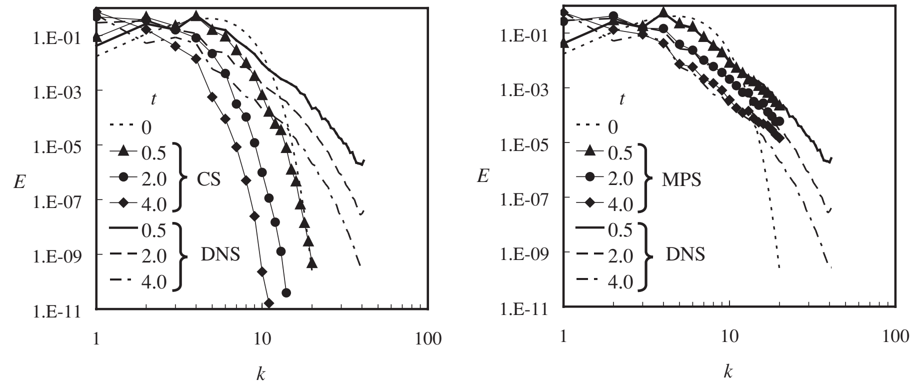
$$\zeta = \frac{1}{2}\int \omega^2\, ds. \tag{7}$$

Figure 1(a) compares the decay of ζ for DNS and core spreading model. The vertical axix is normalized by the initial value ζ_0. The core spreading model



(a) (b)

Figure 1. Decay of enstrophy at $Re = 1300$.



(a) (b)

Figure 2. Energy spectra at $Re = 1300$.

provides too early decay compared to DNS. The tendency is unaffected by the change of initial core radius ϵ_0 which was varied from 0.5 r_0 to 1.5 r_0. On the other hand, the MPS Laplacian model agrees well with DNS as shown in Figure 1 (b).

4.3 Energy Spectra

Figure 2 presents the evolution of energy spectra, comparing the results of two diffusion models and DNS. The core spreading model shown in Figure 2 (a) does not represent the distribution in higher wave number range as DNS does. On the other hand, the MPS Laplacian model in 8 2 (b) well follows the variation calculated by DNS. The change in ϵ_0/r_0 did not alter the tendency given by either model. The apparent superiority of the MPS Laplacian model as demonstrated above reflects the fact that the addition of vortex elements during the simulation is necessary to represent the increase of energy in high wave number range. Such treatment could also be introduced to the core spreading model, though the vorticity of the individual element is too difficult to control.

5. CONCLUDING REMARKS

The vortex method is capable of simulating two-dimensional homogeneous isotropic turbulence. The MPS Laplacian model is superior to the core spreading model in representing increase of energy spectra in high wave number range that is indicated by DNS. The representation of the interaction between vortex elements seems to be the key for realistic calculation of dissipation process.

REFERENCES

1. A. Ojima and K. Kamemoto (2001), Numerical simulation of unsteady flow through a horizontal axis wind turbine by a vortex method. In: *The Second International Conference on Vortex Methods*, Istanbul, pp. 173–180.
2. M. Kiya and H. Ishii (1991), Vortex interaction and Kolmogorov spectrum. *Fluid Dynamics Research*, 8, pp. 73–83.
3. G.-H. Cottet, B. Michaux, S. Ossia and G. VanderLinden (2002), A comparison of spectral and vortex methods in three-dimensional incompressible flows. *Journal of Computational Physics*, 175, pp. 702–712.
4. A. Leonard (1980), Vortex methods for flow simulation. *Journal of Computational Physics*, 37, pp. 289–335.
5. S. Koshizuka and Y. Oka (1996), Moving-particle semi-implicit method for fragmentation of incompressible fluid. *Nucl. Sci. Eng.*, 123, pp. 421–434.
6. T. Tatsumi and S. Yanase (1981), The modified cumulant expansion for two-dimensional isotropic turbulence, *Journal of Fluid Mechanics*, 110, pp. 475–496.

AN APPLICATION OF THE LOCAL PETROV–GALERKIN METHOD IN SOLVING GEOMETRICALLY NONLINEAR PROBLEMS

Y.B. Xiong[1], S.Y. Long[1], D.A. Hu[1] and G.Y. Li[2]
[1]*Department of Engineering Mechanics, Hunan University, Changsha 410082,China*
[2]*College of Mechanical and Automotive Engineering, Hunan University, Changsha 410082, China*

Abstract The local Petrov–Galerkin method (MLPG) is applied to solve the geometrically nonlinear problem in this paper. The local Petrov–Galerkin method uses the moving least square approximation as a trial function, and involves only integrations over a regular local subdomain and on a local subboundary centered at a node in question. These special properties lead to a more convenient formulation in dealing with nonlinear problems. An incremental and iterative solution procedure is used to solve the geometrically nonlinear problem. Formulations for the geometrically nonlinear problems are obtained from virtual work principle. All measures are related back to the original configuration. Several examples are given to show that the local Petrov–Galerkin method has good accuracy in solving the geometrically nonlinear problem.

Keywords: local Petrov–Galerkin method, moving least square approximation, the geometrically nonlinear problem, virtual work principle.

1. INTRODUCTION

The local Petrov–Galerkin method is a new numerical technique presented in recent years. It requires only nodal information and no element connectivity is needed. Nodes can be randomly distributed in a domain under consideration. The local Petrov–Galerkin method uses the moving least square (MLS) as an approximation function, which is different from the finite element method. MLPG constructs field function based on scatter nodes. The local Petrov–Galerkin method is similar to the element-free Galerkin method

G. R. Liu et al. (eds.), Computational Methods, 1509–1514.

to some extent except integral format, as it involves only integrations over a regular local subdomain and on a local subboundary centered at a node in question. This method is more flexible in solving nonlinear problems than the finite element method, the boundary element method and the element-free Galerkin method.

Atluri *et al.* proposed firstly the local Petrov–Galerkin method and applied it to solving two-dimensional potential problems and bending problems of beams [1, 2]. Then Long applied MLPG to 2-D elasticity problems and bending problems of a thin plate [3, 4]. In this paper, MLPG is developed to solve the 2-D geometrically nonlinear problems.

2. LOCAL PETROV–GALERKIN INTEGRAL FORMULATIONS AND DISCRETIZATION EQUATIONS

The linearization meshless local Petrov–Galerkin integral equation is given as follows for 2-D geometrically nonlinear problems

$$\begin{aligned}
&\int_{0_{\Omega_s}} {}_0D_{ijkl}\, {}_0e_{kl}\, {}_0^tB_L w_i\, {}^0d\Omega_s + \int_{{}^0\Omega_s} {}_0^tS_{ij}\, {}_0^tB_{NL} w_i\, {}^0d\Omega_s \\
&\quad + 2\alpha \int_{{}^0\Gamma_{su}} w_i(u_i - \bar{u}_i)^0 d\Gamma = {}^{t+\Delta t}Q_s - \int_{{}^0\Omega_s} {}_0^tS_{ij}\, {}_0^tB_L w_i\, {}^0d\Omega_s \qquad (1)
\end{aligned}$$

$$ {}^{t+\Delta t}Q_s = \int_{\partial^0\Omega_s} {}_0^{t+\Delta t} t_k w_i\, {}^0d\Gamma_s + \int_{{}^0\Omega_s} {}^0\rho_0^{t+\Delta t} b_k w_i\, {}^0d\Omega_s \qquad (2)$$

where ${}_0D_{ijkl}$ is a tangent constitutive tensor, α a penalty parameter used to impose the essential boundary conditions and w_i the weight function of the moving least square approximation. ${}_0e_{ij}$ and ${}_0\eta_{ij}$ are the linear and quadratic terms of incremental displacements u_i; ${}_0^tS_{ij}$ is the Kirchhoff stress tensor obtained by last solution step; ${}^0\rho$ is density; ${}^{t+\Delta t}_0t_k$ is the surface traction and ${}^{t+\Delta t}_0b_k$ is the body force; ${}^{t+\Delta t}Q_s$ is the virtual work of external loads in subdomain ${}^0\Omega_s$; ${}^0\Gamma_{su}$ is a part of the local boundary $\partial^0\Omega_s$, over which displacement increments are prescribed. Where symbol "0" at the upper or lower left corner denotes that the initial configuration is selected as a reference configuration in MLPG formulations.

Using MLS interpolation function to approximate the displacement increments $\boldsymbol{u}_i$, the discretization equation can be written in matrix form

$$ {}_0^t\boldsymbol{K}\boldsymbol{u} = \mathbf{f} \qquad (3)$$

where

$$
\begin{aligned}
{}^t_0\boldsymbol{K} &= \int_{{}^0\Omega_s} {}^t_0\boldsymbol{B}_L^{\mathrm{T}}\ {}_0\boldsymbol{D}\ {}^t_0\boldsymbol{B}_L \boldsymbol{W}\,{}^0\mathrm{d}\Omega_s - \int_{{}^0\Gamma_{su}} {}^{t+\Delta t}_{0}\boldsymbol{t}_i \boldsymbol{W}^0\mathrm{d}\Gamma \\
&\quad + \int_{{}^0\Omega_s} {}^t_0\boldsymbol{B}_{NL}^{\mathrm{T}}\ {}^t_0\boldsymbol{S}\ {}^t_0\boldsymbol{B}_{NL}\boldsymbol{W}\,{}^0\mathrm{d}\Omega_s + 2\alpha \int_{{}^0\Gamma_{su}} \boldsymbol{W}\boldsymbol{C}\ {}^t_0\boldsymbol{\Phi}_j\ {}^0\mathrm{d}\Gamma \qquad (4)\\
\boldsymbol{f} &= \int_{{}^0\Gamma_{st}} {}^{t+\Delta t}_{0}\bar{\boldsymbol{t}}_i \boldsymbol{W}\,{}^0\mathrm{d}\Gamma + \rho \int_{{}^0\Omega_s} {}^{t+\Delta t}_{0}\boldsymbol{b}_i \boldsymbol{W}\,{}^0\mathrm{d}\Omega_s \\
&\quad + 2\alpha \int_{{}^0\Gamma_{su}} \boldsymbol{W}\boldsymbol{C}\ {}^t_0\bar{\boldsymbol{u}}^0\mathrm{d}\Gamma - \int_{{}^0\Omega_s} {}^t_0\boldsymbol{B}_L^{\mathrm{T}}\ {}^t_0\hat{\boldsymbol{S}}\ {}^0\boldsymbol{W}\,\mathrm{d}\Omega_s \qquad (5)
\end{aligned}
$$

and C is the parameter of boundary; ${}^t_0\boldsymbol{B}_L$ and ${}^t_0\boldsymbol{B}_{NL}$ are the transformation matrices between linear strain ${}_0e_{ij}$, nonlinear strain ${}_0\eta_{ij}$ and displacements, respectively. ${}_0\boldsymbol{D}$ is the constitutive matrix of material. ${}^t_0\boldsymbol{S}$, ${}^t_0\hat{\boldsymbol{S}}$ are the matrix and vector of Kirchhoff stresses, respectively. ${}^t_0\boldsymbol{\Phi}_i$ is the shape function of MLS approximation. All these measures refer to the initial configuration.

In the present computation, both quartic spline weight function [1] and exponential weight function proposed by Zhou [5] are employed.

3. NUMERICAL EXAMPLES

Example 1. A cantilever beam is now analysed under uniformly distributed loads, as shown in Figure 1. Geometric and material parameters are: $\mu = 0.2$, $L = 10.0\,\mathrm{cm}$, $h = 1.0\,\mathrm{cm}$, one unit thickness, $E = 1.2 \times 10^4\,\mathrm{N/cm}^2$. 60(12 × 5) nodes are regularly distributed. The radius of support domain r_i is equal to 5cm. The quartic spline weight function is used as the test function and its radius is equal to 0.2 which is also the radius of a subdomain. 10 and (8 × 8) Gauss integral points are adopted on the boundary Γ_s and in subdomain Ω_s, respectively. Loads are always kept at the vertical direction and linear elastic material is considered here. In the solution process, 30

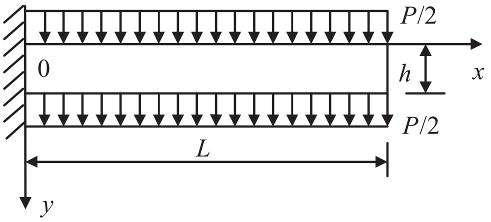


Figure 1. A cantilever beam under uniformly distributed loads.

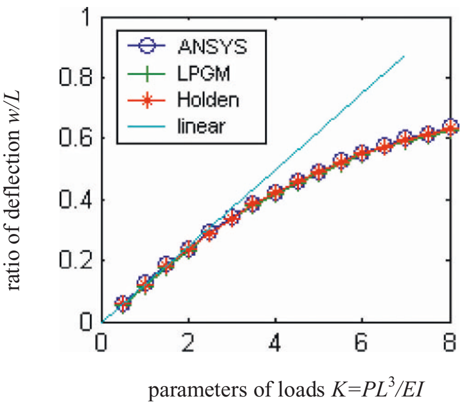
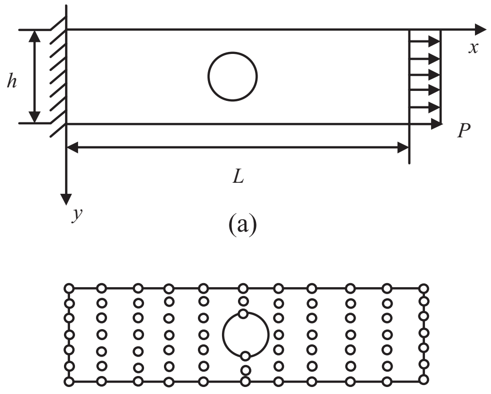


Figure 2. Results of a cantilever beam under uniformly distributed loads.



(a)

(b)

Figure 3. A holed square body subjected to a tension.

steps of load are imposed and equilibrium iterations are used in each load step. The results obtained are shown in Figure 2. The results show that responses in the large deformation appear more rigid than in the small deformation and results of MLPG are in good agreement with those of Holden and ANSYS.

Example 2. A holed square body subjected to a tension is analysed, as shown in Figure 3. $L = 10.0\,\text{m}$ $h = 4.0\,\text{m}$, $\mu = 0.2$, $E = 1.2 \times 10^4$ N/m^2, one unit thickness, the radius of a hole is 1.0 m, and the centre of the hole is at (5.0, 2.0). One end of the square body is fixed and the other end is subjected to a uniform tension P = 1000 N/m. We use 76 nodes, as shown in Figure 3(b). All

Table 1. Geometrically nonlinear solutions of a holed square body subjected to a tension.

Coordinate (m)	u (mm)		v (mm)	
	ANSYS	LPGM	ANSYS	LPGM
(1,0)	90.789	91.014	22.452	22.754
(5,0)	565.670	566.210	155.280	156.740
(10,0)	1129.900	1130.420	30.628	31.532
(1,4)	90.789	90.565	−22.432	−22.950
(5,4)	563.630	564.630	−153.130	−154.750
(10,4)	1129.300	1130.600	−27.920	−28.643

other data are the same as in example 1. Results obtained are shown in Table 1. Results of MLPG show good agreement with those of ANSYS.

4. CONCLUSION

The formulation and implementation of the Local Petrov–Galerkin method (MLPG) have been presented in this work for the geometrically nonlinear problem. The quartic spline function and the exponential weight function serve as the test function and the trial function in MLS approximation, respectively. The results of numerical examples show that MLPG can solve the geometrically nonlinear problems effectively and accurately.

ACKNOWLEDGEMENTS

The support of this work by Natural Science Foundation of China (No. 10372030), National 863 Project of China (No. 2003 AA 411230) and Foundation of The Ministry of Education of China (No. 20020532021, [2002] 350) is gratefully acknowledged.

REFERENCES

1. A. Majorek (1992), The Influence of Heat Transfer on the Development of Stresses and Distortions in Martensitically Hardened SAE 1045 and SAE 4149, In: *Proceedings of the First International Conference on Quenching and Distortion*, Chicago, USA, p. 171.
2. C. Heming (2002), Solution of an inverse problem of heat conduction of 45 steel with martensite phase transformation in high pressure during gas quenching, *Journal of Materials Science and Technology*, 18, 4, p. 372.

3. C. Heming (1999), Calculation of the residual stress of a 45 steel cylinder with a non-linear surface heat-transfer coefficient including phase transformation during quenching, *Journal of Materials Processing Technology*, 80–90, p. 339.
4. H. Ghoneim (1986), Constitutive modeling and thermoviscoplasticity, *Journal of Thermal Stress*, 9, p. 345
5. S.R. Bobner and Y. Par (1975), A non-linear constitutive modeling of thermoviscoplasticity, *ASME Journal of Applied Mechanics*, 42, p. 385.
6. Ivan Tzitzelkov and H. Paul (1974), Eine mathematische methode zur beschreibung des umwandlungsverhaltens eutektoidischer staehle, *Archiv Eisenhuttenwes*, 45, 5, p. 254.

ELASTIC-PLASTIC LARGE DEFORMATION ANALYSIS USING SPH

Y. Sakai[1] and A. Yamasita[2]
[1]*Faculty of Education and Human Sciences, Yokohama National University, Japan*
[2]*Faculty of Engineering, Nihon Bunri University, Japan*

Abstract Smoothed particle hydrodynamics (SPH) is extended to the strain rate independent elastic-plastic large deformation analysis for low velocity contact problems. The volume constant condition is imposed on the plastic deformation process using a pressure equation given by the particle density condition in a unit volume. Test problems show that these improvements lead to good stability and accuracy of large deformation analysis.

Keywords: particle method, SPH, large deformation, elastic-plastic.

1. INTRODUCTION

During the past decade there has been a effort to apply SPH (Smoothed Particle Hydrodynamics) to the problem of high velocity impact. The superior points of SPH for high velocity impact is that it is based on Lagrangian form and is substantially one of meshless methods. SPH technique can easily be applied to severe distortions and fragmentation process of impact bodies. However there has been almost no study for applying SPH to the elastic-plastic large deformation analysis of relatively low velocity contact problems. This paper presents 2D and 3D elastic-plastic large deformation algorithms using SPH method, where the D_{p} matrix method and the volume constant condition are used for the plastic process of materials.

2. THE SPH METHOD

The foundation of SPH [1] is one of the interpolation technique. The equation of motion and the conservation laws of substance, in the form of partial

G. R. Liu et al. (eds.), Computational Methods, 1515–1519.

differential equations, are introduced into integral equations through the use of an interpolation function (weight function) that gives the estimate of the field variables at a point. In numerical process, information is given at discrete points, so that the integrals are evaluated as summing over neighboring particles. Consider a function f and a kernel W which has a radius (support domain) h, the kernel estimate is

$$f(x) \approx \int f(x')\, W(x - x', h)\, dx' \tag{1}$$

As a typical weight function we used the spline function of three degrees. If we identify $\rho(x')dx'$ as the differential mass dm and make summation over neighbouring particles, the discrete kernel estimate becomes

$$f(x) \approx \sum_{J=1}^{N} \frac{m^J}{\rho^J} f(x^J) W(x - x^J, h) \tag{2}$$

3. ANALYSIS

3.1 Elastic-Plastic Models for Particle Methods

It is well known that the constitutive relation for elastic-plastic field varies in regard to the strain rate. For high velocity impact Johnson and Cook [2] have successfully described for a variety of materials.

$$\sigma = (\mathrm{A} + \mathrm{B} \cdot \varepsilon^{\mathrm{n}})(1 + \mathrm{C} \ln \acute{\varepsilon})(1 - \mathrm{T}^{*\mathrm{m}}) \tag{3}$$

Where $\acute{\varepsilon}$ is the strain rate, and A, B, C, n and m are material constants; T* is the ratio $(\mathrm{T}-\mathrm{T}_{\mathrm{room}})/(\mathrm{T}_{\mathrm{melt}}-\mathrm{T}_{\mathrm{room}})$, and T is the absolute temperature. SPH structural analyses have been performed mainly for high strain rates and large strain as space-debris impact problems.

In this study as a constitutive relation for low velocity deformation of metals we use. D_{p} matrix method which is a standard algorithm for plastic deformation independent on strain rate. The incremental form of a constitutive relation is

$$\{d\sigma\} = [D]\{d\varepsilon\} \tag{4}$$

The plastic flow regime is determined by the von Mises criterion and using stress invariant the plastic constitutive model is described using D_{p} matrix.

$$\{d\sigma\} = [D_{\mathrm{p}}]\{d\varepsilon\} \tag{5}$$

The details of D_{p} matrix method is described in the standard FEM text book [3]. In a particle method the evaluation of yield condition and the calculation of strain components are executed at each particle position.

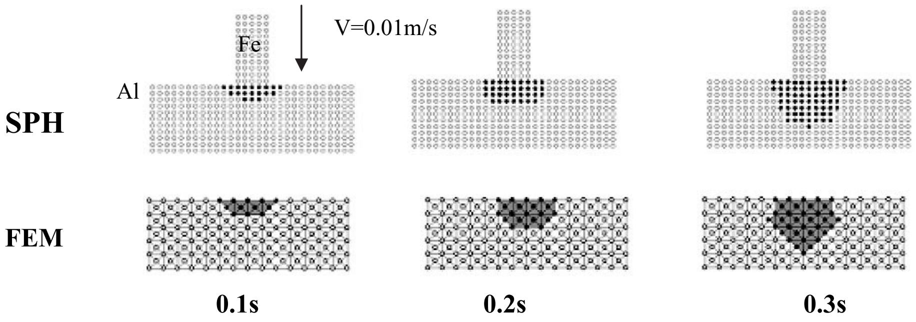


Figure 1. The growth of plastic regions.

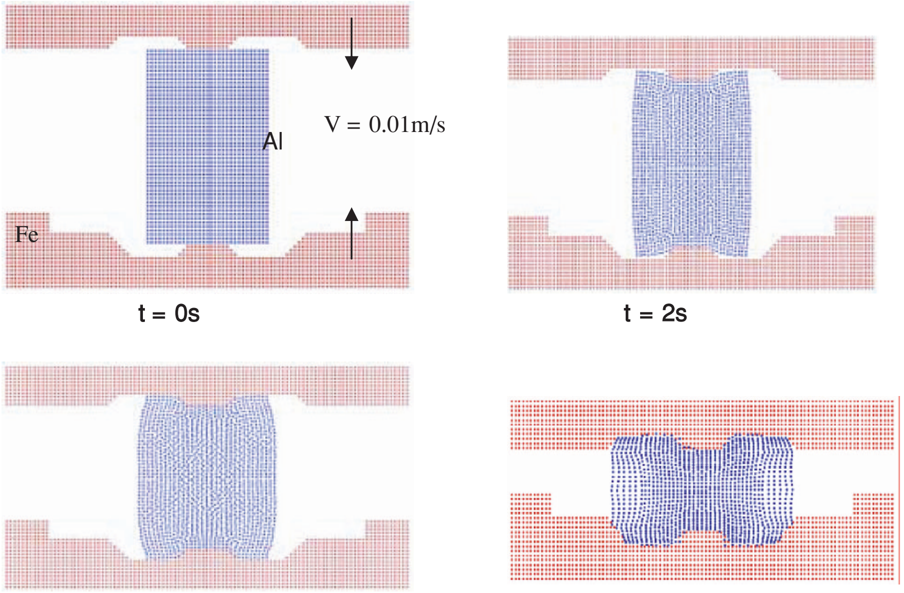
Figure 1 shows the growth of plastic regions in an impact model using FEM and SPH. An aluminum plate is loaded by an iron rod with the constant velocity 0.01 m/s. The both materials are modeled as a von Mises elastic-plastic material with a Young' of 160 GPa, 55 GPa, an initial yield strength of 5 GPa, 0.55 GPa, and a constant hardening slope of 40 GPa, 5 GPa, Poisson's ratio is 0.3. The FEM analysis is performed using triangular elements. As shown in the figures the growth of plastic regions are very similar and the stress-strain relations at any points in a materials are also in good agreement in FEM and SPH simulations.

3.2 The Elastic-Plastic Analysis Under Large Deformation

There are the growing demands for computing super large deformation of materials in mold manufacturing techniques. Up to now the analysis has been done by FEM, however the large deformation causes severe mesh distortion in elements near the boundary, thereby dramatically reducing the stable time increment during solution. Moreover rezoning and adaptive meshing technique are necessary during FEM large deformation analysis.

On the other side particle methods are meshless techniques that no longer use connectivity data, so users are free from the troublesome problems to decide how frequently remeshing should be done and what method to use to map the solution from the old mesh to the new mesh as the solution processes.

During the plastic deformations of metals under sufficiently low velocity it is noted that the volume of the deformed material has been nearly constant throughout the process. This is the typical feature of plastic flow theory and this effect should be taken into account to the analysis of plastic large deformation of metals.



t = 0s t = 2s

t = 7s t = 20s

Figure 2. The deformed shapes of a rod under compression.

In a particle simulation the constant volume condition (incompressibility condition) is obviously maintained by considering the following equation [4].

$$\frac{\partial \rho}{\partial t} = 0 \tag{6}$$

The constant of density of a material becomes the constant of density of particles in SPH simulation. When the temporal particle number density n_i^* is different from the constant value n^0, the volume constant condition is

$$n_i^* + n_i' = n^0 \tag{7}$$

where n_i' is an adjusting value. In plastic flow theory the conservative of mass is

$$\frac{1}{n^0}\frac{\partial n'}{\partial t} + \nabla \cdot \bar{u}' = 0 \tag{8}$$

Where u is the velocity of plastic flow caused by the temporally occurred pressure from the change of density n_i'. The gradient of pressure causes the temporal velocity,

$$\frac{\partial \bar{u}'}{\partial t} = -\frac{1}{\rho}\nabla P^{n+1} \tag{9}$$

Using Equations (6)–(9) the following Poisson' equation is obtained.

$$\nabla^2 P^{n+1} = -\frac{\rho}{\Delta t^2}\frac{n_i^* - n^0}{n^0} \tag{10}$$

By solving above equation we can calculate temporal pressure due to the variation of density of metals and then coordinates of particles are corrected by using the pressure gradient. Figure 2 shows the elastic-plastic large deformation behaviours of metals using above algorithm. The deformed shapes show good general agreement and thee is no large void or split between particles. All these results appear to be reasonable and well-behaved.

REFERENCES

1. J.J. Monaghan and R.A. Gingold (1983), Shock simulation by the particle method SPH. *Journal of Computational Physics*, 52, pp. 374–389.
2. G.R. Johnson and W.H. Cook (1985), Francture characteristics of three metals subjected to various strains, strain rates, temperatures and pressures. *Engineering Fracture Mechanics*, 21, pp. 31–48.
3. O.C. Zienkjewicz and R.L. Taylor (1996), *The Finite Element Method*, McGraw-Hill Book Company, New York.
4. S. Koshizuka, H. Tamako and Y. Oka (1995), A particle method for incompressible viscous flow with fluid fragmentation, *Computational Fluid Dynamics Jounal*, 4, 1, pp. 29–46.

TWO ALGORITHMS FOR SUPERCONVERGENT STRESS RECOVERY BASED ON MLS AND FINITE POINTS METHOD

S. Amirian[1], A. Kiasat[2], H. Hashemolhosseini[3], H. Dalayeli[2]
and B. Koosha[1]
[1]*Department of Civil Engineering, Isfahan University of Technology, Isfahan 84156, Iran*
[2]*Department of Mechanical Engineering, Isfahan University of Technology, Isfahan 84156, Iran*
[3]*Department of Mining Engineering, Isfahan University of Technology, Isfahan 84156, Iran*

Abstract Two methods of stress recovery have been suggested and investigated in this paper. The first one is introduced by using moving least square method (MLS) and superconvergent points. The second method is based on the satisfaction of equilibrium equations at some nodes for which the recovery is applied. Simultaneous solution of these equations increases computational time. So, the second method is more expensive than the first one. A numerical example is used to compare the stresses recovered by these two methods with corresponding FEM, well-known SPR method and analytical solutions. The effect of various orders of basis functions and the values of weight functions on stress recovery by the proposed methods is also investigated. The present research indicates that the two methods, and especially the first one, represent acceptable accuracy over the domain and even on the boundaries, in comparison with SPR method, and also good convergency is achieved.

Keywords: stress recovery, moving least square (MLS), SPR method and FEM.

1. INTRODUCTION

Due to the nature of conventional displacement type FEM, the gradient of obtained field becomes discontinue crossing the borders of elements. This discontinuity reduces as mesh system of problem becomes finer and the approximate solution approaches to exact one. However for better interpretation of obtained results it is often necessary to engage certain procedures for smoothing

G. R. Liu et al. (eds.), Computational Methods, 1521–1525.

these approximate fields gradient. It has been shown [1] that in certain points of elements the precision of field gradient is distinctively superior to the other points. Zhu and Zienkiewicz [2, 3] used this phenomenon for obtaining a smooth field gradient in certain subdomain called as patch. Each patch contains a certain number of elements. The mentioned algorithm which is called Superconvergent Patch Recovery method (SPR) uses a conventional least square method to relate an approximate gradient field to each patch. This method not only yields a smooth field in each patch but also effectively improves the precision of predicted field. Unfortunately the SPR method is much less effective at the borders of problems. Wiberg et al. [4] try to improve the performance of SPR method by incorporating the equilibrium and boundary conditions. Although they engage a sophisticated and expensive procedure which needs numerical integration over each patch, their results show no significant improvement to the standard SPR method. Tabbara et al. [5] recovered the stress field by moving least square (MLS) technique. They obtained local approximation of displacement field by MLS technique from nodal values. By differentiating from this approached field and equilibrium equation they are able to obtain stress field. This method is limited to the problems with linear constitutive equations. By establishing the equilibrium between external and internal forces, Boroomand and Zienkiewicz [6] recovered the stress field. Their method seems to be quite complicated but highly efficient. Hashemolhosseini et al. [7] introduced an algorithm based on finite point method for recovery the hydrostatic pressure. The present work can be somehow considered as an extension of this algorithm. Hereby two different stress recovery methods based on MLS and finite point technique are presented.

2. FORMULATION

In the first purposed method the recovered stresses in each point (often in the nodes) is obtained by interpolating from the values of stresses of superconvergent points in its vicinity. The interpolation is done by using the Moving Least Square Method. Here the choice of influence domain, basis function and weight function can influence the results. In the second method the stresses at each point is approximated by MLS technique from superconvergent and the recovered points as (Formula (1)):

$$\hat{\sigma}_{ij} = \sum_{I=1}^{M} \mathrm{N}_I^{(1)}(\mathbf{x})\, \bar{\sigma}_{ij_I}^{(1)} + \sum_{J=1}^{N} \mathrm{N}_J^{(2)}(\mathbf{x})\, \bar{\sigma}_{ij_J}^{(2)} = \sum_{K=1}^{N+M} \mathrm{N}_K(\mathbf{x})\, \bar{\sigma}_{ij_K} \qquad (1)$$

Here $\hat{\sigma}_{ij}$ are the recovered stresses at desired point, $\bar{\sigma}_{ij_I}^{(1)}$ are the known stresses at the superconvergent points, $\bar{\sigma}_{ij_I}^{(2)}$ are the unknown stresses at recoverd points and $N(\mathbf{x})$ are corresponding shape functions. The residue of equilibrium at each point can be obtained as (Formula (2)):

$$\sum_{K=1}^{N+M} \frac{\partial N_K(\mathbf{x})}{\partial \mathbf{x}_j}\bar{\sigma}_{ij_K} + f_i = R_i^{eq} \tag{2}$$

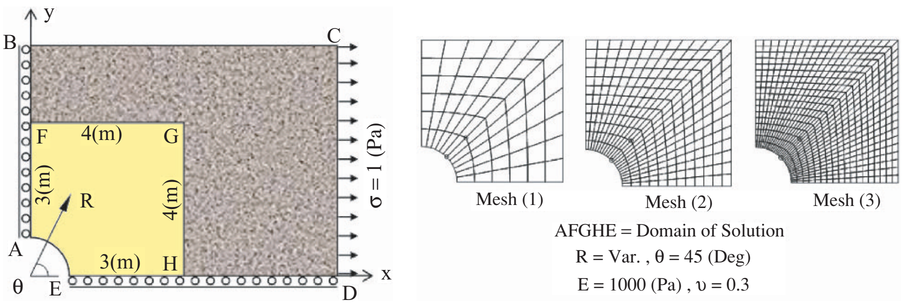
While at the boundary points the residue of boundary condition equation becomes as (Formula (3)):

$$\sigma_{ij}n_j - T_i = 0 \text{ OR } \sum N_K(\mathbf{x})\bar{\sigma}_{ij_K} n_j - T_i = R_i^{bo} \tag{3}$$

Here T_i and n_j are traction vector and normal vector to boundary surface. At last by minimizing the total square of residues, the recovered stresses at the desired points are computed.

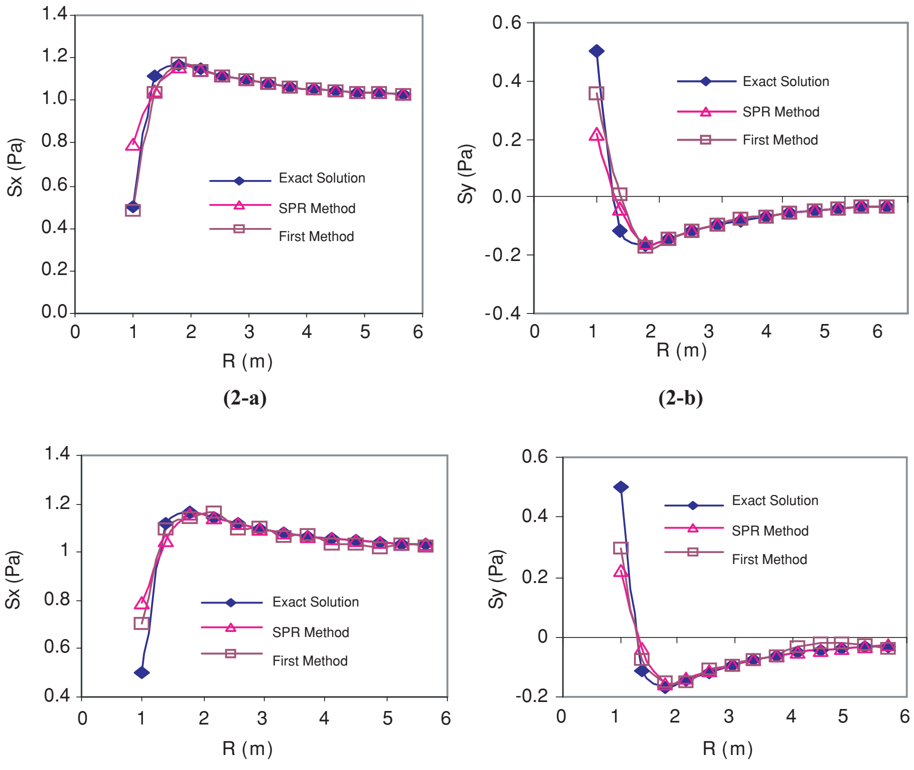
3. NUMERICAL EXAMPLE AND RESULTS

The well-known elastic problem shown in Figure 1 is selected to compare the methods. This problem is often considered as a benchmark for stress recovery algorithms. The problem is solved by three systems of meshes; consist of bilinear Q4 elements with one point of integration for each element. The results show (Figure 2) that both introduced methods are more efficient than SPR approach. Especially they yield much more precise results near the borders of problem. The second method is more complicated than the first one, but not as efficient as it. The first method is a simple and cheap algorithm which at least can be used with SPR method for recovery of stresses at the points which are near the boundaries.



AFGHE = Domain of Solution
R = Var. , θ = 45 (Deg)
E = 1000 (Pa) , υ = 0.3

Figure 1. Infinite plate with a hole.



(2-a) (2-b)

(2-c) (2-d)

Figure 2. Normal stresses in x and y directions vs. radius for different methods.

REFERENCES

1. J. Barlow (1981), Optimal stress location in finite element method. *International Journal for Numerical Methods in Engineering*, 10, pp. 243–251.
2. J.Z. Zhu and O.C. Zienkiewicz (1990), Superconvergence recovery technique and a posteriori error estimators. *International Journal for Numerical Methods in Engineering*, 30, pp. 1321–1339.
3. O.C. Zienkiewicz and Z. Zhu (1992), The superconvergent patch recovery and a posteriori error estimates. Part I: the recovery technique. *International Journal for Numerical Methods in Engineering*, 33, pp. 1331–1364.
4. N.E. Wiberg, F. Abdulwahab and S. Ziukas (1994), Enhanced superconvergent patch recovery incorporation equilibrium and boundary conditions. *International Journal for Numerical Methods in Engineering*, 37, pp. 3417–3440.
5. M. Tabbara, T. Blaker and T. Belytschko (1994), Finite element derivative recovery by moving least square interpolants. *Computer Methods in Applied Mechanics and Engineering*, 117, pp. 211–223.

6. B. Boroomand and O.C. Zienkiewicz (1997), An improved REP recovery and effectivity robustness test. *International Journal for Numerical Methods in Engineering*, 40, pp. 3247–3277.
7. H. Hashemolhosseini, H. Dalayeli and M. Farzin (2002), Correction of hydrostatic pressure obtained by the finite element flow formulation using moving least squares method. *Journal of Material Processing Technology*, 125–126, pp. 588–593.

GEOMETRICALLY NONLINEAR ANALYSIS USING MESHFREE RPIM

K.Y. Dai[1], G.R. Liu[1] and X. Han[2]
[1]*Department of Mechanical Engineering, National University of Singapore, 9 Engineering Drive 1, Singapore 117576*
[2]*The Key Laboratory of Advanced Technology for Vehicle Body Design & Manufacture of Ministry of Education, College of Mechanical and Automotive Engineering, Hunan University, Changsha 410082, P. R. China*

Abstract Geometrically nonlinear analysis for structures with finite deformation is presented based on a meshfree radial point interpolation method (RPIM). The method uses a set of distributed nodes to discretize problem domain that may reduce the possibility of mesh distortion or tangling. The formulations are given in material description. The method is implemented using a full Newton–Raphson iteration procedure. Numerical problem is investigated for a cantilever beam made of compressible hyperelstic neo-Hookean materials to demonstrate the effectiveness of this method.

Keywords: meshfee method, nonlinear analysis, point interpolation method, finite deformation, geometrical nonlinearity.

1. INTRODUCTION

Nonlinear analysis becomes now an essential component to investigate the behaviours of structures under abnormal loading conditions, especially in the design of structures of peculiar types or of great importance. In addition, in the presence of pre-stress geometric nonlinearities are of the same order of magnitude as linear elastic effects in structures, which implies that in most cases geometric nonlinearities should be considered, such as the problems of buckling, cable nets, fabric structures, etc. [1]. In the field of manufacturing, nonlinear simulation is also an important process in prototype testing, sheet metal forming, extrusion, forging and casting.

G. R. Liu et al. (eds.), Computational Methods, 1527–1531.

In recent years, the element-free or meshfree methods have been developed to deal with the large deformation or crack propagation problems. They have achieved remarkable progress in computational mechanics and related fields. The main purpose of this paper is to further the RPIM [2–4] to the geometrical nonlinear problems in solid mechanics.

2. RADIAL POINT INTERPOLATION METHOD (RPIM)

Consider a field function $u(\mathbf{x})$ defined by a set of arbitrary distributed nodes $\mathbf{x}_i (i = 1, 2, \ldots, N)$ in a domain Ω bounded by Γ. It is assumed that only the neighboring n nodes in the support domain of a considered point $\mathbf{x}_Q$ have effects on $u(\boldsymbol{x}_Q)$. The radial PIM with polynomial basis function interpolates the field variable $u(\mathbf{x})$ in form [5]

$$u^h(\mathbf{x}, \mathbf{x}_Q) = \sum_{i=1}^{n} r_i(\mathbf{x}) a_i + \sum_{j=1}^{m} p_j(\mathbf{x}) b_j = \mathbf{r}^T(\mathbf{x})\mathbf{a} + \mathbf{p}^T(\mathbf{x})\mathbf{b} \tag{1}$$

where a_i and b_j are the unknown coefficients for the radial basis $r_i(\mathbf{x})$ and the polynomial basis $p_j(\boldsymbol{x})$, respectively. Among many others, multi-quadric (MQ) function is used here. The coefficients a_i and b_j are determined by enforcing the Equation (1) to pass through n data nodes in the support domain. Therefore the following equations can be obtained.

$$\mathbf{u}^e = \mathbf{R}_Q \mathbf{a} + \mathbf{P}_m \mathbf{b} \tag{2}$$

To guarantee unique solutions, the following constraints are imposed:

$$\sum_{i=1}^{n} p_j(\mathbf{x}_i) a_i = 0, \quad j = 1, 2, \cdots, m \tag{3}$$

Combining Equations (2) and (3) and solving these equations yield

$$\mathbf{a} = \mathbf{S}_a \mathbf{u}^e, \quad \mathbf{b} = \mathbf{S}_b \mathbf{u}^e \tag{4}$$

Substituting vectors **a** and **b** into Equation (1), the interpolation function $\boldsymbol{\varphi}(\mathbf{x})$can be obtained

$$u^h(\mathbf{x}, \mathbf{x}_Q) = [\mathbf{r}^T \mathbf{S}_a + \mathbf{p}^T \mathbf{S}_b] \mathbf{u}^e = \boldsymbol{\varphi}(\mathbf{x}) \mathbf{u}^e \tag{5}$$

Note that the shape functions obtained through the above procedure possess Delta function property and unity partition property. If at least linear polynomial terms are included in Equation (1), they can reproduce a general linear polynomial exactly.

3. VARIATIONAL FORM AND DISCRETE EQUATIONS

Using the qualities related to the current configuration, the standard equilibrium equation for a solid is given as (see e.g. Belytschko *et al.*, [6])

$$\frac{\partial \sigma_{ji}}{\partial x_j} + \rho b_i = \rho \dot{v}_i \quad in\ \Omega \tag{6}$$

where σ_{ji} is the Cauchy stress component; b_i the body force component per unit mass and ρ is the mass density in the current configuration. Note that $\rho = \rho_0 / J$ and ρ_0 is the mass density in the initial undeformed configuration.

For the current configuration the traction and displacement boundary conditions may be expressed by requiring that

$$\sigma_{ji} n_j = \bar{t}_i \quad on\ \Gamma_t,\ u_i = \bar{u}_i \quad on\ \Gamma_u \tag{7}$$

where n_j is direction cosine of a unit outward normal on the deformed surface. The variational weak form of Equation (6) can be expressed as

$$\int_v \delta D_{ij} \sigma_{ji}\, dv + \int_v \delta v_i \rho \dot{v}_i\, dv = \int_v \delta v_i \rho b_i\, dv + \int_\Gamma \delta v_i \bar{t}_i\, d\Gamma \tag{8}$$

It can be transformed to the expression with respect to the initial or reference configuration

$$\int_V \delta \dot{E}_{ij} S_{ji}\, dV + \int_V \delta v_i \rho_0 \ddot{u}_i\, dV = \int_V \delta v_i \rho_0 b_i\, dV + \int_{\Gamma_0} \delta v_i \bar{T}_i\, d\Gamma_0 \tag{9}$$

After discretization of the linearized equilibrium equation using RPIM shape functions, the discrete form becomes

$$\mathbf{Ku} = \mathbf{R} - \mathbf{M}\dot{\mathbf{v}} \tag{10}$$

where the tangent stiffness matrix consists of material and geometric parts. The residual force vector $\mathbf{R}$ is defined as the deviation of the internal force $\mathbf{L}^S$ from the external force $\mathbf{L}^E$.

In this paper a simple case of an isotropic hyperelastic material is considered, which is termed as the compressible neo-Hookean material. The stored energy function ψ for such a material is defined as

$$\psi = \frac{\mu_0}{2}(I_1 - 3) - \mu_0 \ln J + \frac{\lambda_0}{2}(\ln J)^2 \tag{11}$$

Here the quantities λ_0 and μ_0 are the Lame constants of the linearized theory. The second Piola-Kirchhoff stress tensor $\mathbf{S}$ can be derived as

$$\mathbf{S} = \lambda_0 \ln J \mathbf{C}^{-1} + \mu_0(\mathbf{I} - \mathbf{C}^{-1}) \tag{12}$$

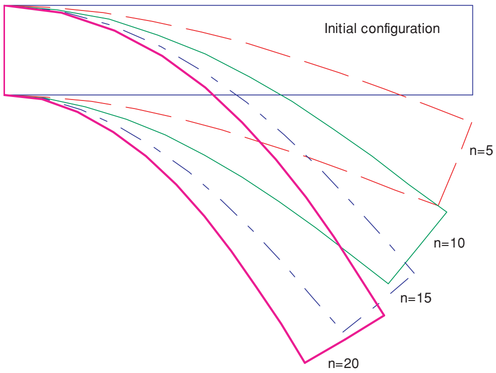


Figure 1. Bending of a cantilever beam according to different steps of load increment.

The elasticity tensors (tangent moduli) are given as follows in component form

$$C^{S}_{ijkl} = \lambda_0 C^{-1}_{ij} C^{-1}_{kl} + (\mu_0 - \lambda_0 \ln J)\,(C^{-1}_{ik} C^{-1}_{jl} + C^{-1}_{il} C^{-1}_{kj}) \tag{13}$$

4. NUMERICAL EXAMPLE

A cantilever beam with large deformation is analysed that subjected to a tip load increment of $\Delta F = 16.0\ N$. The size of the beam is (10 cm × 2 cm) and initially discretized using (11 × 3) regularly distributed nodes. The analysis is carried out using twenty load incremental steps ($n = 20$). The simulation converges very fast and in each load increment, the iteration is performed less than 4 times. Figure 1 illustrates the different stages of deformation for the beam using RPIM. The link between the tip deflections versus the load steps is plotted in Figure 2. It can be seen that, the nonlinear analysis makes the beam stiffer than the linear solutions as the load increases. Some other examples have also been successfully analysed and will be presented at the conference.

5. CONCLUSION

In this paper, the point interpolation method (PIM) is presented for the analysis of geometrically nonlinear problems with finite deformation. With radial basis functions employed as interpolators, the problem domain can be discretized by a set of regular/irregular nodes and no element connectivity is considered. Load incremental procedure is used to ensure stable and accurate solutions along with the full Newton–Raphson algorithm. The equilibrium equation and its corresponding linearizations as well as the applied constitutive

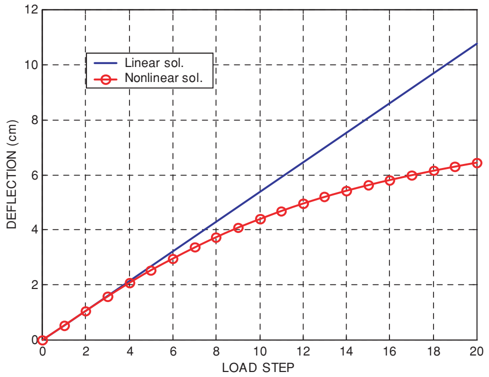


Figure 2. Tip-deflection/load relationships for a cantilever beam.

modes are established in terms of a material description. Comparing with formulations in spatial description, it is found that the present ones show higher convergence speed and are more suitable for geometrical analysis of solids with finite deformation, which is therefore recommended when using RPIM.

ACKNOWLEDGEMENT

This work is supported by the National Science Foundation of China under the grant number 10372031.

REFERENCES

1. R. Levy and R.W. Spillers (1995), *Analysis of Geometrically Nonlinear Structures*. Chapman & Hall, New York.
2. G.R. Liu and Y.T. Gu (2001), A point interpolation method for two-dimensional solids. *International Journal for Numerical Methods in Engineering*, 50, pp. 937–951.
3. J.G. Wang and G. R. Liu (2002), A point interpolation meshless method based on radial basis functions. *International Journal for Numerical Methods in Engineering*, 54, pp. 1623–1648.
4. G.R. Liu (2002), *Mesh Free Methods: Moving Beyond the Finite Element Method*. CRC press, Boca Raton, FL.
5. M.J.D. Powell (1992), The theory of radial basis function approximation in 1990. *Advances in Numerical Analysis*, F.W. Light (ed.), Oxford: Clarendon Press, pp. 303–322.
6. T. Belytschko, W.K. Liu and B. Moran (2000), *Nonlinear Finite Elements for Continua and Structures*. John Wiley & Sons, Chichester.

GALERKIN MESHLESS METHODS BASED ON PARTITION OF UNITY QUADRATURE

Q.H. Zeng and D.T. Lu
Department of Mechanics and Mechanical Engineering, University of Science and Technology of China, Hefei, China 230027

Abstract Partition of unity quadrature (PUQ) for Galerkin meshless methods is presented and Galerkin meshless methods based on PUQ are studied in detail. The mathematic foundations of PUQ, finite covering and partition of unity, are described and the validity of PUQ is proved theoretically. Galerkin meshless methods based on PUQ do not require any mesh information and are 'truly' meshless methods. Numerical examples show the excellent agreement with conventional quadrature methods and exact solution, which exemplifies the accuracy and robustness of Galerkin meshless methods based on PUQ. The relation between accuracy of integral and cover size is also analysed.

Keywords: Galerkin meshless methods, partition of unity quadrature, finite covering, partition of unity.

1. INTRODUCTION

Meshless methods are the computational methods which have been developed rapidly in recent years. According to the ways transforming partial differential equations into algebraic equations, Meshless methods can be classified collectively as Galerkin meshless methods [1], collocation meshless methods [2] and Petrov–Galerkin meshless methods [3]. Due to the high accuracy and stability, Galerkin meshless methods are applied broadly. But it's unavoidable to carry out the numerical integration over the whole physical domain in Galerkin weak form.

To evaluate the integral in Galerkin meshless methods, Belytschko et al. [1] decomposed the physical domain into regular meshes, called cell structure, and Gauss quadrature is employed on each cell structure. This quadrature method is called cell quadrature (CQ) method which is used in most of Galerkin meshless

G. R. Liu et al. (eds.), Computational Methods, 1533–1539.

methods at present. There are two drawbacks in this method. One is that the introduction of cell wastes some of the advantages of meshless methods, the other is that the cells are supposed to be regular without considering the spatial arrangement of nodes. As a result, significant error may be gotten when the integrand is complex.

Beissel et al. [4] and Chen et al. [5] use another quadrature method, nodal integration. This method can be expressed as follows

$$\int_{\Omega} f(x) d\Omega = \sum_{i} f(x_i) \Delta\Omega_i \tag{1}$$

In nodal integration, the entire domain is decomposed into a large number of small subdomains $\Delta\Omega_i$. The integrand $f(\boldsymbol{x})$ in $\Delta\Omega_i$ is supposed to be $f(\boldsymbol{x}_i)$, $\boldsymbol{x}_i$ is a point included in $\Delta\Omega_i$. Nodal integration is easy to perform and the computational effort is small, but it's difficult to assign $\Delta\Omega_i$ to each of nodes, and the solution presents spatial oscillations.

Partition of unity quadrature technique is presented and Galerkin meshless methods based on PUQ are studied in this paper. PUQ seems attractive in several ways. First, the decomposition of quadrature field and mesh information are always not required in this method, which makes Galerkin meshless methods be 'truly' meshless methods. Second, the stiffness matrix is symmetric and positive definite. Third, the influence of the spatial arrangement of nodes are considered which can reduce the integral error according to Dolbow's viewpoint [6].

A brief outline of the paper is as follows. In Section 2, partition of unity quadrature method is presented. Numerical implementation of Galerkin meshless methods with PUQ is demonstrated in Section 3 and numerical examples are given in Section 4. The paper ends with conclusions and discussions in Section 5.

2. PARTITION OF UNITY QUADRATURE

Definition 1 (Finite Convering). Assume Ω is an open smooth domain in $R^d (d = 1, 2, 3)$, $\bar{\Omega}$ is the closure of Ω. $\boldsymbol{Q}_N$ will denote an arbitrarily chosen set of N nodes from Ω

$$\boldsymbol{Q}_N = \left\{ \boldsymbol{x}_1\ \boldsymbol{x}_2 \cdots \boldsymbol{x}_N \right\}, \boldsymbol{x}_k \in \Omega$$

The set $\boldsymbol{Q}_N$ is then used to define a finite open covering $\boldsymbol{O}_N = \{\Omega_k\}_{k=1}^{N}$ of Ω. $\boldsymbol{O}_N$ is composed of N circular (or elliptic, rectangular) covers centred at the nodes $\boldsymbol{x}_k$, $k = 1, \ldots, N$, and the following conditions are satisfied

$$\Omega_I = \{\boldsymbol{x} \in \Omega : ||\boldsymbol{x} - \boldsymbol{x}_k||_{R^d} < r_k\} \tag{2a}$$

$$\bigcup_{k=1}^{N} \Omega_k \supset \bar{\Omega} \tag{2b}$$

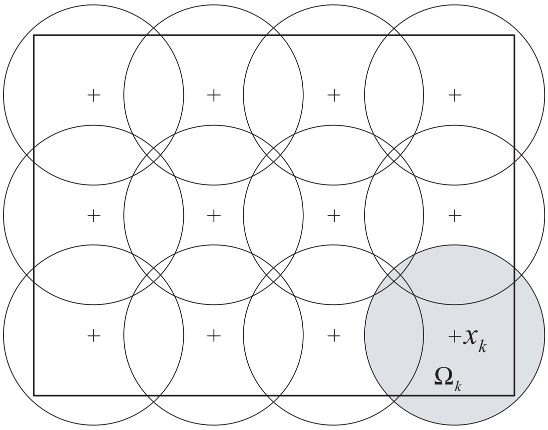


Figure 1. Finite covering.

then $\boldsymbol{O}_N$ is called a finite covering of Ω.

A finite covering $\{\Omega_k\}_{k=1}^N$ of Ω is shown in Figure 1, where $\Omega \subset R^2$, the covers are set to be circular and the centres are marked by sign "+".

Definition 2 (Partition of Unity). Assume $\boldsymbol{O}_N = \{\Omega_k\}_{k=1}^N$ is a finite covering of $\Omega \subset R^d$, a class of functions $\Phi_N = \{\phi_k\}_{k=1}^N$ is called a partition of unity with respect to the finite covering $\boldsymbol{O}_N$ if it possesses the following properties

$$\phi_k(\boldsymbol{x}) \in C_0^\infty(R^d), \quad 0 \le \phi_k(\boldsymbol{x}) \le 1, \quad \boldsymbol{x} \in R^d, k = 1, 2, \ldots, N \tag{3a}$$

$$\sum_{k=1}^N \phi_k(\boldsymbol{x}) = 1, \quad \forall \boldsymbol{x} \in \Omega \tag{3b}$$

$$\text{supp } \phi_k(x) \subset \Omega_k, \quad k = 1, 2, \ldots, N \tag{3c}$$

where supp $\phi_k(\boldsymbol{x})$ denotes the closure of set $\{\boldsymbol{x} \in \Omega : \phi_k(\boldsymbol{x}) \ne 0\}$, called support of $\phi_k(x)$. Condition (3a) can be loosen to $\phi_k(x) \in C^l(\Omega_k), \quad l \in N_0$.

Theorem 1. Assume $\boldsymbol{O}_N = \{\Omega_k\}_{k=1}^N$ is a finite covering of finite open domain Ω, $\Phi_N = \{\phi_k\}_{k=1}^N$ is a partition of unity with respect to $\boldsymbol{O}_N$, $f(x)$ is a integrable function defined on Ω, then

$$\int_\Omega f(\boldsymbol{x})\, d\Omega = \sum_{k=1}^n \int_{\Omega\cap\Omega_k} \phi_k(\boldsymbol{x}) f(\boldsymbol{x})\, d\Omega \tag{4}$$

Proof. Since $\sum\limits_{k=1}^N \phi_k(\boldsymbol{x}) = 1$ for $\forall \boldsymbol{x} \in \Omega$, we have

$$\int_\Omega f(\boldsymbol{x})\, d\Omega = \int_\Omega \sum_{k=1}^N \phi_k(\boldsymbol{x}) f(\boldsymbol{x})\, d\Omega = \sum_{k=1}^N \int_\Omega \phi_k(\boldsymbol{x}) f(\boldsymbol{x})\, d\Omega \tag{5}$$

From Equation (3c), we have

$$\int_{\Omega} \phi_k(\boldsymbol{x}) f(\boldsymbol{x})\, d\Omega = \int_{\Omega\cap\Omega_k} \phi_k(\boldsymbol{x}) f(\boldsymbol{x})\, d\Omega \tag{6}$$

By Substituting Equation (6) into (5), we will obtain Equation (4). Because of formula (3b), we called this integral technique partition of unity quadrature (PUQ).

3. NUMERICAL IMPLEMENTATION

Consider 2D Poisson equation in a domain Ω with boundary $\partial\Omega$

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} &= f(x, y) \quad \text{in } \Omega \\ u &= \bar{u} \quad \text{on } \Gamma_u \\ \frac{\partial u}{\partial n} &= -\bar{q} \quad \text{on } \Gamma_q \end{aligned} \tag{7}$$

where the superposed bar denotes prescribed boundary value, Γ_u and Γ_q are the essential boundary and natural boundary, respectively. $\Gamma_u \cup \Gamma_q = \partial\Omega$, $\Gamma_u \cap \Gamma_q = \phi$, $\partial\Omega$ is the boundary of Ω, and n is the unit outward normal to $\partial\Omega$.

Using modified variation principle, the equation and boundary conditions mentioned above are equivalent to minimize the following functional

$$\begin{aligned} \Pi = \int_{\Omega} \left\{ \frac{1}{2} \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right] + uf \right\} \mathrm{d}\Omega \\ - \int_{\Gamma_u} \left[\frac{\partial u}{\partial n} (u - \bar{u}) \right] \mathrm{d}\Gamma + \int_{\Gamma_q} u\bar{q}\, \mathrm{d}\Gamma \end{aligned} \tag{8}$$

The stationarity of Π, $\delta\Pi = 0$ can lead to following system of linear equations

$$[\boldsymbol{K}]\{\boldsymbol{u}\} = \{\boldsymbol{F}\} \tag{9}$$

where

$$\begin{aligned} \boldsymbol{K}_{ij} = -\int_{\Omega} \left(\frac{\partial N_i}{\partial x} \frac{\partial N_j}{\partial x} + \frac{\partial N_i}{\partial y} \frac{\partial N_j}{\partial y} \right) \mathrm{d}\Omega + \int_{\Gamma_u} \left[\left(\frac{\partial N_i}{\partial x} n_x + \frac{\partial N_i}{\partial y} n_y \right) N_j \right. \\ \left. + \left(\frac{\partial N_j}{\partial x} n_x + \frac{\partial N_j}{\partial y} n_y \right) N_i \right] \mathrm{d}\Gamma \end{aligned} \tag{10a}$$

$$\boldsymbol{F}_i = \int_{\Omega} N_i \mathrm{d}\Omega + \int_{\Gamma_u} \left[\left(\frac{\partial N_i}{\partial x} n_x + \frac{\partial N_i}{\partial y} n_y \right) \bar{u} \right] \mathrm{d}\Gamma + \int_{\Gamma_q} N_i \bar{q}\, \mathrm{d}\Gamma \tag{10b}$$

n_x, n_y are the Cartesian components of n.

It is noted that MLS shape function $N(x)$ [1, 7] and partition of unity (PU) function $\Phi(x)$ can be constructed by MLS, but the PU function is not necessary to keep coincident with MLS shape function [8]. For simpleness, Shepard function [7] is usually acted as the PU function.

4. NUMERICAL EXAMPLES

We will consider Poisson equation:

$$\begin{aligned} \nabla^2 u(x, y) &= -2(x + y - x^2 - y^2) \quad \text{in}\, \Omega : x \in [0, 1], y \in [0, 1] \\ u(x, y) &= 0 \quad \text{on}\, \partial\Omega \end{aligned} \tag{11}$$

the exact solution is given by

$$u(x, y) = (x - x^2)(y - y^2) \tag{12}$$

This problem is analysed for 11 × 11 nodes as shown in Figure 2 and the numerical results are obtained by using EFG with PUQ and CQ, respectively (Figure 3). The domain surrounded by nodes is taken as the subfield of integration in CQ, which is a square and the length of side is denoted by L. In each cell 4 × 4 Gauss points are used to evaluate the stiffness of EFG. The square covers are used in PUQ with Shepard PU function, the length of side of the cover is $1.6L$ and 4 × 4 Gauss points are distributed in each cover. It can been seen

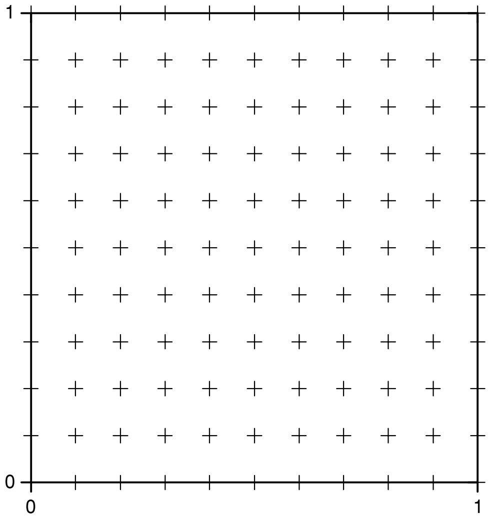


Figure 2. Nodal arrangement.

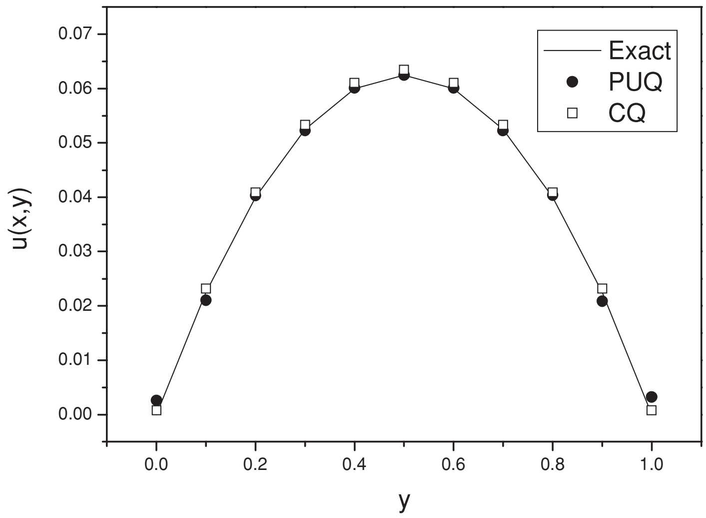


Figure 3. Results comparison along $x = 0.5$.

that the numerical results of PUQ are in good agreement with CQ and exact solution.

The accuracy of PUQ is closely related to the cover size. For this study, we define L_2 error as follows:

$$error = \sum_{i=1}^{N} \left[u^{NUM}(x_i, y_i) - u^{EXC}(x_i, y_i)\right]^2 \tag{13}$$

where $u^{NUM}(x_i, y_i)$ and $u^{EXC}(x_i, y_i)$ are numerical solution and exact solution at (x_i, y_i), respectively. The L_2 error is shown in Table 1 while the length of side of square covers ranges from $1.0L$ to $2.4L$.

The finite covering can 'just' cover the computational domain if the length of side of covers is $1.0L$, which is equivalent to CQ. It can be seen from Table 1 that the CQ results are not as accurate as PUQ in general. A significant error appears when the length of side increases to be $1.2L$ due to the absence of Gauss points in the overlapping of covers, which makes the integration results larger. Remarkably accurate results are achieved when the length of side is $1.4L$–$2.2L$.

Table 1. The size of quadrate covers and L_2 error ($\times 10^{-3}$).

Length of side	$1.0L$	$1.2L$	$1.4L$	$1.6L$	$1.8L$	$2.0L$	$2.2L$	$2.4L$
L_2 error	15.50	117.20	0.39	0.26	0.09	1.92×10^{-6}	0.56	2.30

5. CONCLUSIONS

A numerical integral technique for Galerkin meshless methods, partition of quadrature (PUQ) is presented and Galerkin meshless methods based on PUQ are studied in detail. The relation between accuracy and covers size is analysed, and it is observed that the numerical results is better than CQ when the cover size is $1.4L$–$2.2L$. PUQ is also appropriate to Petrov–Galerkin meshless methods, and it is easy to extend to linear or nonlinear elastic problems.

REFERENCES

1. T. Belytschko, Y.Y. Lu and L. Gu (1994), Element-free Galerkin methods. *International Journal for Numerical Methods in Engineering*, 37, pp. 229–256.
2. X. Zhang, K.Z. Song, M.W. Lu and X. Liu (2000), Meshless methods based on collocation with radial basis functions. *Computational Mechanics*, 26, pp. 333–343.
3. S.N. Atluri and T. Zhu (2000), New concepts in meshless methods. *International Journal for Numerical Methods in Engineering*, 47, pp. 537–556.
4. S. Beissel and T. Belytschko (1996), Nodal integration of the element-free Galerkin method. *Computer Methods in Applied Mechanics and Engineering*, 139, pp. 49–74.
5. J.S. Chen, C.T. Wu, S. Yoon and Y. You (2001), A stabilized conforming nodal integration for Galerkin mesh-free methods. *International Journal for Numerical Methods in Engineering*, 50, pp. 435–466.
6. J. Dolbow and T. Belytschko (1999), Numerical integration of the Galerkin weak form in meshfree methods. *Computational Mechanics*, 23, pp. 219–230.
7. T. Belytschko, Y. Krongauz, D. Organ, M. Fleming and P. Krysl (1996), Meshless methods: an overview and recent developments. *Computer Methods in Applied Mechanics and Engineering*, 139, pp. 3–47.
8. D. Marc and N.D. Hung (2002), A truly meshless Galerkin method based on a moving least squares quadrature. *Communications in Numerical Methods in Engineering*, 18, pp. 1–9.

RADIAL POINT INTERPOLATION COLLOCATION METHOD (RPICM) USING UPWIND BIASED LOCAL SUPPORT SCHEME FOR SOLVING CONVECTION-DOMINATED EQUATIONS

Xin Liu[1], G. R. Liu[2] and Kang Tai[3]
[1]*Department of Mechanics, Zhejiang University, Hangzhou, P. R. China, 310027*
[2]*Department of Mechanical Engineering National University of Singapore, Singapore, 117576*
[3]*Division of Engineering Mechanics, School of Mechanical & Productional Engineering, Nanyang Technological University, Singapore*

Abstract This paper discusses a scheme for choosing local upwind biased support domains in the RPICM for solving convection-dominated problems. A numerical example of a 2D steady state convection-domainted problem is presented. It is demonstrated that significant improvement on the accuracy can be obtained after using this scheme for convection-dominated problems.

Keywords: meshfree, convection-dominated equation, collocation.

1. INTRODUCTION

Point interpolation method (PIM) and radial PIM are based on weak-forms and local support formulation and were proposed by Liu [1], and it has been further studied for fluid problems [2]. So far, most applications of PIM are based on Galerkin or Petrov-Galerkin weak forms, which needs numerical integrations globally or locally. A collocation form of RPIM or RPICM has also been studied and applied to solve different kind of PDEs [3, 4]. This study applies RPICM to the solution of convection-dominated equations.

The convection-diffusion equation appears in a variety of practical applications involving the modeling of transport phenomenon. The convection-dominated equation, where the magnitudes of the ratio of the diffusion

G. R. Liu et al. (eds.), Computational Methods, 1541–1546.

coefficient to convection coefficient is very small, is very difficult to be solved numerically based on standard formulation. In recent years, a great deal of effort has been devoted to the development of the numerical approximation of convection-dominated problems, as well-summarised by Zienkiewicz and Taylor [5]. In finite element and finite difference method (FDM), a common upwind scheme has been effectively adopted to obtain high accurate solution for such problems. Similar to the upwind difference scheme used in FDM, the local upwind biased support domain scheme is presented and implemented in the RPICM to stabilize the solution in this study. It can be shown that the upwind biased support domain can improve the accuracy and stability for convection-dominated problems because it can fully capture the information from the upstream. Due to the freedom in selecting the support domain in meshfree methods, this scheme will be easily incorporated in RPICM for solving convection-dominated problems.

In this paper, radial point interpolation formulation is first introduced. Then RPICM is applied to solve 2D steady state convection-dominated equation with special emphasis on using local upwind biased support domain scheme. Some concluding remarks are given finally.

2. RADIAL BASIS POINT INTERPOLATION

The approximation of a function $u(\mathbf{x})$ may be written as a linear combination of n radial basis functions, viz.,

$$u(\mathbf{x}) \cong \tilde{u}(\mathbf{x}) = \sum_{i=1}^{n} a_i \phi(||\mathbf{r} - \mathbf{r}_i||) \tag{1}$$

where n is the number of points in a local support domain near $\mathbf{x}$, a_i are coefficients to be determined and the radial basis function ϕ can be the well-known Multiquadrics (MQ), or Gaussian basis function, or thin plate spline (TPS) function.

Gaussian basis function is used in this paper:

$$\phi(||\mathbf{r} - \mathbf{r}_i||) = \exp(-||\mathbf{r} - \mathbf{r}_i||^2) \tag{2}$$

Where $\mathbf{r}$ is the distance between two nodes. In 2D problems, we have

$$||\mathbf{r} - \mathbf{r}_i|| = \sqrt{(x - x_i)^2 + (y - y_i)^2} \tag{3}$$

The coefficients a_i in Equation (1) can be determined by enforcing Equation (1) to be satisfied at all n nodes within the local support domain.

The interpolations of a function at the k^{th} point can have the form of

$$\begin{aligned}\tilde{u}(\mathbf{x}_k) = a_1\phi(||\mathbf{r}_k - \mathbf{r}_1||) + a_2\phi(||\mathbf{r}_k - \mathbf{r}_2||) + \cdots \\ + a_n\phi(||\mathbf{r}_k - \mathbf{r}_n||), \quad k = 1, 2, \ldots n\end{aligned} \tag{4}$$

From Equation (4), the approximated function and shape function can be obtained:

$$\tilde{u}(\mathbf{x}) = \Psi \tilde{u}^e, \quad \Psi = [\Psi_1 \ldots \Psi_i \ldots \Psi_n] \tag{5}$$

More details can be found in a book by Liu [1].

3. NUMERICAL EXAMPLE

In this section, a 2D steady state convection-dominated equation is numerically analysed. The results are obtained and compared with different size of local biased support domain. Large improvement on accuracy can be observed from these results as the results of using a biased support domain.

The error indicator used in Table 1 is defined as follows:

$$e = \sqrt{\sum_{i=1}^{N} (u_i^{ex} - \hat{u}_i)^2 \Big/ \sum_{i=1}^{N} (u_i^{ex})^2} \tag{6}$$

Example: 2D steady-state convection-diffusion equation. Governing equation:

$$-\nabla \cdot (\mathbf{D}\nabla u) + \mathbf{v} \cdot \nabla u + \beta u = f(x, y), (x, y) \in \Omega = [0, 1] \times [0, 1] \tag{7}$$

$$\mathbf{D} = \begin{bmatrix} \varepsilon & 0 \\ 0 & \varepsilon \end{bmatrix}, \quad \mathbf{v} = [(3 - x), (4 - y)], \quad \beta = 1 \tag{8}$$

Dirichlet boundary conditions:

$$u|_{\partial\Omega} = 0 \tag{9}$$

The exact solution is given by

$$u^{ex} = \sin(x)\left(1 - e^{\frac{2(1-x)}{\varepsilon}}\right) y^2 \left(1 - e^{\frac{3(1-y)}{\varepsilon}}\right) \tag{10}$$

Table 1. Error result obtained with different diffusion coefficients and different biased distance.

$\varepsilon = 10^{-2}$		$\varepsilon = 10^{-4}$		$\varepsilon = 10^{-6}$	
α_u	e (%)	α_u	e (%)	α_u	e (%)
0.0	55.53	0.0	494.63	0.0	518.07
1.0	32.81	1.0	38.88	1.0	38.95
1.5	12.01	1.5	14.43	1.5	14.46
2.0	2.00	2.0	2.15	2.0	2.15

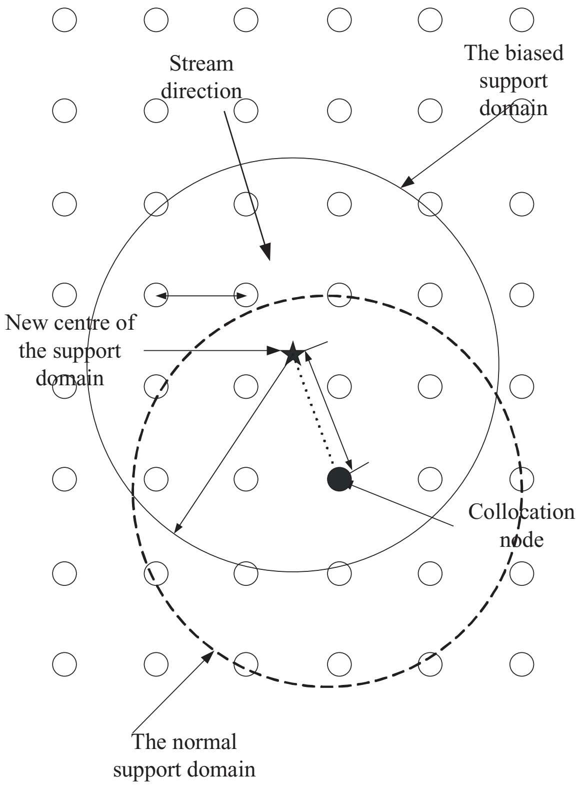


Figure 1. Schematics of a biased support domain.

The following N_d equations at internal domain nodes and N_b equations on the Dirichlet boundary Γ_b are satisfied:

$$\begin{aligned} L(\tilde{u}_i) &= -\nabla \cdot (\mathbf{D}\nabla\tilde{u}_i) + \mathbf{v} \cdot \nabla\tilde{u}_i + \beta\tilde{u}_i \\ &\quad - f(x_i, y_i), \quad i = 1, 2, \ldots, N_\mathrm{d} \\ \tilde{u}_j &= 0, \quad j = 1, 2, \ldots, N_\mathrm{b} \end{aligned} \tag{11}$$

An effective and simple way to establish an upwind biased support domain is deliberately selecting more nodes in the upstream direction while forming the local support domain for a collocation node. Figure 1 shows a biased support domain constructed based on a normal support domain by moving its centre along upstream direction. Figure 2 plots results of $u(\mathbf{x})$ obtained using the Guassian RPICM with the normal and biased support domains.

Table 1 lists the errors of $u(\mathbf{x})$ with different parameter choices for the biased distance. These results are obtained by 121-node regular model with Gaussian radial basis (e^{-r^2}) using local upwind biased support domain ($\mathrm{d_s} = 2.0\mathrm{d_c},\ \mathrm{d_u} = \alpha_\mathrm{u}\mathrm{d_c}, \mathrm{d_c} = 0.1$). The results show that the RPICM with the normal support domain gives very poor accuracy for convection-dominated

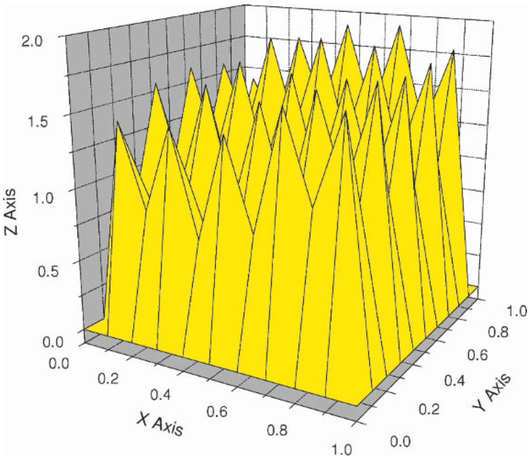


Figure 2 (a). $\alpha_u = 0.0$, $\varepsilon = 10^{-6}$.

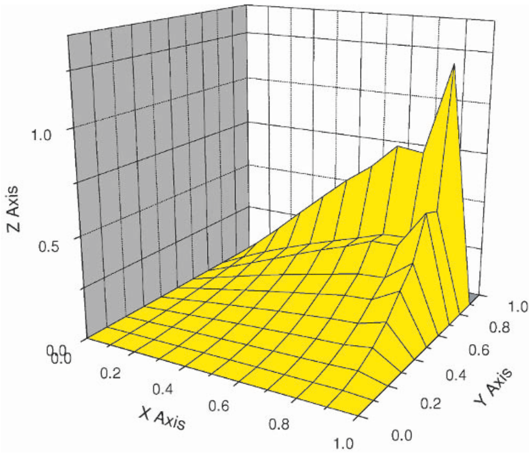


Figure 2 (b). $\alpha_u = 1.0$, $\varepsilon = 10^{-6}$.

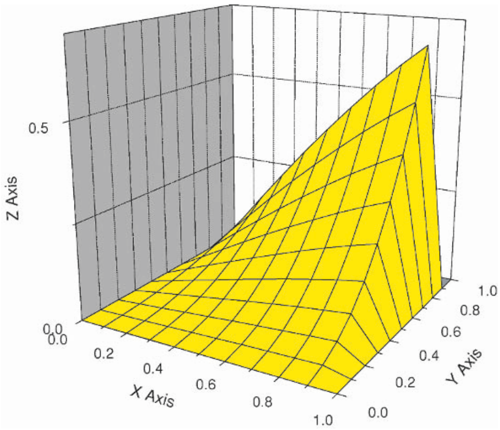


Figure 2 (c). $\alpha_u = 2.0$, $\varepsilon = 10^{-6}$.

cases. In other words, the RPICM with the biased support domain gives more accuracy when the biased distance increases.

4. CONCLUSIONS

A local upwind biased support domain scheme is used in a RPICM to solve convection-dominated equation in this paper. Much more accurate solution has been obtained. The results demonstrate that this scheme is very promising for dealing with convection-dominated problems. Further improvement can be made to adaptively move the biased domain based on the ratio of convection to diffusion, and to stabilize the collocation procedure without the use of a weak form.

REFERENCES

1. G.R. Liu (2002), *Mesh free Methods, Moving beyond the Finite Element Method*, CRC Press, Boca Raton, FL
2. Y.L. Wu and G.R. Liu (2003), A meshfree formulation of local radial point interpolation method (LRPIM) for incompressible flow simulation, *Computational Mechanics* 30, 5–6, pp. 355–365.
3. X. Liu, G.R. Liu, K. Tai and K.Y. Lam (2003a), Radial Basis Point Interpolation Collocation Method For 2d Solid Problem, In: *Advances in Meshfree and X-FEM Methods, Proceedings of the 1st Asia Workshop on Meshfree Methods*, World Scientific, Singapore pp. 35–40.
4. X. Liu, G.R. Liu, K. Tai and K.Y. Lam (2003b), Radial Point International Collocation Method (RPICM) for the solution of two phase flow through porous media. In: *Third International Conference on Computational Fluid Dynamics in the Minerals & Process Industries, 10–12 December, 2003*.
5. O.C. Zienkiewicz and R.L. Taylor (1989), *The Finite Element Method*, 4th ed., McGraw-Hill Book Company Boca Raton, FL.

MESHFREE NUMERICAL SOLUTION OF TWO-PHASE FLOW THROUGH POROUS MEDIA

Xin Liu and Y.P. Xiao
Department of Mechanics, Zhejiang University, Hangzhou, P. R. China, 310027

Abstract Thin plate spline radial point interpolation collocation method (TPS-RPICM) has been applied to the numerical simulation of two-phase flow in porous media. In the numerical testing, it was applied to solve the Five-spot Waterflood problem, which is a classical benchmark in this field of petroleum reservoir simulation. The results obtained by the TPS-RPICM for coarse and fine models were compared, and it demonstrates the accuracy and convergence of the proposed method for multiphase flow problems.

Keywords: TPS, meshfree, two-phase flow, porous media, reservoir simulation.

1. INTRODUCTION

Point interpolation method (PIM) and Radial PIM (RPIM) proposed by Liu and Wang [1] are based on Galerkin or Petrov–Galerkin weak forms, and numerical integrations are required. Its collocation form radial point interpolation collocation method (RPICM) has also been presented and applied to 2D linear elastic problems and other problems [2,3]. MQ-RPICM was applied to the numerical simulation of two-phase immiscible flow in porous media by Liu [3]. It was found that water saturation fronts obtained with coarse and fine models are lightly different. The reason is because the h-convergence is strictly dependent on the choice of shape parameter in MQ. This was also shown in the work by Lee et al. [4]. In contrast, the h-convergence is very good through some numerical tests while TPS-RPICM is employed [5]. The object of this paper is to apply TPS-RPICM to solve an immiscible flow equation

G. R. Liu et al. (eds.), Computational Methods, 1547–1553.

in reservoir simulation. The numerical results show that TPS-RPICM is more available for the solution of transient problem due to its good h-convergence feature.

In this study, the governing equations for immiscible flow and radial point interpolation approximation are briefly introduced. The collocation schemes for the pressure and saturation equations are formulated with the present method. Then, an immiscible flow equation in reservoir simulation is solved by this present method TPS-RPICM. Some concluding remarks are given finally.

1.1 Governing Equation for Two-Phase Immiscible Flow

The governing equations describing two-phase immiscible flow through porous media can be found in [3].

2. RADIAL POINT INTERPOLATION

The approximation of a function $u(x)$ may be written as

$$u(\mathbf{x}) \cong \tilde{u}(\mathbf{x}) = \sum_{i=1}^{n} a_i \varphi_i + p(\mathbf{x}) \tag{1}$$

where n is the number of points in the support domain near $\mathbf{x}.a_i$ are coefficients to be determined and φ are the TPS radial basis function, $p(\mathbf{x})$ is additional linear or quadratic polynomials.

2.1 Thin Plate Spline

$$\varphi_i = ||\mathbf{r} - \mathbf{r}_i||^{2m} \log(||\mathbf{r} - \mathbf{r}_i||), \quad ||\mathbf{r} - \mathbf{r}_i|| = \sqrt{(x - x_i)^2 + (y - y_i)^2} \tag{2}$$

The interpolations of the function at the kth point have the form:

$$\tilde{u}_k = \tilde{u}(\mathbf{x}_k) = \sum_{i=1}^{n} a_i \varphi_{ki} + p(\mathbf{x}_k), \quad \varphi_{ki} = \varphi(||\mathbf{x}_k - \mathbf{x}_i||), \quad k = l \ldots n \tag{3}$$

From (3), the approximated function and shape function can be obtained:

$$\tilde{u}(\mathbf{x}) = \Psi \tilde{u}^e, \quad \Psi = [\Psi_1 \ldots \Psi_i \ldots \Psi_n] \tag{4}$$

For more details, see for Liu [6] and X. Liu [2–5]

2.2 RPICM for Pressure and Saturation Equations

From Equation (4), the point interpolation and their derivatives forms for pressure and saturation functions can be expressed as:

$$\tilde{p}(\mathbf{x}) = \Psi\tilde{\mathbf{p}}^e, \quad \frac{\partial \tilde{p}(\mathbf{x})}{\partial \mathbf{x}} = \frac{\partial \Psi}{\partial \mathbf{x}}\tilde{\mathbf{p}}^e, \quad \tilde{S}_w(\mathbf{x}) = \Psi\tilde{S}_w^e, \quad \frac{\partial \tilde{S}_w(\mathbf{x})}{\partial \mathbf{x}} = \frac{\partial \Psi}{\partial \mathbf{x}}\tilde{S}_w^e \tag{5}$$

At time t^{m+1}, the following equations are satisfied in internal domain nodes:

$$-\nabla \cdot (\lambda^{m+1}\nabla\tilde{p}^{m+1}) + \frac{r_c}{2}\nabla(\nabla \cdot (\lambda^{m+1}\nabla\tilde{p}^{m+1})) = 0 \tag{6}$$

$$\begin{aligned}\phi\frac{\tilde{S}_w^{m+1} - \tilde{S}_w^m}{\Delta t} &+ \mathbf{v}_T^{m+1}\left(\frac{df_w^{m+1}}{d\tilde{S}_w}\right) \cdot \nabla\tilde{S}_w^{m+1} - \frac{r_c}{2}\nabla \cdot \left(\mathbf{v}_T^{m+1}\frac{df_w^{m+1}}{d\tilde{S}_w} \cdot \nabla\tilde{S}_w\right)\\ &= \varepsilon\Delta\tilde{S}_w^{m+1}\end{aligned} \tag{7}$$

It should be noted that a small diffusion term has been added to (7) to handle shocks and ε is the adjustable diffusion coefficient. Here, implicit time integration scheme has been adopted. The additional term involving in r_c is taken as numerical stable term, r_c is the characteristic length for the discretization model. For uniform model, r_c is taken as the uniform distance between nodes; For scattered point model, r_c is taken as the size of support domain at collocation point.

The following equations are satisfied on reflected boundaries for both pressure and saturation equations:

$$\mathbf{n} \cdot \nabla\tilde{p}^{m+1} - \frac{r_c}{2}(\nabla \cdot (\lambda^{m+1}\nabla\tilde{p}^{m+1})) = 0, \quad \mathbf{n} \cdot \nabla\tilde{S}_w^{m+1} = 0 \tag{8}$$

The following equations are satisfied on Dirichlet boundaries for both pressure and saturation:

At source well for injected fluid (water),

$$\tilde{p}^{m+1} = p_0, \qquad \tilde{S}_w^{m+1} = 1.0, \tag{9}$$

At sink well for produced fluid (oil),

$$\tilde{p}^{m+1} = 0, \qquad \tilde{S}_w^{m+1} = 0.0, \tag{10}$$

The last system algebraic equation can be solved by Newton–Raphson iteration scheme.

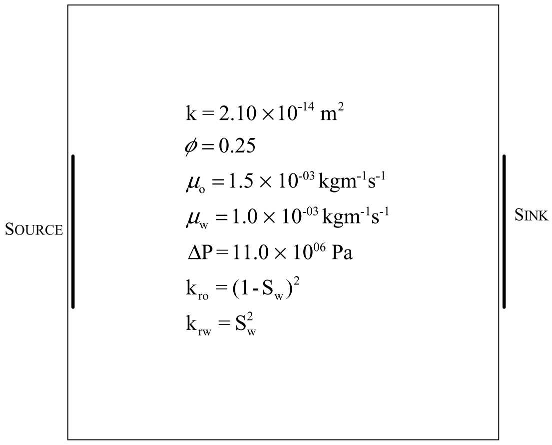


Figure 1. 2D horizontal homogeneous reservoir.

2.3 Numerical example

In this section, a 2D homogeneous two-phase immiscible reservoir problem [7] has been tested by TPS-RPICM in which linear additional term in Equation (1) and m = 4 in Equation (2) are used. Owing to symmetry, only the upper half of the reservoir needs to be discretized. Source and sink nodes with specified p values are indicated by the thick lines in Figure 1. The size of the reservoir is [0, 400.0 m] × [0, 400.0 m]. S_w was kept equal to unity at the source. The 25 ×13 = 325-node coarse model, 49 ×25 = 1225-node fine uniform model, 1225-node Halton scattered point model shown as Figure 2 are used to solve the pressure and saturation equations. Figures 3–5 show the wetting fluid saturation distributions at 2000 days for three different models. There are 25 points in every support domain. For 325-node uniform model, time step numbers: 500; time interval: 4 day; diffusion coefficient $\varepsilon = 0.01$, stable parameter $r_c = 16.0$

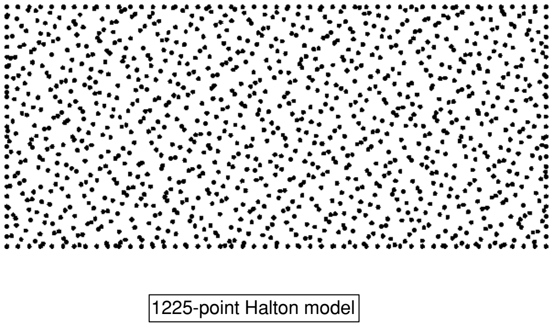


Figure 2. 1225-node Halton model.

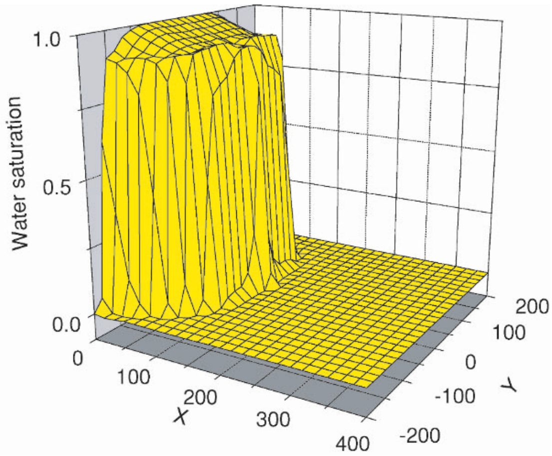


Figure 3. Saturation distribution obtained with 25 × 13 uniform model.

for dealing with Neumann boundary in pressure equation. For 1225-node uniform and scattered point models time step numbers: 2000; time interval: 1 day; diffusion coefficient $\varepsilon = 0.005$. For 1225-node uniform model, stable parameter $r_c = 8.0$ for dealing with Neumann boundary in pressure equation. For 1225-node scattered point model, stable parameter r_c is taken as the size of support domain for dealing with Neumann boundary in pressure equation. It can be observed that results obtained with coarse and fine models are very close at water saturation front. However, the numerical oscillation is smaller for fine model, and higher accurate water saturation front has been obtained with fine model. On the other hand, it is also seen that the results with the uniform and

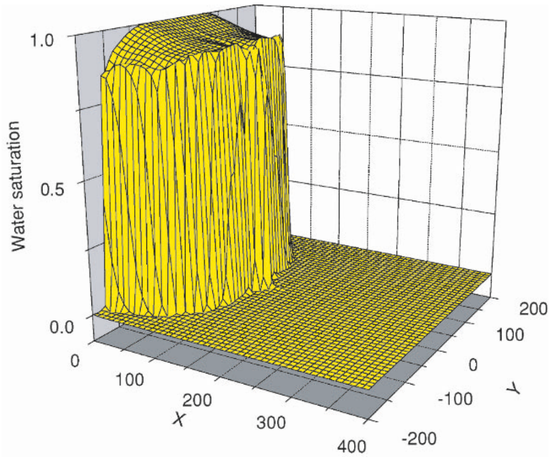


Figure 4. Saturation distribution obtained with 49 × 25 uniform model.

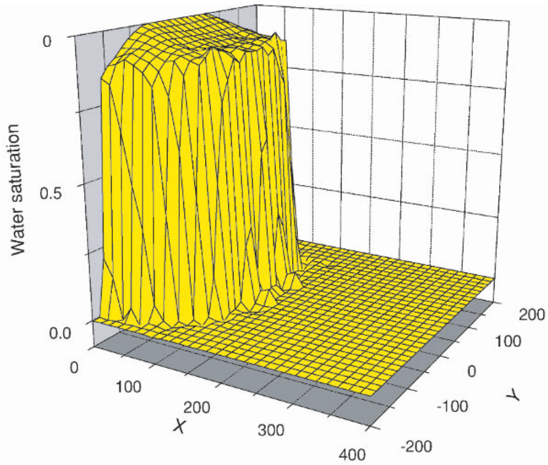


Figure 5. Saturation distribution obtained with 1225-node Holton model.

scattered point models show little difference, and it shows this present method is suitable for the scattered point model. In addition, the results obtained by TPS-RPICM are compared with Figure 13 in Langtangen's paper [7], and show good agreement.

3. CONCLUSIONS

We have employed TPS-RPICM to solve immiscible two-phase flow in porous media. The results computed demonstrate that the h-convergence is good when TPS is used in RPICM. The good h-convergence feature will be very important for further numerical simulation research in practical engineering problems.

REFERENCES

1. G. R. Liu and J.G. Wang (2002), A point interpolation meshless method based on radial basis functions. *International Journal for Numerical Methods in Engineering*, 54, pp. 1623–1648.
2. Xin Liu, G. R. Liu, K. Tai and K. Y. Lam (2003a), Radial Basis Point Interpolation Collocation Method For 2-d Solid Problem. In: *Advances in Meshfree and X-FEM Methods, Proceedings of the 1st Asin Workshop on Meshfree Methods*, World Scientific, Singapore, pp. 35–40.
3. Xin Liu, G. R. Liu, K. Tai and K. Y. Lam (2003b), Radial Point International Collocation Method (RPICM) for the Solution of Two Phase Flow Through Porous Media. In: *Third*

International Conference on Computational Fluid Dynamics in the Minerals and Process Industries, 10-12 December, 2003.

4. C. K. Lee, X. Liu, and S. C. Fan (2003), Local multiquadric approximation for solving boundary value problems. *Computational Mechanics*, 20, pp. 396–409.
5. Xin Liu, G. R. K. Tai and K. Y. Lam (2004), Radial point interpolation collocation method for the solution of partial differential equations. *Computers and Mathematics with Applications*, (submitted).
6. G. R. Liu (2002), *Mesh free Methods, Moving beyond the Finite Element Method*, CRC Press Boca Raton, FL.
7. H.P. Langtangen (1990), Implicit finite element methods for two-phase flow in oil reservoirs. *International Journal of Numerical Methods in fluids*, 10, pp. 651–681.

STRESS ANALYSIS OF 3-D SOLIDS USING A MESHFREE RADIAL POINT INTERPOLATION METHOD

G.R. Liu[1,2], G.Y. Zhang[1] and Y.T. Gu[1]
[1] *Centre for Advanced Computations in Engineering Science (ACES), Department of Mechanical Engineering, National University of Singapore, 10 Kent Ridge Crescent, 119260, Sigapore*
[2] *Singapore MIT Alliance, National University of Singapore, 9 Engineering Drive 1, Singapore 117576*

Abstract A meshfree radial point interpolation method (RPIM) is presented for stress analysis of three-dimensional (3-D) problems. In this method, the radial basis function augmented with polynomial is used to construct shape functions. As the shape functions so constructed have the Kronecker delta function property, the RPIM enforces essential boundary conditions much more easily than the meshfree methods based on the moving least-square approximation. The Galerkin weak form of the governing equations is used to derive the discretized system equations. The stress analysis of a numerical example is studied and it is concluded that the present meshfree RPIM is very accurate and efficient for stress analysis of 3-D solids.

Keywords: meshfree, meshless, radial point interpolation, three-dimensional.

1. INTRODUCTION

Meshfree methods have been developed and achieved remarkable progress in recent years. There are some well-known ones such as smooth particle hydrodynamics (SPH) [1], the diffuse element method (DEM) [2], the element free Galerkin (EFG) method [3] and the meshless local Petrov-Galerkin (MLPG) method [4], and so on.

The point interpolation method (PIM) [5–8] is a meshfree method developed using local shape functions and Galerkin weak form. In PIM, the shape function is constructed based only on a group of nodes arbitrarily distributed in

G. R. Liu et al. (eds.), Computational Methods, 1555–1559.

a local support domain by means of interpolation. A background cell structure is required to evaluate the integration in the Galerkin weak form. The major advantage of PIM is that the shape functions created possess the Kronecker delta function property, and hence many numerical treatments are as simple as in the finite element method (FEM).

PIM using radial basis functions (RBFs) is termed as radial PIM (RPIM) [7, 8]. In the RPIM, RBFs are used for constructing shape functions and it has been proved that the moment matrix of the RBF interpolation is usually invertible for arbitrary scattered nodes [9]. Therefore, the RPIM is very stable and robust for arbitrary nodal distributions.

In this paper, the three-dimensional (3-D) RPIM formulation is first developed. A numerical example is then presented to demonstrate the accuracy and efficiency of RPIM for 3-D solids.

2. RADIAL POINT INTERPOLATION

Consider an approximation function $u(\mathbf{x})$ in a local support domain that discretized by a set of arbitrarily distributed nodes. We have the following RPIM form of the function $u(\mathbf{x})$using radial basis function $R_i(\mathbf{x})$ augmented with polynomial basis function $p_j(\mathbf{x})$ [8, 9].

$$u(\mathbf{x}) = \sum_{i=1}^{n} R_i(\mathbf{x}) a_i + \sum_{j=1}^{m} P_j(\mathbf{x}) b_j = \mathbf{R}^T(\mathbf{x})\mathbf{a} + \mathbf{P}^T(\mathbf{x})\mathbf{b} \tag{1}$$

where $n$is the number of RBFs and is also identical to the number of nodes in the local support domain of interest point $\mathbf{x}$, and m is the number of polynomial basis functions. There are a number of types of RBFs and the multi-quadrics (MQ) function is adopted in this paper. In order to determine a_i and b_j in Equation (1), a local support domain must be formed. In this paper, the $n$nodes which are the nearest to the interest point $\mathbf{x}$ are adopted in the support domain. It leads to n linear equations by enforcing Equation (1) to be satisfied at these n nodes in the support domain. Finally, the approximation function can be rewritten as the follows

$$u(\mathbf{x}) = \Phi^T(\mathbf{x})\mathbf{U}_s = \sum_{i=1}^{n} \phi_i u_i \tag{2}$$

where $\Phi(\mathbf{x})$ is the vector of shape functions corresponding to the nodal displacements,

$$\Phi^T(\mathbf{x}) = \left\{ \phi_1(\mathbf{x})\ \phi_2(\mathbf{x}) \ldots \phi_n(\mathbf{x}) \right\} \tag{3}$$

3. GLOBAL WEAK FORM FOR THREE-DIMENSIONAL SOLIDS

Consider the following 3-D problem defined in the domain Ω boundary by Γ [10].

$$\mathbf{L}^T\boldsymbol{\sigma} + \mathbf{b} = 0 \quad \text{in } \Omega \tag{4}$$

The natural and the essential boundary conditions are given as the follows, respectively,

$$\boldsymbol{\sigma} \cdot \mathbf{n} = \bar{\mathbf{t}} \quad \text{on } \Gamma_t; \qquad \mathbf{u} = \bar{\mathbf{u}} \quad \text{on } \Gamma_u \tag{5}$$

The unconstrained Galerkin weak form of Equation (4) is posed as the follows [8].

$$\int_{\Omega} (\boldsymbol{L}\delta\boldsymbol{u})^T(\boldsymbol{D}\boldsymbol{L}\boldsymbol{u})\, d\Omega - \int_{\Omega} \delta\boldsymbol{u}^T\boldsymbol{b}\, d\Omega - \int_{\Gamma_t} \delta\boldsymbol{u}^T\bar{\boldsymbol{t}}\, d\Gamma = 0 \tag{6}$$

Substituting Equation (2) into Equation (6) yields

$$\mathbf{K}\mathbf{u} = \mathbf{f} \tag{7}$$

where

$$\boldsymbol{K}_{ij} = \int_{\Omega} \boldsymbol{B}_i^T\boldsymbol{D}\boldsymbol{B}_j\, d\Omega; \quad \boldsymbol{f}_i = \int_{\Gamma_t} \phi_i\bar{\boldsymbol{t}}\, d\Gamma + \int_{\Omega} \phi_i\boldsymbol{b}\, d\Omega \tag{8}$$

4. A NUMERICAL EXPERIMENT

In this section, a 3-D cantilever beam, shown in Figure 1, is studied as a testing example using the RPIM. In this study, the units are all taken as international standard units and the material adopted is linear elastic with $E = 3.0 \times 10^7$ and $\nu = 0.3$. The parameters of the beam are taken as $L = 50$, $D = 10$, $B = 1$ and $p = 1000$. For the RPIM, $q = 1.03$, $\alpha_c = 4.0$ and 52 nodes

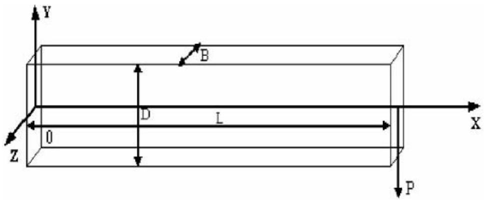


Figure 1. A cantilever beam.

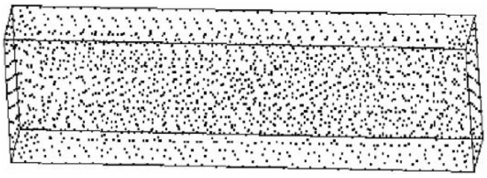


Figure 2. Irregular nodal distribution.

are adopted in the support domain. The cantilever beam is represented by 1620 irregularly distributed nodes (shown in Figure 2) and 4447 tetrahedron-shaped background integration cells are used. There are 11 Gauss points in each tetrahedron cell for integration. As the beam has a unit thickness and a plane stress problem can be considered to yield the analytical solution [10], this analytical solution is adopted to act as the reference solution for the cantilever.

Figure 3 illustrates a comparison of normal stress in x direction along a particular line between the analytical solutions and the RPIM results. The plot shows that the results obtained using RPIM in 3-D are in great agreement with

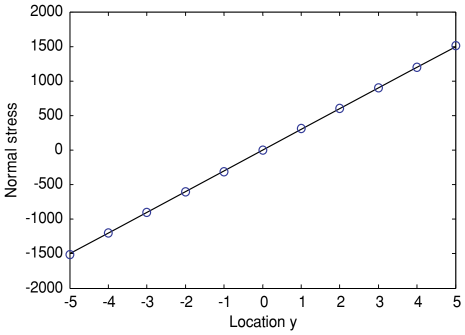


Figure 3. Normal stress (σ_{xx}) distribution along the line located at $x = L/2, z = 0.0$.

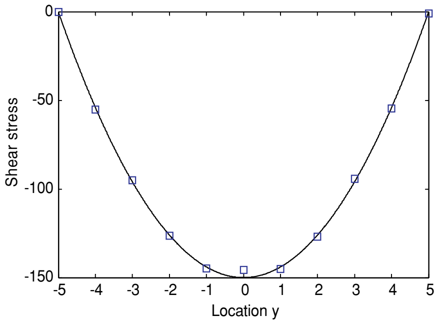


Figure 4. Shear stress (σ_{xy}) distribution along the line located at $x = L/2, z = 0.0$.

the analytical solutions. Figure 4 shows that the shear stresses obtained using the RPIM have a good agreement with the analytical solutions.

The present RPIM code is also used for analyzing complicated 3-D mechanical components, and the results will be presented in oral presentation.

5. CONCLUSIONS

A RPIM formulation for 3-D solids is presented in this paper. In this method, the combination of radial basis functions and the polynomial is adopted to construct shape functions based on a 3-D local support domain. As the shape functions so derived possess Kronecker delta function property, the essential boundary conditions can be imposed as easily as in FEM. The RPIM is coded and applied for stress analysis of 3-D solids. It is found that the meshfree RPIM is very easy to implement, very accurate and efficient for stress analysis of three-dimensional problems, and with equivalent efficiency as the standard FEM. The code is being developed further for dynamic and non-linear problems.

REFERENCES

1. L. Lucy (1977), A numerical approach to testing the fission hypothesis. *Astronomical Journal*, 82, pp. 1013–1024.
2. B. Nayroles, G. Touzot and P. Villon (1992), Generalizing the finite element method: diffuse approximation and diffuse elements. *Compututational Mechanics*, 10, pp. 307–318.
3. T. Belyschko, Y.Y. Lu and L. Gu (1994), Element-free Galerkin methods. *Int. J. Numer. Meth. Engng*, 37, pp. 229–256.
4. S.N. Atluri, T. Zhu (1998), A new meshless local Petrov-Galerkin (MLPG) approach in computational mechanics. *Compututational Mechanics*, 22, pp. 117–127.
5. G.R. Liu and Y.T. Gu (2001a), A local point interpolation method for stress analysis of two-dimensional solids. *Struct. Eng. Mech*, 11, 2, pp. 221–236.
6. G.R. Liu and Y.T. Gu (2001b), A point interpolation method for two-dimensional solids. *Int. J. for Numerical Methods in Engineering*, 50, pp. 937–951.
7. J.G. Wang and G.R. Liu (2002), A point interpolation meshless method based on radial basis functions. *Int. J Numer Meth Engrg*, 54, 11, pp. 1623–1648.
8. G.R. Liu (2002), *Mesh free methods: Moving beyond the finite element method.* CRC Press, Boca Raton, USA, pp. 87–107.
9. M.J.D. Powell (1992), The theory of radial basis function approximation in 1990. In: *Advanced in Numerical Analysis*, F.W. Light (ed.), pp. 303-322.
10. S.P. Timoshenko and J.N. Goodier (1970), *Theory of Elasticity*, 3rd ed., McGraw-Hill International Editon, New York, pp. 41–46.

3-D HEAT TRANSFER ANALYSIS USING A COLLOCATION METHOD TOGETHER WITH RPIM SHAPE FUNCTIONS AND FIC BOUNDARY CONDITIONS

B.H. Zhou and G.R. Liu
Department of Mechanical Engineering, National University of Singapore, 9 Engineering Drive 1, Singapore 117576

Abstract This paper describes a meshfree collocation procedure for heat transfer analysis. In this procedure, radial point interpolation method (RPIM) is employed to construct shape functions and a finite calculus treatment is used for problems with derivative boundary conditions to improve the stability and accuracy of numerical solution. Numerical examples of 2-D and 3-D heat transfer problems are presented to study the stabilization parameter to demonstrate the efficiency of the present numerical procedure.

Keywords: RPIM-numerical analysis, finite calculus, heat transfer, meshfree, collocation.

1. INTRODUCTION

Meshfree methods can be generally classified into three major categories: methods based on strong forms (e.g., collocation), those based on weak forms and those based on the combination of weak and strong forms [1]. A comprehensive review for meshfree methods can be found in the books by Liu [2, 3].

Meshfree collocation method has been used extensively due to its simplicity in implementation and computational efficiency. However, the poor stability and low accuracy for problems with derivative boundary conditions (DBCs) constitutes the main drawback for this type of methods. Many treatments aiming to alleviate such a flaw have been proposed in recent years, and one of which is the use of finite calculus (FIC) [4, 5].

In this paper, the radial point interpolation method (RPIM) augmented with polynomials is used to construct shape functions with Kronecker delta function

G. R. Liu et al. (eds.), Computational Methods, 1561–1566.

properties, which allows easy handling of essential boundary conditions. A collocation procedure using the RPIM shape functions is then employed to analyse heat transfer problem numerically. Finite calculus (FIC) concept is adopted for the treatment of DBCs. Several numerical examples of heat transfer problem are solved via the present collocation method, the results show a significant improvement in accuracy and efficiency, when a proper stabilization parameter is used.

2. RADIAL POINT INTERPOLATION

2.1 Formulation Using Radial-Polynomial Basis

Interpolation of functions with pure radial basis functions fails to capture the linear field exactly due to the inconsistency nature (see, e.g., [2]). Adding polynomial terms up to linear order can ensure the reproduction of a linear field, and it can usually improve the accuracy of the results.

Consider a temperature field function $T(\mathbf{x})$ defined in a domain represented by a set of nodes. The approximate interpolation $T(\mathbf{x})$ by using the surrounding nodes in a local support domain of a point at $\mathbf{x}$ can be written as

$$T(\mathbf{x}) = \sum_{i=1}^{n} \phi_i(\mathbf{x})\, T_i \tag{1}$$

where ϕ_i is a shape function and T_i is the function value at node i in the support domain. The shape functions can be obtained through a standard procedure (see, e.g., [2]), using different types of radial basis functions. In this paper, the multiquadric (MQ) radial functions are used. Once the shape functions are obtained, Equation (1) is then used in the standard collocation procedure to discretize the governing equations, which leads to a set of algebraic equations. Since the RPIM shape functions possess the delta function property, the essential boundary conditions can be imposed as in the conventional FEM or FEM.

2.2 Finite Calculus for the DBCs

Collocation method together with RPIM shape functions can usually produce accurate results for PDEs, for cases when only essential boundary conditions exist. However, the solution accuracy can drop significantly when a DBC presents. To improve the accuracy in such a case, a stabilization technique called finite calculus approach has been suggested by Oñate *et al.* [4, 5]. The FIC procedure, imposing balance laws of mechanics over a domain of finite size, is used here only to deal with the derivative boundary conditions.

Consider the following 3-D problem of heat transfer defined in a cubic domain $[0 \leq x \leq L, 0 \leq y \leq L, 0 \leq z \leq L]$, where L is a given constant.

$$\nabla^2 T + \frac{3\pi^2}{L^2} \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{\pi y}{L}\right) \sin\left(\frac{\pi z}{L}\right) = 0 \tag{2}$$

where ∇^2 denotes Laplace operator, and T is temperature variable in the cubic domain. The derivative boundary condition on $x = 0$ or $x = L$ is given as follows.

$$\nabla^2 T + \frac{3\pi^2}{L^2} \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{\pi y}{L}\right) \sin\left(\frac{\pi z}{L}\right) = 0 \tag{3}$$

On the rest of five surfaces of the cubic domain, the temperature is fixed at

$$\overline{T} = \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{\pi y}{L}\right) \sin\left(\frac{\pi z}{L}\right) \tag{4}$$

Using finite calculus procedure, the DBC equations can be rewritten as

$$\begin{aligned} &\left(\frac{\partial T}{\partial x} - \frac{\pi}{L} \cos\left(\frac{\pi x}{L}\right) \sin\left(\frac{\pi y}{L}\right) \sin\left(\frac{\pi z}{L}\right)\right) \\ &\quad + \alpha d_c \mathbf{n} \left(\nabla^2 T + \frac{3\pi^2}{L^2} \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{\pi y}{L}\right) \sin\left(\frac{\pi z}{L}\right)\right) = 0 \end{aligned} \tag{5}$$

where $\mathbf{n}$ is the outwards normal on the boundary, d_c is the distance in the normal direction between the sampling node on the boundary and the nearest node within the local support domain. We introduce α as a stabilization parameter to be adjusted to achieve better accuracy.

The least square norm is defined as error indicator in our accuracy study.

$$error = \frac{1}{N} \sqrt{\left(\sum_{i=1}^{N} \left(T_i^{\text{exact}} - T_i^{\text{num}}\right)^2\right) \Big/ \left(\sum_{i=1}^{N} \left(T_i^{\text{exact}}\right)^2\right)} \tag{6}$$

where T_i^{exact} is exact values of the temperature, T_i^{num} is numerical values obtained using the present numerical method, and N is total number of nodes in the cubic domain.

The derivation for 2-D heat transfer problem based on the above formulation is straight, and hence is omitted here.

3. RESULTS AND DISCUSSION

Some benchmarking numerical examples with exact analytical solutions are studied to test the present collocation procedure.

Figure 1 shows the temperature distribution for a 2-D problem defined in $[0 \leq x \leq 100, 0 \leq y \leq 100]$. The problem has been solved with the derivative

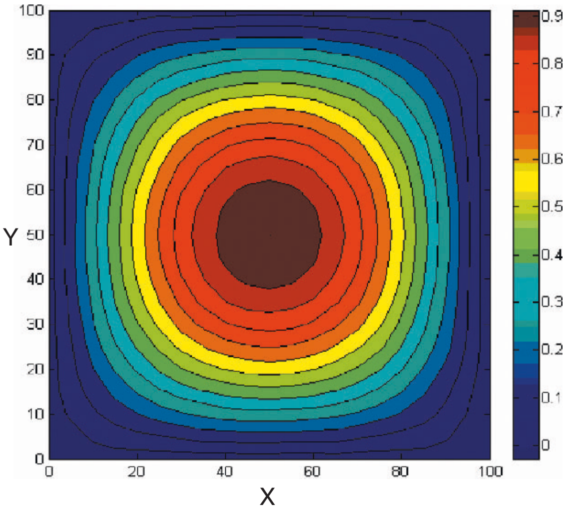


Figure 1. Temperature distribution in a 2-D square domain.

boundary condition at $x = 100$ and 121 evenly distributed nodes are used. Figure 2 shows the least square error obtained using Equation (6) when different stabilization parameter α is used. It can be found that better results have been obtained when α is around 0.4. An example of 3-D heat transfer problem

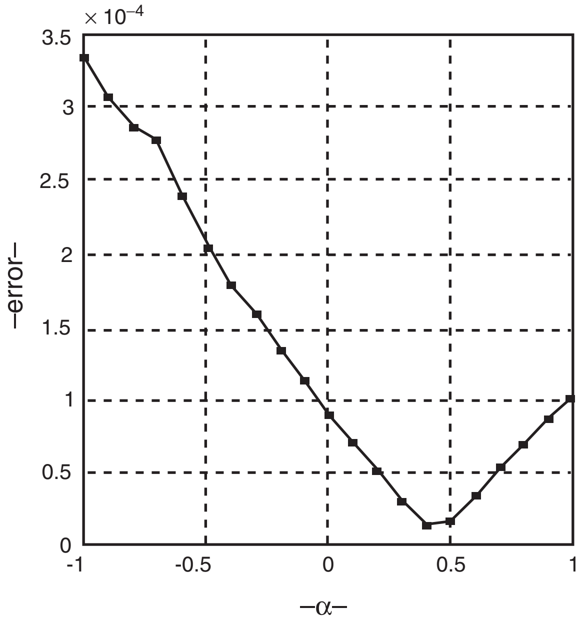


Figure 2. Least square error in the temperature for different parameter α used.

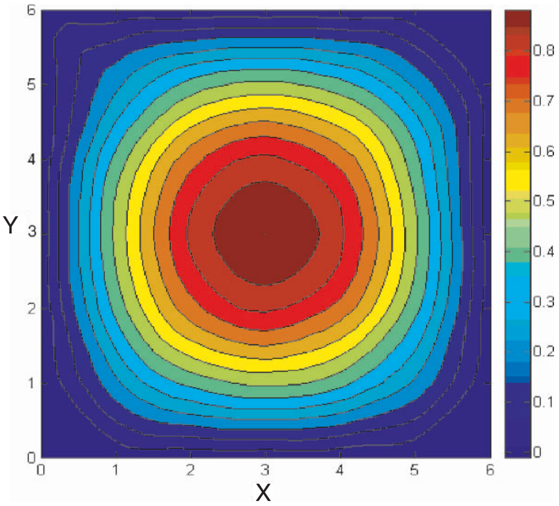


Figure 3. Temperature distribution on a plane at z = 3 in a cubic domain.

is defined in a cubic domain $[0 \leq x \leq 6, 0 \leq y \leq 6, 0 \leq z \leq 6]$. $7 \times 7 \times 7$ regularly distributed nodes are used for the numerical analysis. Figure 3 shows the temperature distribution on the cross -section at $z = 3$. The least square error for different stabilization parameter α is shown in Figure 4. It is found that accurate solution can be obtained when $\alpha = 0.4 \sim 1.0$ in this case.

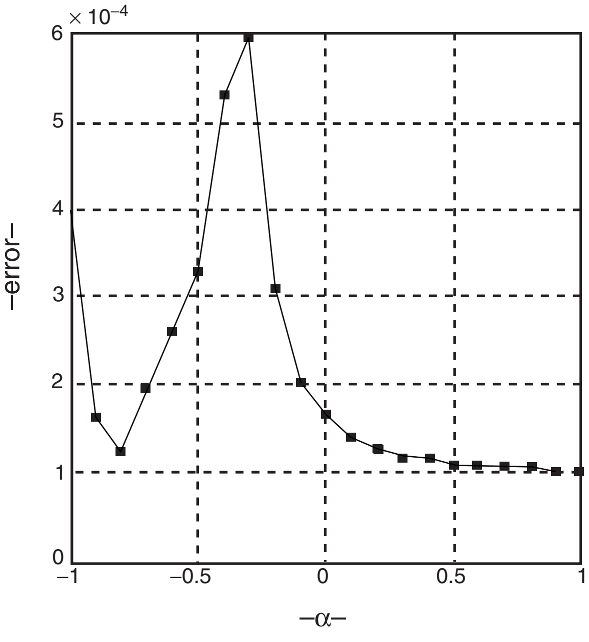


Figure 4. Least square error in the temperature for different parameter α used.

4. CONCLUSIONS

A collocation method, together with RPIM shape functions and a FIC procedure, is used to analyze 3-D heat transfer problem. Numerical examples have shown that the stabilization parameter α affects significantly the accuracy of the solution, and it should be within the range of 0.4 to 1.0 for our example problems.

REFERENCES

1. G.R. Liu and Y.T. Gu (2003), A meshfree method: meshfree weak-strong (MWS) from method for 2-D solids. *Computational Mechanics*, 33(1), pp. 2–14.
2. G.R. Liu (2002), *Mesh Free Methods: Moving Beyond the Finite Element Method*. CRC Press, Boca Raton.
3. G.R. Liu and M.B. Liu (2003), *Smoothed Particle Hydrodynamics: A Meshfree Particle Method,* World Scientific, New Jersey.
4. E. Onate, F. Perazzo and J. Miquel (2001), A finite point for elasticity problems. *Comuperts and Structures*, 79 (2001), pp. 2151–2163.
5. E. Onate (2004), Possibilities of finite calculus in computational mechanics. *International Journal for Numerical Methods in Engineering*, 60, pp. 255–281.

SYMPLECTIC ANALYSIS FOR OPTICAL WAVEGUIDES IN LAYERED MEDIA

Zheng Yao and Wanxie Zhong
Department of Engineering Mechanics, Dalian University of Technology, Dalian, China 116023

Abstract Symplectic analysis is introduced into optical waveguide theory, by using Hamiltonian system theory in which the transverse electric and magnetic field vectors are the dual vectors. The method can accommodate arbitrary anisotropic material. The electro-magnetic stiffness matrix of a typical segment in optical waveguide is generated by means of the combination algorithm of contiguous segments based on energy variational principle. Then the Wittrick–Williams algorithm is used to extract the eigenvalues. Thereafter, an energy band analysis is performed for optical waveguide in layered media.

Keywords: symplectic, optical waveguide, layered media, Wittrick–Williams algorithm.

1. INTRODUCTION

Frequency spectrum analysis of waveguides shows that the eigenvalues exhibit *energy band* behaviour [1], such that a wave with frequency ω which is in a pass-band can propagate along the structure, whereas otherwise ω is in a stop-band and the wave decays to zero over long distances. The energy analysis used to find pass-bands and stop-bands is very important in practical disciplines. In the analysis of waveguides, the layered structure is encountered frequently. The optical wave is electro-magnetic wave, the analysis of optical waveguides is the topic of the present paper.

G. R. Liu et al. (eds.), Computational Methods, 1567–1579.

2. SYMPLECTIC ANALYSIS FOR ELECTRO-MAGNETIC WAVE

The Maxwell equation set is the foundation for this paper. For optical materials there is usually neither source nor current (i.e., $\rho = 0,\ \mathbf{j} = 0$) and the equations are given as [1–3]

$$\nabla \cdot \mathbf{D} = \mathbf{0}, \tag{1}$$

$$\nabla \cdot \mathbf{B} = \mathbf{0} \tag{2}$$

$$\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t = -\mu \partial \mathbf{H}/\partial t, \tag{3}$$

$$\nabla \times \mathbf{H} = \partial \mathbf{D}/\partial t = \varepsilon \partial \mathbf{E}/\partial t \tag{4}$$

where **E** and **H** are, respectively, the electric and magnetic field vectors; **D** and **B** are electric displacement density and magnetic flux density vectors. The constitutive relations are

$$\mathbf{D} = \varepsilon \mathbf{E}, \tag{5}$$

$$\mathbf{B} = \mu \mathbf{H} \tag{6}$$

The Maxwell equations are formulated in the time domain, whereas the corresponding forms in the frequency domain are

$$\mathbf{H} = \mathbf{h} e^{-\mathrm{i}\omega t}, \quad \mathbf{E} = \mathrm{i}\mathbf{e} e^{-\mathrm{i}\omega t} \tag{7}$$

where $\mathbf{e}(x, y, z, \omega)$ and $\mathbf{h}(x, y, z, \omega)$ are to be determined. Equation (3) and (4) transform to

$$\omega \mu_0 \mathbf{h} = \mathbf{R} \cdot \mathbf{e}, \tag{8}$$

$$\omega \varepsilon \mathbf{e} = \mathbf{R} \cdot \mathbf{h} \tag{9}$$

$$\text{where } \mathbf{R} = \begin{bmatrix} 0 & -\partial/\partial z & \partial/\partial y \\ \partial/\partial z & 0 & -\partial/\partial x \\ -\partial/\partial y & \partial/\partial x & 0 \end{bmatrix} \tag{10}$$

is an operator matrix.

If S is the boundary surface of a finite domain V; the boundary condition for a perfect conductor is

$$\mathbf{n} \times \mathbf{e} = 0, \quad \mathbf{n} = \mathbf{i}_x l + \mathbf{i}_y m + \mathbf{i}_z n, \text{ on the boundary } S \tag{11}$$

For a finite domain V with perfect conductor boundary conditions, the variational principle can be expressed as

$$\Pi(\mathbf{e},\mathbf{h}) = \mathrm{Re}\left\{\begin{matrix}\iiint_V \left[\mathbf{h}^H\cdot(\mathbf{R}\cdot\mathbf{e}) - \mu_0\omega\mathbf{h}^H\mathbf{h}/2 - \omega\mathbf{e}^H\boldsymbol{\varepsilon}\mathbf{e}/2\right]\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \\ +\iint_S \mathbf{e}^H\cdot(\mathbf{n}\times\mathbf{h})\mathrm{d}S\end{matrix}\right\},\ \delta\Pi = 0 \tag{12}$$

where superscript H denotes Hermitian transposition. The components of vectors **h** and **e** are treated as independently varying functions in the functional. Note that the boundary condition (11) is also the natural boundary condition derived from the variational principle [4].

The configuration of the longitudinal co-ordinate z is different to that of the transverse ones x and y: Let partial differential with respect to it be denoted by $\partial(\#)\big/\partial z = \dot{(\#)}$ and let the transverse vectors **q** and **p** be written as

$$\mathbf{q} = \{e_x \quad e_y\}^T,\ \mathbf{p} = \{h_y \quad -h_x\}^T,\ q_1 = e_x,\ q_2 = e_y,\ p_1 = h_y,\ p_2 = -h_x \tag{13}$$

These form a pair of dual vectors. The variational principle can be rewritten with dual vectors **q** and **p**:

$$\Pi(\mathbf{q},\mathbf{p}) = \iiint_V \times \begin{bmatrix}\mathbf{p}^T\dot{\mathbf{q}} - \omega\mathbf{q}^T\boldsymbol{\varepsilon}_{tr}\mathbf{q}/2 - \mu_0\omega\mathbf{p}^T\mathbf{p}/2 - (\partial p_1/\partial x + \partial p_2/\partial y)(\boldsymbol{\varepsilon}_{zt}\mathbf{q})/\varepsilon_z \\ +(\partial p_1/\partial x + \partial p_2/\partial y)^2/(2\varepsilon_z\omega) + (\partial q_1/\partial y - \partial q_2/\partial x)^2/(2\mu_0\omega)\end{bmatrix}\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z,$$
$$\delta\Pi = 0 \tag{14}$$

where $$\boldsymbol{\varepsilon}_{tr} = \boldsymbol{\varepsilon}_t - \boldsymbol{\varepsilon}_{zt}\varepsilon_z^{-1}\boldsymbol{\varepsilon}_{zt}^T = \begin{bmatrix}\varepsilon_{t11} & \varepsilon_{t12} \\ \varepsilon_{t21} & \varepsilon_{t22}\end{bmatrix},\ \boldsymbol{\varepsilon}_{zt} = [\,\varepsilon_{xz} \quad \varepsilon_{yz}\,] \tag{15}$$

The variational principle gives

$$\dot{\tilde{\mathbf{q}}}(x,y,z) = \mathbf{A}\tilde{\mathbf{q}} + \mu_0\omega\tilde{\mathbf{p}} + \mathbf{D}_{\mathrm{op}}\tilde{\mathbf{p}} \tag{16a}$$

$$\dot{\tilde{\mathbf{p}}}(x,y,z) = -\omega\boldsymbol{\varepsilon}_{tr}\tilde{\mathbf{q}} - \mathbf{B}_{\mathrm{op}}\tilde{\mathbf{q}} - \mathbf{A}^T\tilde{\mathbf{p}} \tag{16b}$$

where $$\mathbf{D}_{\mathrm{op}} \underset{\mathrm{def}}{=} \frac{1}{\omega\varepsilon_z}\begin{bmatrix}\partial^2/\partial x^2 & \partial^2/\partial x\partial y \\ \partial^2/\partial x\partial y & \partial^2/\partial y^2\end{bmatrix}, \tag{17a}$$

$$\mathbf{B}_{\mathrm{op}} \underset{\mathrm{def}}{=} \frac{1}{\omega\mu_0}\begin{bmatrix}\partial^2\big/\partial y^2 & -\partial^2\big/\partial x\partial y \\ -\partial^2\big/\partial x\partial y & \partial^2\big/\partial x^2\end{bmatrix} \tag{17b}$$

$$\mathbf{A} = -\begin{bmatrix}\partial\big/\partial x \\ \partial\big/\partial y\end{bmatrix}\left(\frac{\boldsymbol{\varepsilon}_{zt}}{\varepsilon_z}\bullet\right), \tag{18a}$$

$$\mathbf{A}^T = \left(\frac{\boldsymbol{\varepsilon}_{zt}^T}{\varepsilon_z}\right)\left[\partial/\partial x \quad \partial/\partial y\right]\bullet \tag{18b}$$

The tilde over dual vectors denotes that **q** and **p** are the functions of (x, y, z).

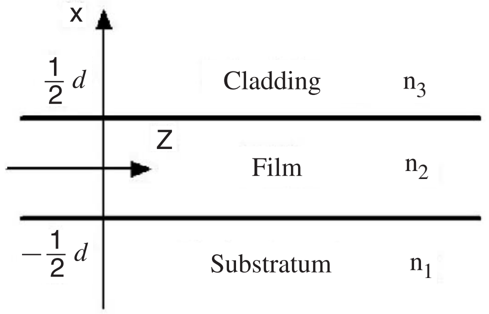


Figure 1. Layered media.

3. SOLUTION FOR A HOMOGENEOUS ISOTROPIC PLANE WAVEGUIDE

In the simplest case the field is invariant in the y direction, see Figure 1.

$$\begin{aligned}\tilde{\mathbf{q}}(x,z) &= \mathbf{q}(x)\cdot\exp(\mathrm{i}k_z z)\\ \tilde{\mathbf{p}}(x,z) &= \mathbf{p}(x)\cdot\exp(\mathrm{i}k_z z)\end{aligned} \tag{19}$$

For isotropic material, substituting (19) into (16a,b) gives

$$\begin{aligned}&\mathrm{d}^2 p_1/\mathrm{d}x^2 = [k_z^2(\varepsilon_z/\varepsilon_{t11}) - \mu_0\omega^2\varepsilon_z]p_1,\ \mathrm{d}^2 q_2/\mathrm{d}x^2 = [k_z^2 - \omega^2\mu_0\varepsilon_{t22}]q_2\\ &\mathrm{i}k_z q_2 = \mu_0\omega p_2,\ \mathrm{i}k_z p_1(x) = -\omega\varepsilon_t q_1; \quad q_2 = e_y,\ p_1 = h_y\end{aligned} \tag{20}$$

This set of equations has been separated the variables, and can be divided into two groups: $q_1 = e_x = 0$, $p_1 = h_y = 0$ (TE solutions) and $q_2 = e_y = 0$, $p_{21} = -h_x = 0$ (TM solutions).

The derivation can be used for all layers. The equations are very simple so that the analysis solution can be given out. The analysis for pass-band in layered media is very necessary. The precise integration method (PIM) [5, 6] in combination with extra W–W (Wittrick–Williams) algorithm [7–11] is applied to solve the problem of isotropic plane waveguide, and gives out all the pass-band eigenvalues.

4. SOLUTION FOR A HOMOGENEOUS ANISOTROPIC PLANE WAVEGUIDE

Now consider this: the media is layered along the z direction and the wave propagates in the x direction. Because the field is invariant in the y direction

$$\mathbf{E}(x,y,z) = \mathbf{e}(z)\cdot\exp[\mathrm{i}(k_x x - \omega t)], \quad \mathbf{H}(x,y,z) = \mathbf{h}(z)\cdot\exp[\mathrm{i}(k_x x - \omega t)] \tag{21}$$

substituting (21) into (16a,b) gives

$$\dot{\mathbf{q}} = -\begin{bmatrix} \mathrm{i}\varepsilon_{xz}k_x/\varepsilon_z & \mathrm{i}\varepsilon_{yz}k_x/\varepsilon_z \\ 0 & 0 \end{bmatrix}\mathbf{q} + \begin{bmatrix} \omega\mu_0 - k_x^2/\omega\varepsilon_z & 0 \\ 0 & \omega\mu_0 \end{bmatrix}\mathbf{p} \tag{23a}$$

$$\dot{\mathbf{p}}(z) = -\begin{bmatrix} \omega\varepsilon_{t11} & \omega\varepsilon_{t12} \\ \omega\varepsilon_{t21} & \omega\varepsilon_{t22} - k_x^2/\omega\mu_0 \end{bmatrix}\mathbf{q} - \begin{bmatrix} \mathrm{i}\varepsilon_{xz}k_x/\varepsilon_z & 0 \\ \mathrm{i}\varepsilon_{yz}k_x/\varepsilon_z & 0 \end{bmatrix}\mathbf{p} \tag{23b}$$

The above equations can be written in matrix form

$$\dot{\mathbf{v}}(z) = \mathbf{H}\mathbf{v}, \quad \mathbf{v} = \begin{Bmatrix} \mathbf{q} \\ \mathbf{p} \end{Bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \mathbf{A} & \mathbf{D} \\ \mathbf{B} & -\mathbf{A}^H \end{bmatrix}, \quad \mathbf{JH} = (\mathbf{JH})^H \tag{24}$$

$$\mathbf{A} = -\begin{bmatrix} \mathrm{i}\varepsilon_{xz}k_x/\varepsilon_z & \mathrm{i}\varepsilon_{yz}k_x/\varepsilon_z \\ 0 & 0 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} \omega\mu_0 - k_x^2/\omega\varepsilon_z & 0 \\ 0 & \omega\mu_0 \end{bmatrix},$$

$$\mathbf{B} = -\begin{bmatrix} \omega\varepsilon_{t11} & \omega\varepsilon_{t12} \\ \omega\varepsilon_{t21} & \omega\varepsilon_{t22} - k_x^2/\omega\mu_0 \end{bmatrix} \tag{25}$$

Equations (23a,b) can be rewritten as

$$\mathbf{K}_{22}\ddot{\mathbf{q}} + (\mathbf{K}_{21} - \mathbf{K}_{12})\dot{\mathbf{q}} - \mathbf{K}_{11}\mathbf{q} = 0, \quad \mathbf{K}_{22} = \mathbf{D}^{-1}, \ \mathbf{K}_{21} = -\mathbf{D}^{-1}\mathbf{A},$$
$$\mathbf{K}_{11} = \mathbf{B} - \mathbf{A}^H\mathbf{D}^{-1}\mathbf{A} \tag{26}$$

Equation (26) is analogous to structural mechanics problems. Hence, the numerical methods developed in computational structural mechanics can be applied. By utilizing PIM and W–W algorithm, all the pass-band eigenvalues can be obtained precisely.

5. COMPUTATIONAL METHOD FOR MATRIX DIFFERENTIAL EQUATIONS

The Computational method of matrix differential equations (24) will be discussed in this section.

5.1 The interval formulation and matrix differential equations

Let the two stations z_a and $z_b (z_b > z_a)$ required by the interval formulation of the precise integration method be selected arbitrarily within a layer to form the interval $[z_a, z_b]$. Evidently, if the vectors $\mathbf{q}_a$ and $\mathbf{p}_b$ are given at the two ends, respectively, the solution $\mathbf{q}$ and $\mathbf{p}$ in the interval $[z_a, z_b]$ is fixed. Thus the solution in the interval $[z_a, z_b]$ is totally determined by the two end boundary conditions, which can be written as [12]

$$\mathbf{q} = \mathbf{q}_a \quad \text{at } z = z_a; \quad \mathbf{p} = \mathbf{p}_b \quad \text{at } z = z_b \tag{27}$$

Because the system is linear, the relationship is

$$\mathbf{q}_b = \mathbf{F}\mathbf{q}_a - \mathbf{G}\mathbf{p}_b \quad \mathbf{p}_a = \mathbf{Q}\mathbf{q}_a + \mathbf{E}\mathbf{p}_b \tag{28}$$

where **F**, **Q**, **G**, **E** are complex matrices to be determined. Differentiating Equation (28) with respect to z and assuming $\mathbf{q}_a$, $\mathbf{p}_a$ are given yields

$$\begin{aligned} \dot{\mathbf{q}}_b &= \dot{\mathbf{F}}\mathbf{q}_a - \dot{\mathbf{G}}\mathbf{p}_b - \mathbf{G}\dot{\mathbf{p}}_b \\ \mathbf{0} &= \dot{\mathbf{Q}}\mathbf{q}_a + \dot{\mathbf{E}}\mathbf{p}_b + \mathbf{E}\dot{\mathbf{p}}_b \end{aligned} \tag{29}$$

The dual equation can then be written as

$$\dot{\mathbf{q}}_b = \mathbf{A}\mathbf{q}_b + \mathbf{D}\mathbf{p}_b \quad \dot{\mathbf{p}}_b = \mathbf{B}\mathbf{q}_b + \mathbf{C}\mathbf{p}_b \tag{30}$$

Combining Equation (29) and Equation (30) gives

$$\begin{aligned} &(\dot{\mathbf{F}} - \mathbf{AF} - \mathbf{GBF})\mathbf{q}_a + (-\dot{\mathbf{G}} - \mathbf{D} - \mathbf{GC} + \mathbf{AG} + \mathbf{GBG})\mathbf{p}_b = \mathbf{0} \\ &(\mathbf{EBF} + \dot{\mathbf{Q}})\mathbf{q}_a + (-\mathbf{EBG} + \dot{\mathbf{E}} + \mathbf{EC})\mathbf{p}_b = \mathbf{0} \end{aligned} \tag{31}$$

Noting that the vectors $\mathbf{q}_a$, $\mathbf{p}_b$ are mutually independent, yields the equations

$$\begin{aligned} &\dot{\mathbf{F}} = (\mathbf{A} + \mathbf{GB})\mathbf{F}, \quad \dot{\mathbf{E}} = \mathbf{E}(\mathbf{BG} - \mathbf{C}), \\ &\dot{\mathbf{G}} = \mathbf{AG} - \mathbf{GC} - \mathbf{D} + \mathbf{GBG}, \quad \dot{\mathbf{Q}} = -\mathbf{EBF} \end{aligned} \tag{32}$$

By going to the limit as $z_b \to z_a$ the boundary conditions for these equations are

$$\mathbf{G}(z_a, z_b) = \mathbf{Q}(z_a, z_b) = \mathbf{0}, \quad \mathbf{F}(z_a, z_b) = \mathbf{E}(z_a, z_b) = \mathbf{I}, \quad \text{when } z_a \to z_b \tag{33}$$

5.2 The combination of adjacent intervals

Two adjacent intervals $[z_a, z_b]$ and $[z_b, z_c]$, for which the interval matrices are $(\mathbf{F}_1, \mathbf{Q}_1, \mathbf{G}_1, \mathbf{E}_1)$ and $(\mathbf{F}_2, \mathbf{Q}_2, \mathbf{G}_2, \mathbf{E}_2)$, respectively, can be combined to form a longer interval $[z_a, z_c]$ with interval matrices $(\mathbf{F}_c, \mathbf{Q}_c, \mathbf{G}_c, \mathbf{E}_c)$.

Applying Equation (28) to intervals 1 and 2 gives

$$\mathbf{q}_b = \mathbf{F}_1\mathbf{q}_a - \mathbf{G}_1\mathbf{p}_b \ \text{for} [z_a, z_b] \tag{34a}$$

$$\mathbf{p}_a = \mathbf{Q}_1\mathbf{q}_a + \mathbf{E}_1\mathbf{p}_b \ \text{for} [z_a, z_b] \tag{34b}$$

$$\mathbf{q}_c = \mathbf{F}_2\mathbf{q}_b - \mathbf{G}_2\mathbf{p}_c \ \text{for } [z_b, z_c] \tag{35a}$$

$$\mathbf{p}_b = \mathbf{Q}_2\mathbf{q}_b + \mathbf{E}_2\mathbf{p}_c \ \text{for } [z_b, z_c] \tag{35b}$$

and the requirement is to find the equations of the combined interval in the form

$$\mathbf{q}_c = \mathbf{F}_c\mathbf{q}_a - \mathbf{G}_c\mathbf{p}_c \text{ for } [z_a, z_c] \tag{36a}$$

$$\mathbf{p}_a = \mathbf{Q}_c\mathbf{q}_a + \mathbf{E}_c\mathbf{p}_c \text{ for } [z_a, z_c] \tag{36b}$$

Solving for the vectors $\mathbf{q}_b$, $\mathbf{p}_b$ by using Equations (34a) and (35b) gives

$$\mathbf{q}_b = (\mathbf{I} + \mathbf{G}_1\mathbf{Q}_2)^{-1}\mathbf{F}_1\mathbf{q}_a - (\mathbf{G}_1^{-1} + \mathbf{Q}_2)^{-1}\mathbf{E}_2\mathbf{p}_c \tag{37a}$$

$$\mathbf{p}_b = (\mathbf{Q}_2^{-1} + \mathbf{G}_1)^{-1}\mathbf{F}_1\mathbf{q}_a + (\mathbf{I} + \mathbf{Q}_2\mathbf{G}_1)^{-1}\mathbf{E}_2\mathbf{p}_c \tag{37b}$$

Substituting these back into Equations (34b) and (35a), and then comparing with Equations (36b) and (36a), gives the desired interval combination equations as

$$\mathbf{F}_c = \mathbf{F}_2(\mathbf{I} + \mathbf{G}_1\mathbf{Q}_2)^{-1}\mathbf{F}_1 \tag{38a}$$

$$\mathbf{G}_c = \mathbf{G}_2 + \mathbf{F}_2(\mathbf{G}_1^{-1} + \mathbf{Q}_2)^{-1}\mathbf{E}_2 \tag{38b}$$

$$\mathbf{Q}_c = \mathbf{Q}_1 + \mathbf{E}_1(\mathbf{Q}_2^{-1} + \mathbf{G}_1)^{-1}\mathbf{F}_1 \tag{38c}$$

$$\mathbf{E}_c = \mathbf{E}_1(\mathbf{I} + \mathbf{Q}_2\mathbf{G}_1)^{-1}\mathbf{E}_2 \tag{38d}$$

These equations are important for eigenvalue problems or the solution of ordinary differential equations.

5.3 Initialization of interval matrices

Equations (38abcd) have shown how interval matrices operate, but so far no interval matrices have been derived, so that only the system matrices **A, B, C, D** are available. Therefore it is now necessary to generate a set of interval matrices from **A, B, C, D**.

The present case constitutes a system which is independent of the coordinate z within each layer of material, which can therefore be considered as an interval. It is then possible to subdivide this interval into any number of equal length subintervals. For convenience, the procedure adopted is a two step process. The layer thickness $Dz_i = (z_i - z_{i-1})$, can be first divided into 64 identical sublayers, so that each one has thickness $subDz_i = Dz_i/64$. This provides the points at which the elements of subsequent eigenvectors will be calculated if required. Each of these sublayers can then be further divided into 2^N sublayers, each with the extremely small thickness

$$\tau = subDz_i/2^N = subDz_i/1,048,576 \quad \text{for } N = 20$$

For this interval τ, the interval matrices $\mathbf{F}$, $\mathbf{Q}$, $\mathbf{G}$, $\mathbf{E}$ can be found as follows. Based on the differential Equation (32), and using the interval boundary conditions of Equation (33), the Taylor series expansion can be expressed as

$$\mathbf{Q}(\tau) = \boldsymbol{\theta}_1\tau + \boldsymbol{\theta}_2\tau^2 + \boldsymbol{\theta}_3\tau^3 + \boldsymbol{\theta}_4\tau^4 \tag{39a}$$

$$\mathbf{G}(\tau) = \boldsymbol{\gamma}_1\tau + \boldsymbol{\gamma}_2\tau^2 + \boldsymbol{\gamma}_3\tau^3 + \boldsymbol{\gamma}_4\tau^4 \tag{39b}$$

$$\mathbf{F}'(\tau) = \boldsymbol{\phi}_1\tau + \boldsymbol{\phi}_2\tau^2 + \boldsymbol{\phi}_3\tau^3 + \boldsymbol{\phi}_4\tau^4 \quad \mathbf{F}(\tau) = \mathbf{I} + \mathbf{F}'(\tau) \tag{39c}$$

$$\mathbf{E}'(\tau) = \boldsymbol{\psi}_1\tau + \boldsymbol{\psi}_2\tau^2 + \boldsymbol{\psi}_3\tau^3 + \boldsymbol{\psi}_4\tau^4 \quad \mathbf{E}(\tau) = \mathbf{I} + \mathbf{E}'(\tau) \tag{39d}$$

where $\boldsymbol{\theta}_i, \boldsymbol{\phi}_i, \boldsymbol{\psi}_i, \boldsymbol{\gamma}_i (i = 1, 2, 3, 4)$ are all 2×2 coefficient matrices yet to be determined. Substituting Equation (39) into Equation (32), and comparing coefficients of various powers of τ gives the following equations

$$\boldsymbol{\theta}_1 = -\mathbf{B} \quad \boldsymbol{\gamma}_1 = -\mathbf{D} \quad \boldsymbol{\phi}_1 = \mathbf{A} \quad \boldsymbol{\psi}_1 = -\mathbf{C} \tag{40}$$

$$\begin{aligned} \boldsymbol{\theta}_2 &= -(\boldsymbol{\psi}_1\mathbf{B} + \mathbf{B}\boldsymbol{\phi}_1)/2 \quad \boldsymbol{\gamma}_2 = (\mathbf{A}\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_1\mathbf{C})/2 \quad \boldsymbol{\phi}_2 = (\mathbf{A}\boldsymbol{\phi}_1 + \boldsymbol{\gamma}_1\mathbf{B})/2 \\ \boldsymbol{\psi}_2 &= (\mathbf{B}\boldsymbol{\gamma}_1 - \boldsymbol{\psi}_1\mathbf{C})/2 \quad \mathbf{a}_2 = (\mathbf{A}\mathbf{a}_1 + \boldsymbol{\gamma}_1\mathbf{s}_p)/2 \quad \mathbf{b}_2 = -(\mathbf{B}\mathbf{a}_1 + \boldsymbol{\psi}_1\mathbf{s}_p)/2 \end{aligned} \tag{41}$$

$$\begin{aligned} \boldsymbol{\theta}_3 &= -(\boldsymbol{\psi}_2\mathbf{B} + \mathbf{B}\boldsymbol{\phi}_2 + \boldsymbol{\psi}_1\mathbf{B}\boldsymbol{\phi}_1)/3 \quad \boldsymbol{\gamma}_3 = (\mathbf{A}\boldsymbol{\gamma}_2 - \boldsymbol{\gamma}_2\mathbf{C} + \boldsymbol{\gamma}_1\mathbf{B}\boldsymbol{\gamma}_1)/3 \\ \boldsymbol{\phi}_3 &= (\mathbf{A}\boldsymbol{\phi}_2 + \boldsymbol{\gamma}_2\mathbf{B} + \boldsymbol{\gamma}_1\mathbf{B}\boldsymbol{\phi}_1)/3 \quad \boldsymbol{\psi}_3 = (\mathbf{B}\boldsymbol{\gamma}_2 + \boldsymbol{\psi}_1\mathbf{B}\boldsymbol{\gamma}_1 - \boldsymbol{\psi}_2\mathbf{C})/3 \end{aligned} \tag{42}$$

$$\begin{aligned} \boldsymbol{\theta}_4 &= -(\boldsymbol{\psi}_3\mathbf{B} + \mathbf{B}\boldsymbol{\phi}_3 + \boldsymbol{\psi}_2\mathbf{B}\boldsymbol{\phi}_1 + \boldsymbol{\psi}_1\mathbf{B}\boldsymbol{\phi}_2)/4 \\ \boldsymbol{\gamma}_4 &= (\mathbf{A}\boldsymbol{\gamma}_3 - \boldsymbol{\gamma}_3\mathbf{C} + \boldsymbol{\gamma}_2\mathbf{B}\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_1\mathbf{B}\boldsymbol{\gamma}_2)/4 \\ \boldsymbol{\phi}_4 &= (\mathbf{A}\boldsymbol{\phi}_3 + \boldsymbol{\gamma}_3\mathbf{B} + \boldsymbol{\gamma}_2\mathbf{B}\boldsymbol{\phi}_1 + \boldsymbol{\gamma}_1\mathbf{B}\boldsymbol{\phi}_2)/4 \\ \boldsymbol{\psi}_4 &= (\mathbf{B}\boldsymbol{\gamma}_3 + \boldsymbol{\psi}_1\mathbf{B}\boldsymbol{\gamma}_2 + \boldsymbol{\psi}_2\mathbf{B}\boldsymbol{\gamma}_1 - \boldsymbol{\psi}_3\mathbf{C})/4 \end{aligned} \tag{43}$$

These matrices can be computed successively without iteration. However, the precision must be considered carefully as follows. The Taylor series expansions are truncated after the τ^4 terms, and the first term is of order τ. Therefore the relative order of the neglected terms is τ^4. However, if $N = 20\tau$ has been divided by rather more than 10^6. Thus τ^4 will be of the order of 10^{-24}, which is well beyond the double precision accuracy of the computer used.

All derivations in the previous sections are exact, in the sense that the sole approximation made, namely the truncation of the Taylor series expansion of Equation (39), has just been shown to cause a numerical error which is less than the round-off error of double precision computation. Therefore the method is exact in the sense that any method can be exact, i.e., it is as exact as the computer precision permits. However, because $\mathbf{F} = \mathbf{I} + \mathbf{F}'$ and $\mathbf{F}'$ is a matrix of small quantities, the addition must not be executed when the interval is very small, as otherwise unnecessary numerical errors would be induced and the exactness lost. Therefore it is vitally important that only $\mathbf{F}'$ is generated and stored in the computer memory, and never $\mathbf{F}$. Hence it is necessary to replace

Equation (38) by

$$\mathbf{F}_c' = (\mathbf{F}' - \mathbf{GQ}/2)(\mathbf{I} + \mathbf{GQ})^{-1} + (\mathbf{I} + \mathbf{GQ})^{-1}(\mathbf{F}' - \mathbf{GQ}/2) + \mathbf{F}'(\mathbf{I} + \mathbf{GQ})^{-1}\mathbf{F}' \tag{44a}$$

$$\mathbf{E}_c' = (\mathbf{F}' - \mathbf{GQ}/2)(\mathbf{I} + \mathbf{GQ})^{-1} + (\mathbf{I} + \mathbf{GQ})^{-1}(\mathbf{F}' - \mathbf{GQ}/2) + \mathbf{F}'(\mathbf{I} + \mathbf{GQ})^{-1}\mathbf{F}' \tag{44b}$$

$$\mathbf{G}_c = \mathbf{G} + (\mathbf{I} + \mathbf{F}')(\mathbf{G}^{-1} + \mathbf{Q})^{-1}(\mathbf{I} + \mathbf{E}') \tag{44c}$$

$$\mathbf{Q}_c = \mathbf{Q} + (\mathbf{I} + \mathbf{E}')(\mathbf{Q}^{-1} + \mathbf{G})^{-1}(\mathbf{I} + \mathbf{F}') \tag{44d}$$

5.4 The 2^N algorithm

After the generation of the interval matrices $\mathbf{E}(\tau)$, $\mathbf{F}(\tau)$, $\mathbf{Q}(\tau)$, $\mathbf{G}(\tau)$ of length τ, combination must be used to obtain the matrices $\mathbf{E}'(subDz_i)$, $\mathbf{F}'(subDz_i)$, $\mathbf{Q}(subDz_i)$, $\mathbf{G}(subDz_i)$, of the given interval $subDz_i$. There are $2^N (= 1,048,576$ for $N = 20)$, small τ intervals. Since all of these small intervals are identical, the interval combination of Equation (44) can be applied with $\mathbf{Q}_1 = \mathbf{Q}_2$,etc., and the combination of the 2^N identical intervals requires N 'doubling up' executions of Equation (44), as given by the instructions

```
{$\mathbf{E}'(\tau)$, $\mathbf{F}'(\tau)$, $\mathbf{Q}(\tau)$, $\mathbf{G}(\tau)$} generated by Equation (39)
for (itera = 0; itera < $N$; itera++) {
   { The execution of Equations (44abcdef); }
   $\mathbf{Q} = \mathbf{Q}_c$; $\mathbf{G} = \mathbf{G}_c$; $\mathbf{F}' = \mathbf{F}_c'$; $\mathbf{E}' = \mathbf{E}_c'$
}
$\mathbf{Q}(subDz_i) = \mathbf{Q}_c$; $\mathbf{G}(subDz_i) = \mathbf{G}_c$; $\mathbf{F}(subDz_i) = \mathbf{I} + \mathbf{F}'_c$;
   $\mathbf{E}(subDz_i) = \mathbf{I} + \mathbf{E}'_c$                                        (45)
```

The algorithm (45) gives the computation for a sublayer of thickness $subDz_i$. However, it is required further to find the layer interval matrices $\mathbf{Q}(Dz_i)$; $\mathbf{G}(Dz_i)$; $\mathbf{F}(Dz_i)$; $\mathbf{E}(Dz_i)$. The reason for setting up a sublayer of thickness $subDz_i$ is that if the thickness of say, the i-th layer, $Dz_i = z_i - z_{i-1}$, is too large, a wave could possibly occur within it. The division into 64 small sublayers ensures that internal waves are impossible, i.e., for each such sublayer its eigenvalue count $J_m(\omega) = 0$. Subsequent combinations of these layers involves keeping track of the eigenvalue count by using the equation

$$J_{mc}(\omega) = J_{m1}(\omega) + J_{m2}(\omega) - s\{G_1\} + s\{(G_1^{-1} + Q_2)\} \tag{46}$$

where $s\{\cdots\}$ is the sign count of the matrix within the brackets. Equation (46) is a necessary complement to Equation (38).

The following instructions give the computation of layer interval matrices $\mathbf{Q}(Dz_i)$; $\mathbf{G}(Dz_i)$; $\mathbf{F}(Dz_i)$; $\mathbf{E}(Dz_i)$

```
{$\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{Q}(subDz)$; $\mathbf{G}_1 = \mathbf{G}_2 = \mathbf{G}(subDz)$; $\mathbf{F}_1 = \mathbf{F}_2 = \mathbf{F}(subDz)$
 $\mathbf{E}_1 = \mathbf{E}_2 = \mathbf{E}(subDz)$; $J_{m1} = J_{m2} = 0$ }
for (itera = 0; itera < 6; itera++) {
   { The execution of Equations (38) and (46)}
   $\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{Q}_c; \mathbf{G}_1 = \mathbf{G}_2 = \mathbf{G}_c; \mathbf{F}_1 = \mathbf{F}_2 = \mathbf{F}_c$;
   $\mathbf{E}_1 = \mathbf{E}_2 = \mathbf{E}_c$; $J_{m1} = J_{m2} = J_{mc}$
}
{$\mathbf{Q}(Dz) = \mathbf{Q}_c; \mathbf{G}(Dz) = \mathbf{G}_c; \mathbf{F}(Dz) = \mathbf{F}_c$;
 $\mathbf{E}(Dz) = \mathbf{E}_c$; $J_m = J_{mc}$ }                                   (47)
```

The algorithm for combining the matrices of all of the intervals into one overall interval is similar. Let $z = 0$ be treated as end a, then the algorithm is

```
for ($layer = 1; layer \le layers; layer++$) {
   if ($layer == 1$){
      $\mathbf{Q}_c = \mathbf{Q}(layer); \mathbf{G}_c = \mathbf{G}(layer); \mathbf{F}_c = \mathbf{F}(layer)$;
      $\mathbf{E}_c = \mathbf{E}(layer)$; $J_{mc} = J_{m1}$;
   }
   else { $\mathbf{Q}_1 = \mathbf{Q}(layer); \mathbf{G}_1 = \mathbf{G}(layer); \mathbf{F}_1 = \mathbf{F}(layer)$;
      $\mathbf{E}_1 = \mathbf{E}(layer)$; $J_{m1} = J_{m1}$;
      $\mathbf{Q}_2 = \mathbf{Q}_c; \mathbf{G}_2 = \mathbf{G}_c; \mathbf{F}_2 = \mathbf{F}_c$;
      $\mathbf{E}_2 = \mathbf{E}_c$; $J_{m2} = J_{mc}$
   The execution of Equations (38) and (46);]
   }
}                                                              (48)
```

So far $\mathbf{Q}_c$, $\mathbf{G}_c$, $\mathbf{F}_c$, $\mathbf{E}_c$, J_{mc} are the overall interval matrices and EC, where EC stands for the eigenvalue count and J_{m1} means the EC for the layer. From the EC computation in combination with a binary search, one can find all of the eigenvalues without ever missing any.

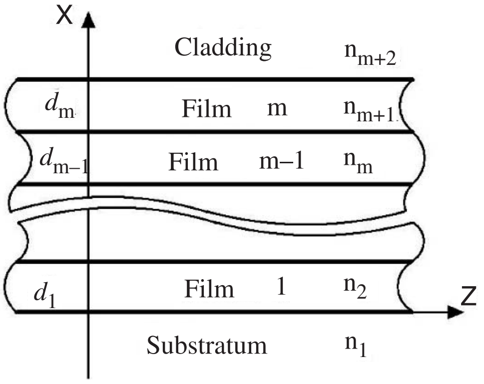


Figure 2. Multi-layered media.

6. NUMERICAL EXAMPLES

First, consider the case of isotropic plane waveguide: the thickness of the films is d, and the wave propagates in the z direction with propagation constant k_z, see Figure 2.

For pass-band:

$$\begin{aligned} k_z &> \frac{\max(n_1, n_{m+1}) \cdot \omega}{c}; \\ k_z &< \frac{\max(n_i) \cdot \omega}{c}; \quad i = 2, 3, \ldots, m \end{aligned} \tag{49}$$

In the simplest case: $m = 1, n_2 = 2.5, n_1 = n_3 = 1.5, d_1 = 600\,nm, k_z = 1 \times 10^7$. So the maximum of ω is $\omega_{MAX} = 2 \times 10^{15}$. All the pass-band eigenvalues can be obtained via W–W algorithm:

$$\begin{aligned} \omega_1 &= 1.279857264 \times 10^{15}, \\ \omega_2 &= 1.506912669 \times 10^{15}, \\ \omega_3 &= 1.834222574 \times 10^{15} \end{aligned}$$

The corresponding wave functions are shown in Figure 3:

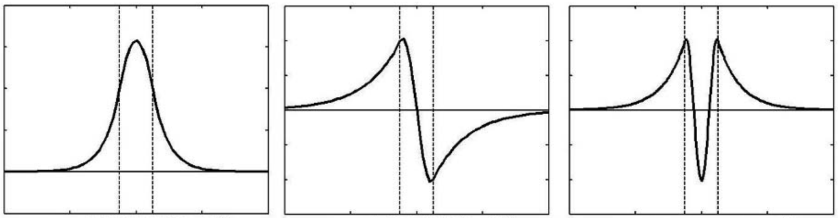
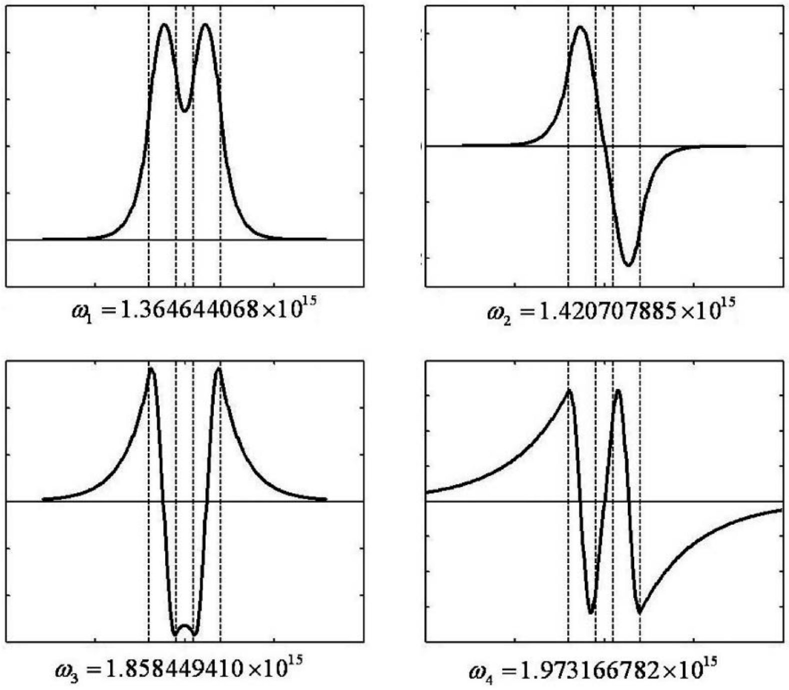

$\omega_1 = 1.279857264 \times 10^{15}$ $\omega_2 = 1.506912669 \times 10^{15}$ $\omega_3 = 1.834222574 \times 10^{15}$

Figure 3. Wave functions of single film.



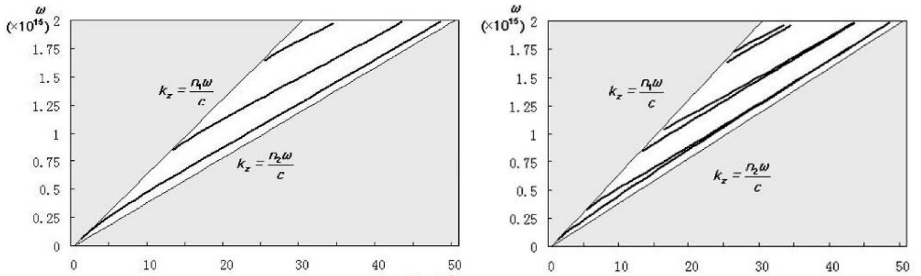
$\omega_1 = 1.364644068 \times 10^{15}$ $\omega_2 = 1.420707885 \times 10^{15}$

$\omega_3 = 1.858449410 \times 10^{15}$ $\omega_4 = 1.973166782 \times 10^{15}$

Figure 4. Wave functions of three films.

If $m = 3$, and take:, $n_2 = n_4 = 2.5$, $n_1 = n_3 = n_5 = 1.5$, $d_1 = 300\ nm$, $d_2 = 200\,nm, d_3 = 300\,nm, k_z = 1 \times 10^7$. The corresponding pass-band eigenvalues and wave functions are shown in Figure 4.

By using W–W algorithm the relationship between propagation constant k_z and pass-band eigenvalue ω can be obtained, and satisfies Equation (49). The relationship is shown in Figure 5. (The thickness of all the films is $d = 600\ nm$; for single layer: $n_1 = n_3 = 1.5,\ n_2 = 2.5$; for three layers: $n_1 = n_3 = n_5 = 1.5, n_2 = n_4 = 2.5$)



(a) Single layer (b) Three layers

Figure 5. Relationship between k_z and ω.

7. CONCLUSION

In present paper, the problem of optical waveguides in layered media is solved via symplectic analysis. A lot of numerical examples have been computed by using PIM and W–W algorithm, and highly precise results were given out. The merit of using the Hamiltonian system and symplectic geometry is that doing so enables the mathematical method of separation of variables, symplectic eigensolutions, expansion solution, etc., to be used to solve a wide range of problems.

REFERENCES

1. S. Ramo, J.R. Whinnery and Th. Van Duzer (1984), *Fields and Waves in Communication Electronics*, 2nd ed., J. Wiley & Sons, NY.
2. J.M. Jin (2002), *The Finite Element Method in Electromagnetics*. J. Wiley & Sons, New York.
3. P.C. Yeh (1988), *Optical Waves in Layered Media*. J. Wiley & Sons, New York.
4. W.X. Zhong (2001), Symplectic system of electro-magnetic wave-guide. *Journal of Dalian University of Techology*, 41, pp. 379–387 (in Chinese).
5. W. Zhong (2004), *Duality System in Applied Mechanics and Optimal Control*. Kluwer, Boston.
6. W.X. Zhong, H.J. Ouyang and Z.C. Deng (1993), *Computational Structural Mechanics and Optimal Control*. DalianUniversity Technology Press, Dalian (in Chinese).
7. W.H. Wittrick and F.W. Williams (1971), A general algorithm for computing natural frequencies of elastic structures. *Quartely Journal of Mechanical and Applied. Mathematics* , 24, pp. 263–84.
8. W. Zhong and F.W. Williams (1991), On the localization of the vibration mode of a substructural chain-type structure. In: *Proceedings of the Institute of Mechanical Engineering*, *Part C*, 205, pp. 281–8.
9. W.X. Zhong and F.W. Williams (1992), Wave propagation for repetitive structures and symplectic mathematics. In: *Proceedings of the Institute of Mechanical Engineering*, *Part C*, 206, pp. 371–9.
10. W.X. Zhong and F.W. Williams (1993), The eigensolutions of wave propagation for repetitive structures. *Structural Engineering and Mechanics*, 1, 1, pp. 47–60.
11. W.X. Zhong and F.W. Williams (1995), On the direct solution of wave propagation for repetitive structures. *Journal of Sound and Vibration*, 181, 3, pp. 485–501.
12. W.X. Zhong, W.P. Howson and F.W. Williams (2001), Precise solutions for surface wave propagation in stratified material. *Transactions of ASME: Journal of Vibration and Acoustics*, 123, pp. 198–204.

A COMPUTATIONAL ANALYSIS OF THERMAL RESIDUAL STRESS DURING MAGNETIC QUENCHING

Z.L. Li, H.M. Cheng, J.Y. Li, Y. Hu and L.J. Hou
Faculty of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming 650093, Yunnan, P. R. China

Abstract In this paper, based on the analysis of coupling action of temperature, phase transformation, magnetic field and stress, the phase transformation condition and equation related to material properties are discussed, a new constitutive equation considering effects of phase transformation and magnetic field is proposed and solved by means of Finite Element Method (F.E.M). The thermal residual stress is obtained and the influencing factors on the thermal stress of magnetic field are analysed and discussed.

Keywords: thermal stress, magnetic quenching, constitutive equation, F.E.M.

1. INTRODUCTION

Magnetic quenching is a new heat treatment method adding additional magnetic field during quenching process of ferromagnetic material, which can well improve the mechanics properties of material. However, compared with the experimental research, the theoretical study on the new heat treatment method is far behind the demands of actual uses. During the course of magnetic quenching, residual stress is a consequence of interactions among time, temperature, deformation, microstructure and additional magnetic field (Figure 1). All of these make the numerical solution more difficult, and research on the above is reported in this paper.

G. R. Liu et al. (eds.), Computational Methods, 1581–1586.

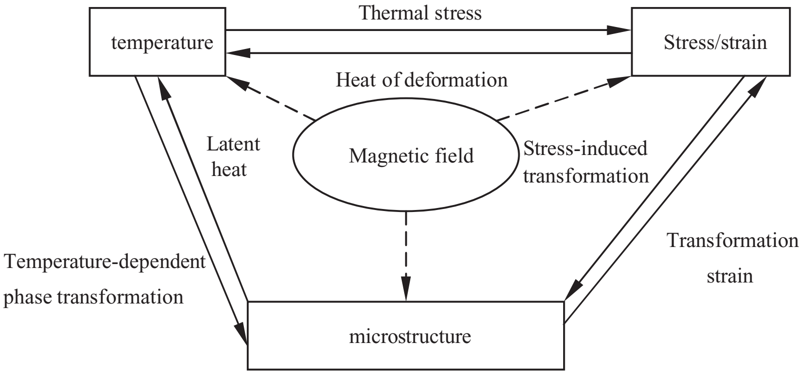


Figure 1. Phenomenological coupling during magnetic quenching.

2. THE INFLUENCE OF MAGNETIC FIELD ON THE PHYSICAL PROPERTIES OF FERROMAGNETIC

2.1 Condition of phase transformation and calculation of the phase composition

According to phase transformation theory, the volume fraction of diffusion phase transformation can be calculated as follows [1]:

$$f_J = 1 - \exp(-b_n t^{a_n}), \tag{1}$$

where f_J is the certain phase composition's volume fraction, $b_n(T)$ and $a_n(T)$ are two coefficients relevant to the temperature that can be obtained in terms of the isotherm transfer curve of the material. The transformation quantity from austenite to martensite, which is non-diffusion transformation, can be calculated by following formula [1]:

$$f_\mathrm{M} = 1 - \exp[-1.10 \times 10^{-2}(M_\mathrm{s} - T)], \tag{2}$$

where f_M is the volume fraction of Martensite, M_s is the critical temperature at which the martensite transformation starts and T is the temperature. Let $H(x)$ express a stepped function as follows:

$$H(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}. \tag{3}$$

And the condition of phase transformation can be written as:

$$F_{A\to M}(T, \dot{T}) = H(\dot{T} - V_c)H(M_\mathrm{s} - T). \tag{4}$$

2.2 Abrupt change of specific heat and heat expansion coefficient

It is the basic feature of ferromagnetic that there exists abrupt change of specific heat and heat expansion coefficient at Curie temperature. According to the theory of ferromagnetism [2], the calculating expressions of the magnetic material's specific heat $c_{\rho H}$ and heat expansion coefficient α_H can be written as:

$$c_{\rho H} = c_\rho - \frac{1}{2}\lambda \frac{d}{dT}[M(T)]^2, \quad \alpha_H = \alpha_I - \frac{1}{3}\left(\frac{\partial \omega}{\partial H}\right)_T \left(\frac{\partial H}{\partial T}\right)_I, \tag{5}$$

where c_ρ and α_I are the common specific heat and heat expansion coefficient, respectively,λ is the molecule field coefficient, ω and $M(T)$ stand for spontaneous magnetization and volume contract induced by magnetic field.

2.3 Magneto-elasticity effect

Generally, the Young's modulus of ferromagnetic material will appear change when it is magnetized by the magnetic field and may be expressed as [2]:

$$E_H = E_0 - \Delta E = E_0\left[1 - E_0\left(\frac{9\mu_0\lambda_s^2}{20\pi I_s^2}\right)\right], \tag{6}$$

where E_0 is the Young's modulus without magnetized, μ_0 is the initial magnetic permeability, I_s is saturated intensity of magnetization and λ_s is the saturated magnetostriction.

3. THE HEAT CONDUCTION GOVERNING EQUATION SET

In order to depict the course of heat transfer of the workpiece under the magnetic quenching more accurately, the following governing equation set [3] is used:

$$\left.\begin{array}{l} [k(T)T_{,i}]_{,i} = c_{\rho H}(T)\dot{T} - \sum\limits_{J=1}^{N} F_J(T, \dot{T})\dot{f}_J L_J + \dfrac{\partial H}{\partial h}\dot{h} \quad \text{in } \Omega \times (0, t] \\ -k(T)\dfrac{\partial T}{\partial n} = a_1\Delta T_w + a_2\Delta T_w^2 + \cdots + a_l\Delta T_w^l \quad \text{at } \partial\Omega \times (0, t] \\ T(x, 0) = T_0 \quad \text{in } \Omega \end{array}\right\}, \tag{7}$$

where the sign H is enthalpy, h is magnetic field strength, M is spontaneous magnetiz-ation, the sign T_0 means sample's initial temperature and $c_{\rho H}$ is the specific heat.

4. THERMAL ELASTOPLASTIC CONSTITUTIVE EQUATION DURING MAGNETIC QUENCHING

The Mises' yield criteria of isotropy strengthen model can be expressed as [4]:

$$\bar{\sigma} - F\left(\int d\bar{\varepsilon}_p, T\right) = 0, \tag{8}$$

where $\bar{\varepsilon}_p$ and $\bar{\sigma}$ represent equivalent plastic strain and equivalent stress, respectively. Differentiating the Equation (8), we can get:

$$\left(\frac{\partial\bar{\sigma}}{\partial\{\sigma\}}\right)^T d\sigma = \frac{\partial F}{\partial\bar{\varepsilon}_p} d\bar{\varepsilon}_p + \frac{\partial F}{\partial T} dT. \tag{9}$$

According to incremental theory, let $d\{\varepsilon\}$ stands for the increments form of the total strain, thus:

$$d\{\varepsilon\} = d\{\varepsilon\}_{\mathrm{e}} + d\{\varepsilon\}_{\mathrm{p}} + d\{\varepsilon\}_{\mathrm{T}} + d\{\varepsilon\}_{\mathrm{tr}} \tag{10}$$

In which, $d\{\varepsilon\}_{\mathrm{e}}$, $d\{\varepsilon\}_{\mathrm{p}}$, $d\{\varepsilon\}_{\mathrm{T}}$ and $d\{\varepsilon\}_{\mathrm{tr}}$ are the increments form of the elastic strain, plastic strain, temperature stain and the united expression of all phase transformation strain component, respectively.

According to the mobile criteria and generalized Hooke's law, we can get:

$$\begin{aligned} d\{\sigma\} &= [D]_H\, d\{\varepsilon\}_{\mathrm{e}} = [D]_H \left(d\{\varepsilon\} - d\{\varepsilon\}_p - d\{\varepsilon\}_T - d\{\varepsilon\}_{tr}\right) \\ &= [D]_H \left(d\{\varepsilon\} - \frac{\partial\bar{\sigma}}{\partial\{\sigma\}} d\bar{\varepsilon}_{\mathrm{p}} - d\{\varepsilon\}_{\mathrm{T}} - d\{\varepsilon\}_{\mathrm{tr}}\right). \end{aligned} \tag{11}$$

Through formulas (9) and (11), we can gain the increments form of thermal elastoplastic constitutive equation during magnetic quenching process as follows:

$$\begin{aligned} d\{\sigma\} = [D]_{\mathrm{e}pH} &\left\{ d\{\varepsilon\} - \left[\frac{d[D]_H^{-1}}{dT}\{\sigma\} + \{\alpha\}\right] dT - d\{\varepsilon\}_{\mathrm{tr}} \right\} - [D]_H \frac{\partial\bar{\sigma}}{\partial\{\sigma\}} \\ &+ \frac{[D]_H \dfrac{\partial\bar{\sigma}}{\partial\{\sigma\}} \dfrac{\partial F}{\partial T} dT}{\dfrac{\partial F}{\partial\bar{\varepsilon}_p} + \left\{\dfrac{\partial\bar{\sigma}}{\partial\{\sigma\}}\right\}^T [D]_H \dfrac{\partial\bar{\sigma}}{\partial\{\sigma\}}}, \end{aligned} \tag{12}$$

where $[D]_{\mathrm{e}pH}$ is the thermal elastoplastic matrix during magnetic quenching process. Equation (12) is a nonlinear equation, it should be solved by means of Finite Element Method, and the Finite Element Method equation of Equation set (12) is established as follows:

$$[K]\{\delta\} = \{Q\}, \tag{13}$$

in which, $[K]$ and $\{\delta\}$ are the rigidity matrix and the displacement matrix, respectively, $\{Q\}$ is the heat load vector.

5. EXAMPLE AND DISCUSSION

The metallic workpiece is shown in Figure 2. It is a 45 steel cylinder of $\Phi 30 \times 60$ mm. The quenching initial temperature is 860°C and the magnetic field intensity is 4.4×10^4 A/m. It is continuously cooled in water whose initial temperature is 21°C.

Based on the known transient temperature distribution of a 45 steel cylinder workpiece during magnetic quenching [3], the finite element Equation (13) is solved and the calculated results of thermal residual stress distribution on Section I and Section II are obtained in Figure 3.

Figure 3 shows that during magnetic quenching, σ_z, σ_r, σ_θ all are smaller than the normal quenching process. It indicates that using magnetic quenching can decrease the internal stress of the sample due to the following reasons: first, additional magnetic field may reduce the cooling velocity of workpiece, and lighten it's temperature difference of surface and inner, so the residual

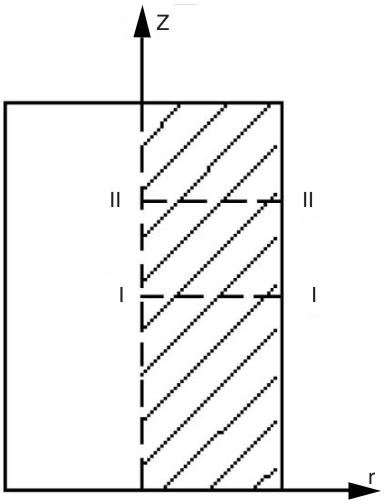
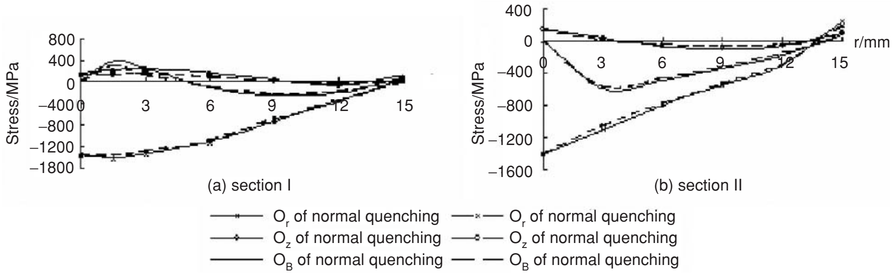


Figure 2. Sample.



(a) section I

(b) section II

Figure 3. Thermal residual stress distribution on Section I and Section II.

thermal stress is small. Second, the heat expansion coefficient of magnetized material take place abrupt change at Curie temperature, which is about 750°C and influence the process of phase transformation, then lead to the less thermal stress. Third, under the effect of additional magnetic field, the Young's modulus of 45 steel get small and make the elastic matrix and the thermal elastoplastic matrix changed and result in the less thermal stress.

ACKNOWLEDGEMENT

This work was supported by the National Science Foundation of China under grant No. 10162002 and the Scientific Research Foundation of Kunming University of Science and Technology under grant No. 2003-27.

REFERENCES

1. Z.Y. Xu (1988), *The Principle of Phase Transformation.* Science Press, Beijing, China.
2. Y.C. Guo (1965), *Ferromagnetism*. Higher Education Press, Beijing, China.
3. M.H. Liu, J.R. Chen and Z.L. Li (2001), A research on heat conduction of heat-magnetic inducing phase transformation in quenching process by finite element method. *Proceedings EPMESC' VIII*, Shanghai, China.
4. B.C. Liu (1990), *Engineering Computational Mechanics*. Mechanical Industry Press, Beijing, China.

PARALLEL FEM ANALYSIS OF HIGH FREQUENCY ELECTROMAGNETIC WAVE IN AN ENVIRONMENT

A. Takei[1], S. Yoshimura[1], B.H. Dennis[2] and H. Kanayama[3]
[1]*Institute of Environmental Studies, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8586, Japan*
[2]*Mechanical and Aerospace Engineering, The University of Texas at Arlington, Box 19018 Arlington, TX 76019-0018 USA*
[3]*Department of Intelligent Machinery and Systems, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan*

Abstract This paper presents a parallel finite element analysis of high frequency electromagnetic wave in an environment. The HDDM (Hierarchical Domain Decomposition Method) is employed as a parallel solver. Nedelec element is employed. The simulation method is tested for a simple model, and the results are compared with exact solutions.

Keywords: ADVENTURE system, HDDM, high frequency electromagnetic wave.

1. INTRODUCTION

In order to estimate high frequency electromagnetic wave density in an environment, we are developing a parallel finite element method (FEM) to solve a high frequency electromagnetic wave field [1]. The present authors have been involved in the ADVENTURE project since 1997 in which a computational mechanics system (ADVENTURE system) for large-scale analysis and design is being developed. The system consists of pre-, main- and post-processing modules and design modules that can be used in various kinds of parallel environment [2]. The ADVENTURE_Magnetic is one of the modules in the ADVENTURE system. The module contains the solver for magnetostatic problems. In this paper, we modify the module so as to solve a high frequency

G. R. Liu et al. (eds.), Computational Methods, 1587–1594.

electromagnetic wave field. We adopt the Hierarchical Domain Decomposition Method (HDDM) for parallel computing [3–4].

2. OUTLINE OF THE HDDM

In the DDM, an analysis model i.e. a FE mesh is subdivided into a number of subdomains. Yagawa and Shioya (1993) proposed a hierarchical technique to implement the DDM on various parallel computers. This technique is called the HDDM. In this particular method, a group of processing elements (PEs) are subdivided into three groups: one Grand Parent PE (Grand), some Parent PEs (Parent) and a number of Child PEs (Child or Children). Figure 1 shows the schematic data flow among PEs.

An analysis model is subdivided into some 'parts' whose number is the same as that of Parents. Then, each part is further subdivided into a number of subdomains. Each Parent stores in its memory a set of the part data that consists of some sets of subdomain data. The subdomain data include coordinates of nodes, material properties and information of the interface between subdomains. In the nonlinear analysis of magnetic fields, magnetic fields are also stored in the Parents. Children are dynamically allocated to Parents, and the data associated with a subdomain stored in Parents are sent to the Children. Each Child receives the data of a subdomain, conducts the computation associated with the subdomain and sends the results back to the Parents. Grand manages the computation; for example, check of the status of Parents and Children and

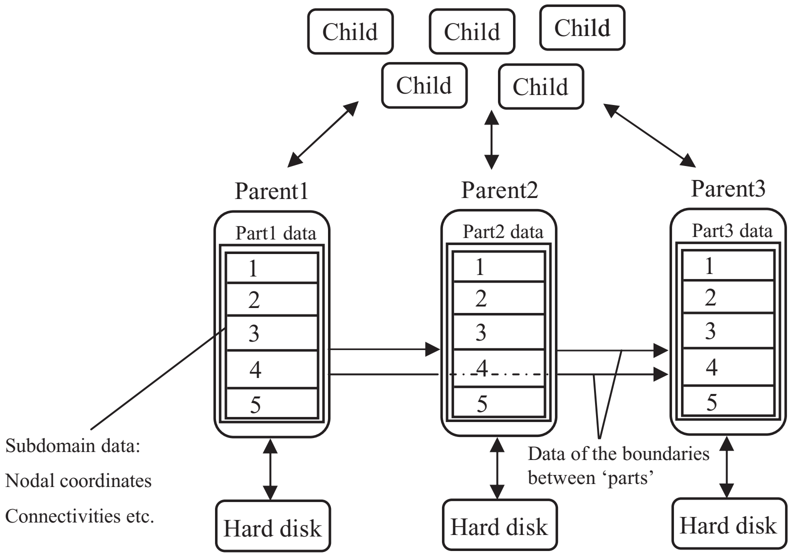


Figure 1. Schematic data flow among PEs in the HDDM.

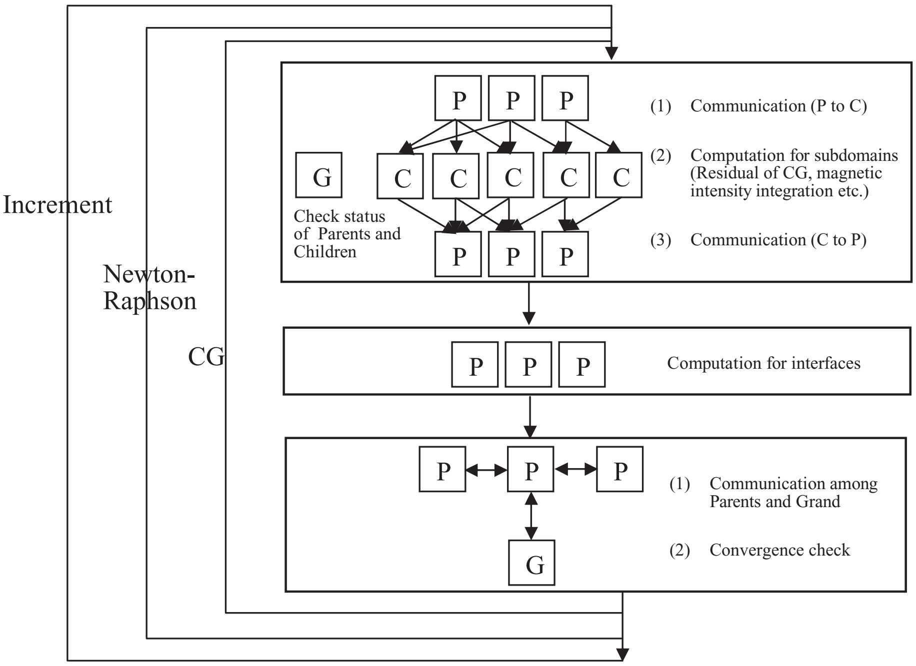


Figure 2. Schematic flow of nonlinear finite element analysis of magnetic fields on the HDDM.

assignment of Children to Parents and judgment of convergence of iterative methods. Figure 2 shows the schematic flow of the computation. The features of the HDDM are summarized as follows:

1. The number of Children can be arbitrary chosen if it is smaller than the number of subdomains. Therefore, robust operations can be done i.e. a user can change the number of PEs without changing that of subdomains.
2. Dynamic workload balancing can automatically be achieved. Therefore, the HDDM is suitable even in heterogeneous parallel environments.
3. Any large-scale analysis data can be handled by increasing the number of Parents.
4. Disk I/Os are conducted only by Parents. A large number of Children need not to directly access hard disks [5].

3. FORMULATION

3.1 Helmholtz wave equation

We consider a wave problem and adopt the magnetic vector potential $\mathbf{A}(\mathbf{x}, t)$ [Wb/m] as an unknown function. The Helmholtz electromagnetic wave equation with a source term is solved. The equation is given as follows:

$$\nabla \times \nabla \times \mathbf{A} + \mu\varepsilon\omega^2\mathbf{A} = -\mu\mathbf{J} \tag{1}$$

Let us consider the Maxwell equation in a frequency domain as:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon} \tag{2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{3}$$

$$\nabla \cdot \mathbf{H} = 0 \tag{4}$$

$$\nabla \times \mathbf{H} = -\frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \quad \text{or} \quad \nabla \times \mathbf{B} = -\mu\varepsilon\frac{\partial \mathbf{E}}{\partial t} + \mathbf{J} \tag{5}$$

In Equations (1)–(5), $\mathbf{E}(\mathbf{x}, \mathrm{t})$ and $\mathbf{H}(\mathbf{x}, \mathrm{t})$ denote field vectors of electric intensity [V/m] and magnetic intensity [A/m], respectively. $\mathbf{D}(\mathbf{x}, \mathrm{t})$ and $\mathbf{B}(\mathbf{x}, \mathrm{t})$ denote electric flux density [C/m^2] and magnetic flux density [Wb/m^2], respectively. ρ and $\mathbf{J}$ denote charge density [C/m^3] and current vector intensity [A/m^2], respectively. ε and μ are permittivity and permeability, respectively. Generally, electromagnetic wave in an environment is solved assuming that the problem becomes stationary. Equations (3) and (5) become Equations (6) and (7), taking the time dependence term of $e^{j\omega t}$. j and ω denote imaginary and angular frequency, respectively.

$$\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H} \tag{6}$$

$$\nabla \times \mathbf{H} = -j\omega\varepsilon\mathbf{E} + \mathbf{J} \tag{7}$$

Equation (9) is derived from Equations (6) and (8). Equation (9) does not contain an eddy current field, and one can get Equation (10) by a scalar potential Φ from Equation (9).

$$\mathbf{B} = \mu\mathbf{H} = \nabla \times \mathbf{A} \tag{8}$$

$$\nabla \times (\mathbf{E} + j\omega\mathbf{A}) = 0 \tag{9}$$

Equation (11) is obtained from Equations (6), (8) and (10), and one can set the second term of the right hand side of Equation (11) to be 0. Then, Equation (11) becomes Helmholtz wave equation given in Equation (1).

$$\mathbf{E} + j\omega\mathbf{A} = -\nabla\Phi \tag{10}$$

$$\left(\Delta + \mu\varepsilon\omega^2\right)\mathbf{A} = -\mu\mathbf{J} + \nabla\left(\nabla \cdot \mathbf{A} + j\omega\mu\Phi\right) \tag{11}$$

$$\nabla \cdot \mathbf{A} + j\omega\mu\Phi = 0 \tag{12}$$

3.2 Weak form

Let us consider the weak form of Equation (1). First order Nedelec element is employed to attain free spurious solutions. Equation (13) is rewritten in the form of to Equation (14). In this study, the perfect reflection boundary condition is employed. The condition is given by Equation (15) [6].

$$F(\mathbf{A}) = \int_{\Omega} \{(\nabla \times \mathbf{A}) \cdot (\nabla \times \mathbf{A}) + \mu\varepsilon\omega^2 \mathbf{A} \cdot \mathbf{A} - \mu \mathbf{A} \cdot \mathbf{J} + \tau(\nabla \cdot \mathbf{A})(\nabla \cdot \mathbf{A})\} ds = 0 \tag{13}$$

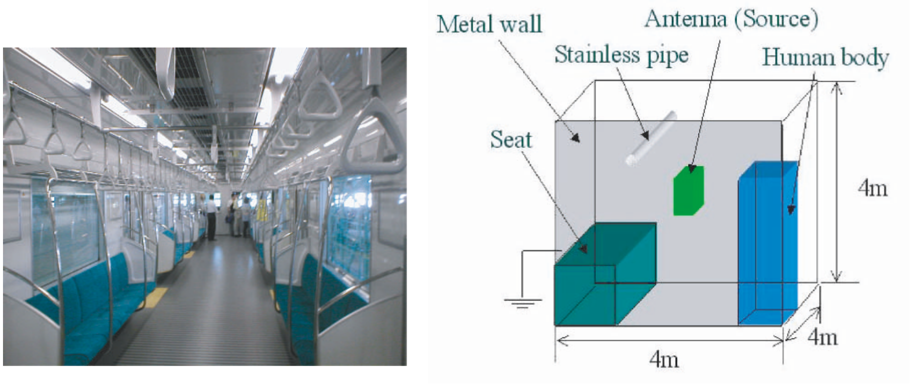
$$\big[[K] + \mu\varepsilon\omega^2[M] + \tau[N]\big]\{\mathbf{A}\} \equiv [D]\{\mathbf{A}\} = -\{\mathbf{J}\} \tag{14}$$

$$\mathbf{A} \times \mathbf{n} = 0 \tag{15}$$

4. ANALYSIS OF THE TEST MODEL

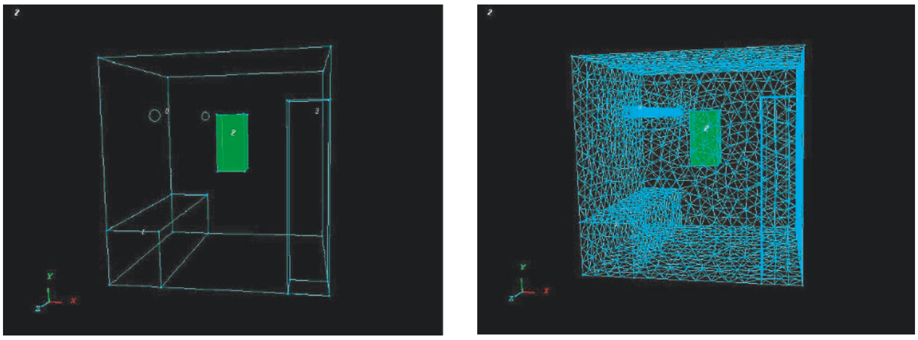
Solving the system equation of Equation (14), input vector i.e. current **J** produces magnetic vector potential in a space. Figure 3b shows a sketch of the test model. The model is simplification of a passenger car shown in Figure 3a. Every wall is grounded. Stainless steel pipe and antenna are made of metal. Seat and human body are assumed employing material parameters of plastic and water, respectively. The space is full of air. Figures 4a and b show the wire frame and the mesh of the input CAD data, respectively.

Calculating condition is shown in Table 1. Figure 5 shows one cycle of the wave of input current J. Its frequency is 500 MHz.



(a) In a passenger car (b) Sketch of the CAD model

Figure 3. The model for test calculation. (a) In a passenger car. (b) Sketch of the CAD model.



(a) Wire frame (b) Mesh

Figure 4. Input data. (a) Wire frame (b) Mesh.

Figure 6 shows the result. Figures 6a and b show the computed magnetic flux density contour and vector on X-Y plane, respectively. Figure 6c shows the magnetic flux vector of full layer. Figure 7 shows the convergence history of the CG iteration. As shown in Figure 7, good convergence has been obtained. The magnetic flux is radiated around the antenna i.e. source **J**. We can obtain that the flux does not permeate the seat and human body perfectly. Obviously, the simulation code can calculate physically correct solution in the model.

5. CONCLUSIONS

In this paper, we have described a parallel finite element method (FEM) to solve a high frequency electromagnetic wave field. We have developed it, modifying the ADVENTURE_Magnetic module. We have shown that the simulation code can calculate the solution in the model. The magnetic flux is radiated around the antenna. The flux does not permeate the seat and human body perfectly. These results are simulated physical phenomenon correctly.

Table 1. Calculation conditions.

Calculation model	Three dimension full layer
Element	First order tetrahedral Nedelec
Number of elements/DOFs	18599/4036*3
Machine spec	Pentium4 2.4GHz 4 parallel.
Calculation time	4.5 [s]

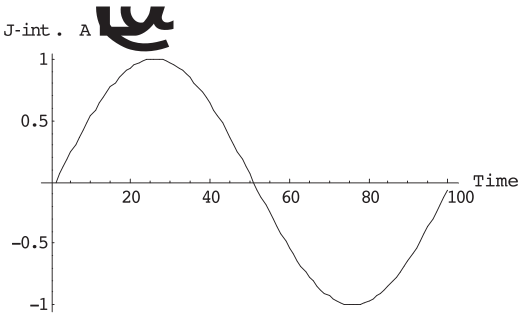
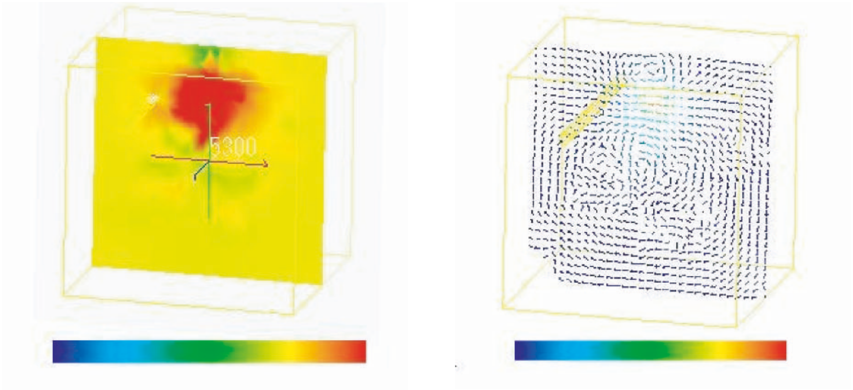
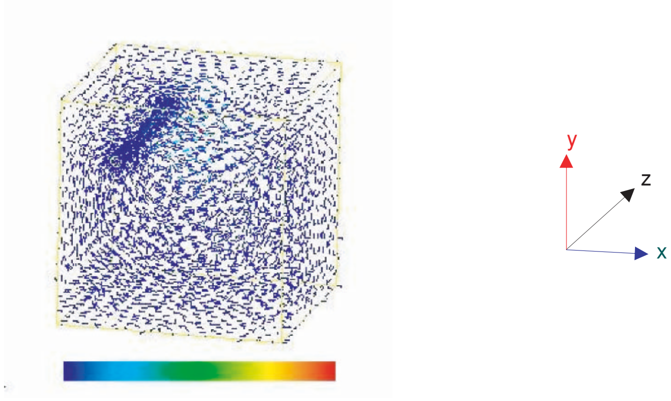


Figure 5. Input current J.



(a) Computed magnetic flux density contour

(b) Magnetic flux density vector on X-Y plane



(c) Vector indication (3D full layer)

Figure 6. Results (Magnetic flux density distributions). (a) Computed magnetic flux density contour. (b) Magnetic flux density vector on X-Y plane. (c) Vector indication (3D full layer).

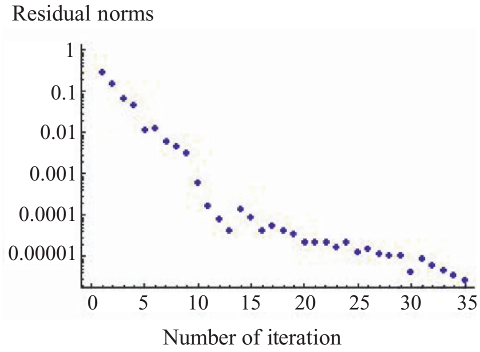


Figure 7. Convergence history of the CG iteration.

REFERENCES

1. A. Takei, S. Hayano and Y. Saito (2000), *IEEE Transactions on Magnetics*, 36, 64, pp. 1031–1034.
2. http://adventure.q.t.u-tokyo.ac.jp
3. H. Kanayama and F. Kikuchi (1999), *INFORMATION*, 2, 1, pp. 37–45.
4. H. Kanayama, D. Tagami, M. Saito and F. Kikuchi (2000), *Transactions of JSCES*, Paper No. 20000033, pp. 201–206.
5. T. Miyamura, H. Noguchi, R. Shioya, S. Yoshimura, G. Yagawa (2002), *Nuclear Engineering and Design*, 212, pp. 335–355.
6. J.D. Jackson (1998), *Classical Electrodynamics*, 3rd ed. John Wiley & Sons Inc., pp. 237–242.

DIFFUSION OF SINGLE Cu AND Ta ATOMS IN SILK-LIKE AMORPHOUS POLYMER

Dai Ling[1], S.W. Yang[2], V.B.C. Tan[1], P. Wu[2]
[1]*Department of Mechanical Engineering, National University of Singapore, 9 Engineering Drive 1, Singapore 117576*
[2]*Institute of High Performance Computing, 1 Science Park Road, Singapore 117528*

Abstract The continuous miniaturization of electronic products has led to new concerns over the resistance of interconnections between transistors. As transistors become smaller and consequently faster, delays in interconnects start to account for a considerable proportion of processing time. One solution is to replace the conventional Al/SiO$_2$ interconnects with Cu/silk interconnects. In order to curb the diffusion of Cu into silk, it has been proposed to coat the Cu with Ta. The suitability of Ta as a diffusion barrier is still not fully established. Ab initio MD simulations were carried out to characterize the diffusion of Cu and Ta atoms inside silk-like polymer structures. The results showed that Cu is much more active in the silk-like polymer than Ta. An analysis of the speed of the Cu atom in the silk polymer showed that the Cu atom experienced impulsive interatomic loads. Such impulses are postulated to cause hoppings, which were highlighted in reports of previous classical MD simulations. In contrast, the motion of Ta atom was relatively stable and can be considered as harmonic oscillations around an equilibrium position. The 'inert' characteristic of Ta suggests it is a good material to serve as a barrier between copper and silk layers to prevent the leakage of Cu into silk.

Keywords: ab initio MD simulation, diffusion coefficient, copper, tantalum, silk.

1. INTRODUCTION

An understanding of the mechanisms involved in the diffusion of copper and tantalum atoms into the silk is important to support the viability using Cu/Ta/Silk structures to replace Al/SiO$_2$ material systems that are widely used now. Although molecular level computations of diffusion phenomena of various material systems have been carried out successfully, reports on the diffusion

G. R. Liu et al. (eds.), Computational Methods, 1595–1599.

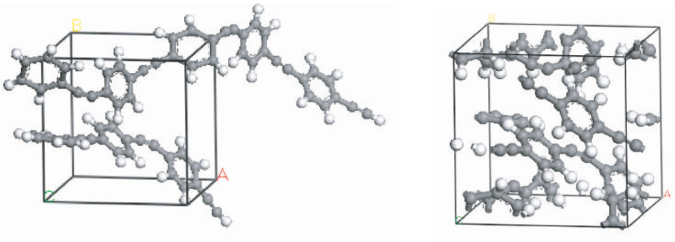
of metal atoms inside the polymer structures are lacking. Current technology can now produce Cu/Ta/Silk layers of nanometer length scale. At such scales, the understanding of interactions among single molecules and atoms is necessary for analysing diffusion phenomena. First-principle ab initio methods are used to study the most basic electronic characteristics of atoms and molecules. They are built on functions that can give reasonable approximate solution to the Schrödinger equation. Therefore, molecular dynamics (MD) simulations can be performed using energy relationships obtained from the ab initio algorithms. In this paper, ab initio MD simulations were carried out to characterize the diffusion of single Cu and Ta atoms in the silk-like amorphous polymer structure.

2. METHODOLOGY

The diffusion model comprised a block of amorphous polymers and single atom of Cu or Ta. The amorphous polymer was built from chains of aromatic hydrocarbon structure comprising about 100 atoms (Figure 1). After building the silk-like amorphous polymer model, a metal atom was placed at the centre of the cell. The ab initio MD simulations were carried out with the CPMD [1] software on an IBM690 parallel cluster. A wave function optimization was then carried out to quench the electrons to the ground-state, after which, ab initio MD simulations were carried out at a temperature of 300 K to study the motion of the metal atoms. By recording the trajectory of the penetrants, the diffusion coefficient, D, can be calculated from,

$$D = \frac{1}{2FN} \lim_{t \to \infty} \frac{d}{dt} \left\langle \sum_{i=1}^{N} [r_i(t) - r_i(0)]^2 \right\rangle \tag{1}$$

Here, F represents the number of dimensions of freedom; N is the number of penetrants; $r_i(t)$ is the position of penetrant i at time t. The angular brackets denote the mean square displacement of penetrants [2].



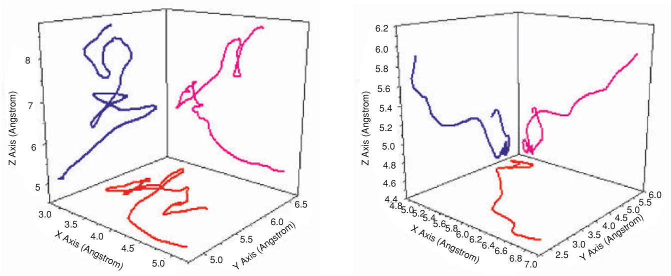
(a) (b)

Figure 1. Model of silk-like amorphous polymer. (a) Original view; (b) In-cell view.

3. RESULTS AND DISCUSSION

The single metal atom, Cu or Ta, was placed at the centre of the amorphous cell as the starting point of the diffusion process. Figure 2 shows the planar projections of the penetrants' diffusion trajectories recorded by the simulations. To further characterize the motions of the metal atom, the total diffusion time was divided into equal intervals, and the average speed of the penetrant at each interval was calculated. Histories of speed of diffusion for the metal atoms are shown in Figure 3. The speed of the Cu atom averages to 0.8 km/s but demonstrates some large sharp peaks in motions. This suggests that the diffusion of Cu in silk can be considered as major wanderings plus some impulsive motions due to interactions from neighbouring elements. The impulsive motion is perceived to contribute to most of the diffusion mechanism. When the impulses are large enough to release the penetrant from the current neighbouring elements, hopping, as observed in classical MD simulations [3], will take place.

In contrast, the Ta atom has an average speed of around 0.5 km/s, and demonstrates less fluctuation in speed. This means that Ta atoms are unlikely



(a) (b)

Figure 2. Planar projections of diffusion trajectories for (a) Cu and (b) Ta.

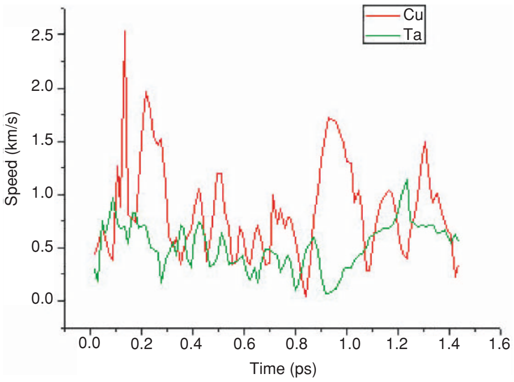


Figure 3. Time histories of speed of diffusion Cu and Ta atoms.

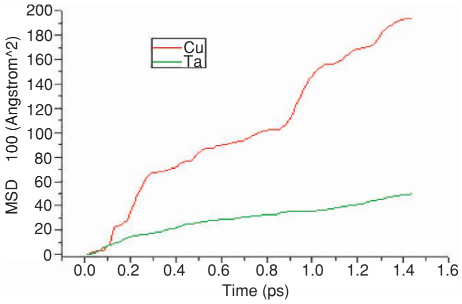


Figure 4. History of mean square displacements of Cu and Ta atoms in silk.

to be excited by the silk-like polymer elements. When the Ta atom is travelling within a reasonably mean dense medium, its trajectory can be considered to be mainly harmonic oscillations around equilibrium position. Thus, the Ta atoms are unlikely to deviate sufficiently far away from their original locations to cause diffusion.

Apart from the speed of the metal atoms inside the amorphous silk polymer, the mean square displacement (MSD) of the atoms, i.e., an average of squared distances summed up over all possible positions from the origin, is also a useful indication of the diffusivity of the metal atoms in the polymer. Figure 4 shows the MSD for the Cu and Ta atoms.

The MSD plot clearly shows that the Ta atom has a much flatter and more linear MSD curve than that of the Cu atom. Comparatively, Cu is quite active inside the silk polymers. The steps in the MSD plot of the Cu atoms occur at the same instances as the large peaks in speed of Figure 3. That the steps in displacements account for more than 50% of the final MSD supports the suggestion that diffusion is dominated by the impulses imparted on the Cu atoms.

With the MSD data, the diffusion coefficient can be calculated from Equation (1). A plot of the diffusion coefficient from Equation (1) against time is shown in Figure 5. The plateau in the plots gives the converged values of the

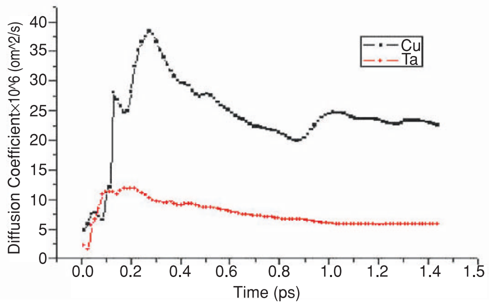


Figure 5. Evolution of diffusion coefficients of Cu and Ta in silk.

diffusion coefficients. Figure 5 shows that the diffusion coefficient of Ta starts to converge at quite early stages while the coefficient for Cu requires a longer duration to converge because of the large impulses. The converged values of diffusion coefficients are estimated to be $2.4 \times 10^{-5}\text{cm}^2/\text{s}$ for Cu and $0.6 \times 10^{-6}\text{cm}^2/\text{s}$ for Ta. The diffusion coefficient of Ta is comparable to those of other small molecule penetrants [2–4] while Cu has a coefficient of around three factors higher.

4. CONCLUSIONS

Preliminary work to simulate the diffusion of metal atoms inside polymers was undertaken to study the viability of Cu/Ta/silk structures for semiconductor applications. Ab initio MD simulations have been carried out to characterize the diffusion of single Cu and Ta atoms inside silk-like amorphous polymer structures. The results showed that Cu is more active in the polymer medium than Ta. Computations based on the simulations gave a diffusion coefficient for Cu which is four times that of Ta. An analysis of the speed of a Cu atom during the diffusion process showed that the atom is subjected to large impulses, which are likely to be the cause of hoppings reported in previous classical MD simulations. The motion of Ta, on the other hand, was quite stable. The Ta atom did not demonstrate the excitations observed for the Cu atom, making Ta a candidate barrier material to prevent the leakage of Cu into silk in electronic applications.

REFERENCES

1. CPMD, Copyright IBM corp 1990–2004, Copyright MPI for Festkorperforschung Stuttgart 1997–2001.
2. T.L. Li, O.K. Dane, P. Kinam (1997), Computer simulation of molecular diffusion in amorphous polymers. *Journal of Controlled Release*, 48, p. 57.
3. S.G. Charati and S.A. Stern (1998), Diffusion of gases in silicone polymers: molecular dynamics simulations. *Macromolecules*, 31, p. 5529
4. R.K. Bharadwaj and R.H. Boyd (1999), Small molecule penetrant diffusion in aromatic polyesters: a molecular dynamics simulation study. *Polymer*, 40, p. 4229.

MULTISCALE COUPLING OF MESHLESS METHOD AND MOLECULAR DYNAMICS

H.Y. Liang[1], G.R. Liu[1] and X. Han[2]
[1]*Centre for ACES, Department of Mechanical Engineering, National University of Singapore, 9 Engineering Drive 1, Singapore 117576, Singapore*
[2]*College of Mechanical and Automotive Engineering, Hunan University, Changsha 410082, P. R. China*

Abstract A multiscale modelling procedure, combining meshless method with molecular dynamics is developed in this paper. An intermediate oscillator is introduced to act as a media for the energy transfer between atom and continuum domains. Very smooth energy transfer is observed in our calculations for both 1D and 2D examples.

Keywords: multiscale, meshless, atomistic, intermediate oscillator.

1. INTRODUCTION

Most physical phenomena in nature involve a hierarchy of both spatial and temporal scales at different levels. Typical cases include protein folding, chemical reaction, turbulence, crack propagation in solid, or shear localization. To solve such kind of problems is often beyond the capability of one theoretical frame valid within a single scale, such as molecular dynamics and continuum mechanics.

Efforts seeking for multiscale methodologies spanning from atomistic to continuum domains can be traced back to the work by Sinclair [1]. Mullins *et al.* [2] used finite element method to model the continuum domain slight away from the region near the crack tip where atomistic calculation is performed. Tadmor *et al.* [3] developed the quasicontinuum (QC) method with capability to remesh according to the variation of the deformation gradient. Rudd and Broughton [4] formulated a coarse-grained molecular dynamics (CGMD) method, derived directly from finite temperature MD through a statistical coarse graining

G. R. Liu et al. (eds.), Computational Methods, 1601–1606.

procedure. Xiao and Belytschko [5] proposed a bridging domain method with linear combination of the continuum and atomic Hamiltonians at the overlapping domain, avoiding spurious wave reflections. Till now, most of the multiscale methods belong to handshaking approach. Recently, Cai *et al.* [6] and Park *et al.* [7] introduced a general Langevin equation as a boundary to atom region, also to eliminate wave reflections.

In this paper, motivated by Xiaoand Belytschko [5], we study wave propagation with a modified multiscale modelling method, coupling meshless method with molecular dynamics. Very smooth energy transfer between atom and continuum domains is achieved.

2. COUPLING MODEL

The coupled model (Figure 1) includes atomistic and continuum domains, and an intermediate oscillator. The oscillator acts as a media for the energy transfer between atomistic and continuum domains. The total Hamiltonian of the extended system is a combination of the Hamiltonians of the atomistic and continuum domains, and the intermediate oscillator,

$$\begin{aligned} H = {} & \sum_{i}^{\text{atom}} \frac{\mathbf{p}_i^{\mathrm{T}} \mathbf{p}_i}{2(1-\alpha)\, m_i} + (1-\alpha)\, V + \sum_{I}^{\text{node}} \frac{\mathbf{p}_I^{\mathrm{T}} \mathbf{p}_I}{2\alpha m_I} + \alpha W \\ & + \sum_{\Theta}^{Media} \left(\frac{\mathbf{p}_{g\Theta}^{\mathrm{T}} \mathbf{p}_{g\Theta}}{2Q} + \boldsymbol{\lambda}_{\Theta}^{\mathrm{T}} \mathbf{g}_{\Theta} + \frac{K}{2} \mathbf{g}_{\Theta}^{\mathrm{T}} \mathbf{g}_{\Theta} \right) \end{aligned} \tag{1}$$

where i stands for the atoms, I for meshless node and Θ for the interpolation point in the overlapping region. $\alpha(\in [0, 1])$ is the scale parameter, allowing for a grade energy transfer between atomistic and continuum domains. In practice, α takes the form of arctangent function. Q, K, $\boldsymbol{\lambda}$, $\mathbf{p}_{g\Theta}$ and $\mathbf{g}_{\Theta}$ are the generalized mass, stiffness, Lagrange force, momentum and coordinate of the intermediate oscillator respectively. $\mathbf{p}_{g\Theta}$ and $\mathbf{g}_{\Theta}$ are obtained by the following,

$$\mathbf{g}_{\Theta} = \sum_{I} N_{\Theta I} \mathbf{u}_I - \sum_{i} N_{\Theta i} \mathbf{u}_i \quad \mathbf{p}_{g\Theta} = Q \left(\sum_{I} \frac{N_{\Theta I}}{\alpha m_I} \mathbf{p}_I - \sum_{i} \frac{N_{\Theta i}}{(1-\alpha)\, m_i} \mathbf{p}_i \right) \tag{2}$$

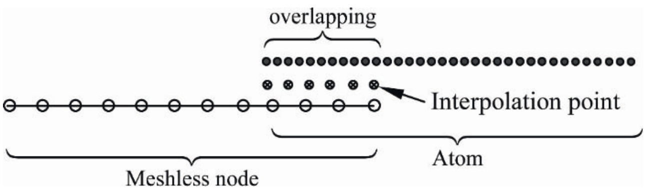


Figure 1. Coupling model.

At each point Θ, there is a ‘particle’ with mass Q (a small value for quick energy transfer in calculations), subjected to a constant force (penalty method) and variable Lagrange force (Lagrange multiplier method). $N_{\Theta I}$ and $N_{\Theta i}$ are shape functions of node I and atom i, respectively, evaluated at point Θ. The shape function is constructed via MLS approach [8]. The equations of motion, derived from Hamiltonian canonical equations, can be integrated with velocity Verlet integrator, using multiple-time step algorithm:

$$\begin{aligned}
\mathbf{u}_{I(n+1)} &= \mathbf{u}_{I(n)} + \dot{\mathbf{u}}_{I(n)}\Delta T + \frac{1}{2}\frac{\mathbf{f}_{I(n)}}{m_I}\Delta T^2 \\
\mathbf{u}_{i\left(n+\frac{j+1}{N}\right)} &= \mathbf{u}_{i\left(n+\frac{j}{N}\right)} + \dot{\mathbf{u}}_{i\left(n+\frac{j}{N}\right)}\Delta t + \frac{1}{2}\frac{\mathbf{f}_{i\left(n+\frac{j}{N}\right)}}{m_i}\Delta t^2 \\
\mathbf{f}_{I\left(n+\frac{j+1}{N}\right)} &= \mathbf{f}^{\mathrm{int}}_{I(n+1)} + \left(\mathbf{f}^{\lambda}_{I\left(n+\frac{j+1}{N}\right)} + \mathbf{f}^{g}_{I\left(n+\frac{j+1}{N}\right)}\right)\Big/\alpha \\
\mathbf{f}_{i\left(n+\frac{j+1}{N}\right)} &= \mathbf{f}^{\mathrm{int}}_{i\left(n+\frac{j+1}{N}\right)} + \left(\mathbf{f}^{\lambda}_{i\left(n+\frac{j+1}{N}\right)} + \mathbf{f}^{g}_{i\left(n+\frac{j+1}{N}\right)}\right)\Big/(1-\alpha) \qquad (3) \\
\dot{\mathbf{u}}_{I\left(n+\frac{j+1}{N}\right)} &= \dot{\mathbf{u}}_{I\left(n+\frac{j}{N}\right)} + \frac{\Delta t}{2m_I}\left(\mathbf{f}_{I\left(n+\frac{j}{N}\right)} + \mathbf{f}_{I\left(n+\frac{j+1}{N}\right)}\right) \\
\dot{\mathbf{u}}_{i\left(n+\frac{j+1}{N}\right)} &= \dot{\mathbf{u}}_{i\left(n+\frac{j}{N}\right)} + \frac{\Delta t}{2m_i}\left(\mathbf{f}_{i\left(n+\frac{j}{N}\right)} + \mathbf{f}_{i\left(n+\frac{j+1}{N}\right)}\right)
\end{aligned}$$

where, $j = 0, 1, 2\cdots, N-1. \Delta T$ is the time step for continuum, Δt for atomistic region. $\mathbf{f}^{\mathrm{int}}_I$,$\mathbf{f}^{\mathrm{int}}_i$ are the internal force for continuum and atom, respectively. The generalized forces are

$$\begin{aligned}
\mathbf{f}^{\lambda}_I &= -\sum_{\Theta}\lambda_{\Theta}N_{\Theta I} \quad & \mathbf{f}^{g}_I &= -K\left(\sum_{\Theta}\sum_{J}N_{\Theta I}N_{\Theta J}\mathbf{u}_J - \sum_{\Theta}\sum_{j}N_{\Theta I}N_{\Theta j}\mathbf{u}_j\right) \\
\mathbf{f}^{\lambda}_j &= \sum_{\Theta}\lambda_{\Theta}N_{\Theta j} \quad & \mathbf{f}^{g}_i &= K\left(\sum_{\Theta}\sum_{J}N_{\Theta i}N_{\Theta J}\mathbf{u}_J - \sum_{\Theta}\sum_{j}N_{\Theta i}N_{\Theta j}\mathbf{u}_j\right)
\end{aligned} \qquad (4)$$

3. RESULTS AND DISCUSSION

In the 1D and 2D models, nearest neighbour (NN) interaction in atom domain is represented with harmonic potential. The elastic properties of continuum domain are equivalent to that of atomistic domain. The units are reduced. The longitudinal wave propagation in 1D model is shown in Figure 2. The

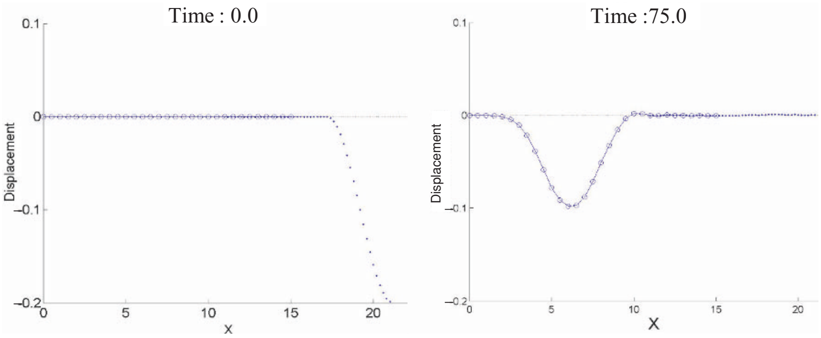


Figure 2. A wave propagates in 1D domain with overlapping area.

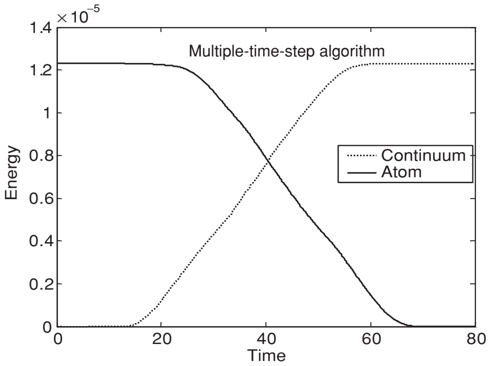


Figure 3. Time history of energy transfer in 1D domain.

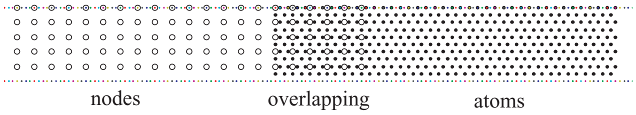


Figure 4. The 2D model.

distances between the NN atoms and NN nodes are 0.2 and 0.5, respectively. There are 30 nodes and 75 atoms in 1D model. The two ends are traction free. An initial displacement of one-quarter of sinusoid is applied as in Figure 2 (time = 0.0). The time steps for atoms and nodes are 0.01 and 0.05, respectively. Figure 3 shows a smooth energy (sum of kinetic and potential energy) transfer from atom to continuum. Figure 4 is a 2D slab model, with 105 nodes and 400 atoms. The distances between the NN atoms and NN nodes are 0.2 and 0.4, respectively. Periodic boundary condition is applied along the vertical direction and the two horizontal ends are traction free. The initial displacement is similar to that in 1D model. Figure 5 demonstrates the plane wave

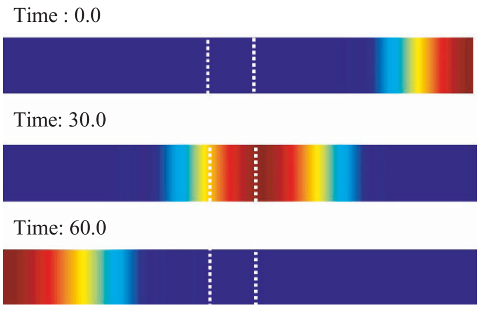


Figure 5. A plane wave propagates in a 2D domain with overlapping area.

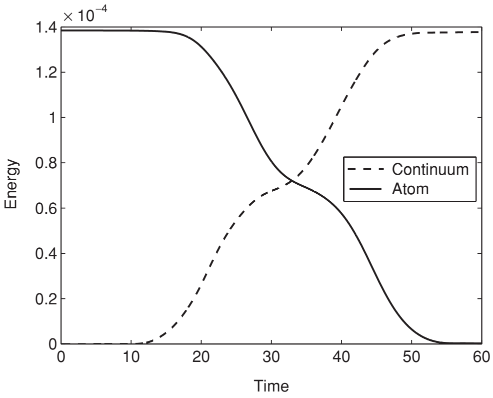


Figure 6. Time history of energy transfer in 2D domain. (plane wave)

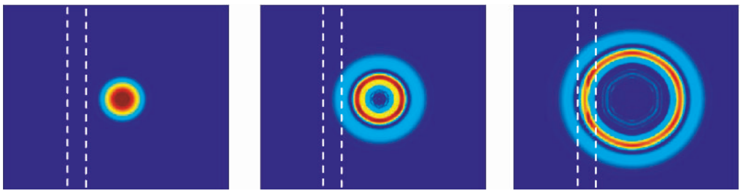


Figure 7. A circular wave propagates in 2D domain with overlapping area.

propagation in the slab model. Again, energy transfers smoothly between the two domains (Figure 6). Figure 7 exhibits the propagation of a circle wave in a square model. The overlapping zone locates between the dashed lines, the left to which is continuum domain, and the right to which is atom domain. The wave crosses the bridging domain in a good manner, with minor wave reflection observed. For comparison, Park et al. [7] adopted generalized Langevin equation as the boundary condition for atom domain, where the refection wave is not neglectable.

4. CONCLUSIONS

A multiscale modelling method, combining meshless method with molecular dynamics, is developed in this paper, in which an intermediate oscillator is introduced acts as a media for the energy transfer between atom and continuum domains. 1D and 2D models are studied with our approach. Smooth energy transfer is observed in our calculations.

ACKNOWLEDGEMENT

This work is supported by the National Science Foundation of China under the grant number 10372031.

REFERENCES

1. J.E. Sinclair (1971), Improved atomistic model of a bcc dislocation core. *Journal of Applied Physics*, 41, pp. 5321–5329.
2. M. Mullins and M.A. Dokainish (1982), Simulation of the (001) plane crack in α-iron employing a new boundary scheme. *Philosophical Magazine A*, 46, pp. 771–787.
3. E.B. Tadmor, M. Ortiz and R. Phillips (1996), Quasicontinuum analysis of defects in solids. *Philosophical Magazine A*, 73, pp. 1529–1563.
4. R.E. Rudd and J.Q. Broughton (1998), Coarse-grained molecular dynamics and the atomic limit of finite elements. *Physical Review B*, 58, pp. 5893–5896.
5. S.P. Xiao and T. Belytschko (2004), A bridging domain method for coupling continua with molecular dynamics. *Computer Methods in Applied Mechanics and Engineering*, 193, pp. 1645–1669.
6. W. Cai, M.de. Koning, V. Bulatov and S. Yip (2000), Minimizing boundary reflections in coupled-domain simulations. *Physical Review Letter*, 85, pp. 3213–3216.
7. H.S. Park, E.G. Karpov, P.A. Klein and W.K. Liu (2004), The bridging scale for two-dimensional atomistic/continuum coupling. *Philosophical Magazine*, in press.
8. T. Belytschko, Y.Y. Lu and L. Gu (1994), Element free Galerkin methods. *International Journal for Numerical Methods in Engineering*, 37, pp. 229–256.

MECHANISMS OF DISINTEGRATION OF MINERAL MEDIA EXPOSED TO HIGH-POWER ELECTROMAGNETIC PULSES

V.A. Chanturiya, I.J. Bunin and A.T. Kovalev
Research Institute of Comprehensive Exploitation of Mineral Resources (IPKON RAS), Russian Academy of Sciences, 4 Kryukovsky Tupik, Moscow 111020, Russia

Abstract This report is devoted to the following plausible mechanisms of disintegration of mineral substances exposed to high-power electromagnetic pulses (HPEMP) with high electric field strength: loosening of mineral structure through electrical breakdown; disintegration due to development of mechanical stresses at the boundary between the dielectric and conductive mineral components; electromagnetic energy absorption by thin metallic films or layers thinner than the characteristic skin layer. Disintegration through these mechanisms would proceed efficiently only provided that the size of the mineral sample exposed to HPEMP exceeds a definite minimum value, which is due to low concentration of the irradiating energy.

Keywords: mineral media, high-power electromagnetic pulses, disintegration.

1. INTRODUCTION

The application of HPEMP irradiation in dressing of resistant gold-containing ores appears attractive as this technique provides for a significant increase in precious metal recovery (30–80% for gold and 20–50% for silver), therewith helping reduce both energy consumption and the cost of products, as shown in previous works by Chanturiya et al. [1–3]. However, the diversified physical processes involved in the interaction of HPEMP with complex mineral media are as yet scantily known.

This study deals with three plausible mechanisms of disintegration of mineral particles under the action of nanosecond HPEMP with high electric field strength $E_p \sim 10^7$ V/m. The first mechanism consists in loosening of the mineral structure due to electrical breakdown effects, which only occurs in cases

G. R. Liu et al. (eds.), Computational Methods, 1607–1614.

where small, highly conductive inclusions are hosted in dielectric media. The second mechanism is related to development of thermomechanical stresses at the boundary (interface) between the dielectric and conductive mineral components, being only realized in cases where these components are comparable in size. The third mechanism, assuming essentially non-thermal action of HPEMP on mineral complexes, is related to electromagnetic energy absorption by thin metallic films or layers much thinner than the characteristic skin layer.

2. MODELING OF DISINTEGRATION PROCESSES IN MINERAL COMPLEXES EXPOSED TO HIGH-POWER PULSE IRRADIATION

2.1 Disintegration of mineral media due to electrical breakdown

The action of HPEMP on a sample with conductive inclusions is accompanied by electric field strengthening in the vicinity of the boundary with maximum curvature. As a result, electrical breakdown occurs at lower E_p values. The as-formed breakdown channels connect the conductive inclusions to each other and to the surface of the sample.

In this report we present experimental data, according to which the boundaries of Au micro-inclusions in FeAsS (arsenopyrite) grains are initially (prior to irradiation) out of contact with joints in the host mineral matrix. Originally the joints in arsenopyrite grains have a spallation (cleavage) appearance, most of them being filled with some dust-like material. After HPEMP irradiation, the cracks prove to be empty of filling, and appear discontinuous. Most of these cracks are confined to boundaries between secondary mineral phase segregations and the arsenopyrite matrix, which suggests that HPEMP have a selective effect on the host mineral matrix.

The average per pulse amount of energy released in the developing breakdown channel is estimated as $w \approx W_0 t_p / t_0$, where $W_0 = \int_V d^3x(\varepsilon \overrightarrow{E}^2/8\pi)$ is electric field energy of the pulse in a particle of volume V, t_p is pulse duration, and t_0 is the time during which the channel develops a length equal to the linear dimension of the particle (L). In essence, t_0 depends on the rate of channel development v_d, which is governed by the drift velocity of electrons in the avalanche at the head cusp of the developing channel, and depends on local field strength as $\propto \sqrt{E}$. For dielectrics with electrical strength about 50 MV/m, $v_d \cong$ 3–5 km/s, such that more than one pulse with $t_p = 30$ ns would be required to break down through a particle with $L > 1$ mm. Once the developing channel reaches the surface of the particle, the discharge becomes creeping, tending to neutralize the electric field throughout the entire particle volume. The total energy released in the channel is limited by the W_0 value (assuming that the particle under consideration is isolated from all other particles).

The size of the zone occupied by cracks (r_c) depends on energy density in the breakdown channel. To calculate r_c, we used the following model considerations from a previous work by Chanturiya et al. [4]. Energy is released in the channel somewhat instantaneously as compared to the time it takes for the channel to expand, stress waves to propagate through the sample and the heated matter to flow out from the channel. The zone occupied by evaporated matter takes on the form of explosion cavity. Beyond the zone of evaporation, the material undergoes crushing, and in farther-away zones cracking due to azimuthal (radial) rupture stresses predominates. While a problem of cylindrical shock wave propagation is being solved in the zones of evaporation and crushing, for the cracking zone we use a quasi-static approximation based on equilibrium equations. The relationship of r_c to initial gas pressure in the channel (p_0) can be approximately expressed as follows:

$$r_c = a_0 \frac{\sigma_* \kappa}{\sigma_c} \left(\frac{p_0}{\sigma_* \sqrt{\kappa}} \right)^{1/2\gamma}, \quad \kappa^2 \equiv \frac{E}{2\sigma_*} \left[1 + \nu + \ln \frac{\sigma_*}{\sigma_c} \right]^{-1}, \tag{1}$$

where a_0 is the initial channel radius, γ is an adiabatic index for the gas in the channel, σ_* and σ_c denote the compression (crushing) and rupture strength parameters of the material exposed to HPEMP, E is Young's modulus and ν is Poisson's ratio.

For quartz samples with $L = 0.3$, 1 and 10 mm and HPEMP with $E_p = 20$ MV/m, we obtained $r_c = 0.005$, 0.08 and 0.5 mm, respectively. In addition to cracks developing in an irradiated sample, there also forms a cavity remaining after matter outflow from the channel, and some voids arise due to opening of the cracks, these latter totalling about 50% of the hollow channel volume. Energy density per unit of channel length depends on particle volume V as $\propto V^{2/3}$, owing to which mechanical stresses around the channel give rise to cracking only in particles larger than $L_{\min}$ varying from 10^{-4} to 10^{-3} m, depending on E_p. This is why this mechanism would only work for relatively large particles subject to disintegration, its efficiency increasing with increasing L and E_p.

2.2 Mechanism of disintegration of mineral aggregates subjected to pulse-heating

Disintegration of mineral particles through irradiation-induced cracking is possible for particles consisting of substances with differing conductivity, when subjected to pulse-heating. The rate of HPEMP energy absorption by the matter is determined as $\eta E_*^2/\rho c_p$, where η and E_* are local electrical conductivity and field strength, respectively, ρ denotes density of particular component of the complex substance exposed to heating pulses and c_p is heat capacity of this material component. Intensity of thermomechanical stresses is governed by heating contrast and dissimilarity of thermoelastic properties of the different material components.

In the simplest case with axial symmetry, where two cylindrical parts (1 and 2) are in contact in the $z = 0$ plane, the quantities $u(r, z)$ and $s(r, z)$, respectively characterizing displacement in the r and z directions, and the $\sigma_{\alpha\beta}(r, z)$ values characterizing stresses within each of the contacting parts are determined by the following equations of motion:

$$\frac{\partial \sigma_{rr}}{\partial r} + \frac{\partial \sigma_{rz}}{\partial z} + \frac{\sigma_{rr} - \sigma_{zz}}{r} = \rho \frac{\partial^2 u}{\partial t^2}, \tag{2}$$

$$\frac{\partial \sigma_{zz}}{\partial z} + \frac{\partial \sigma_{rz}}{\partial r} + \frac{\sigma_{zr}}{r} = \rho \frac{\partial^2 s}{\partial t^2} \tag{3}$$

with the following boundary conjugation (continuity) conditions:

$$u^{(1)}(r, z = 0) = u^{(2)}(r, z = 0).$$

The components of the deformation tensor

$$e_{rr} = \partial u/\partial r, \quad e_{zz} = \partial s/\partial z, \quad e_{\varphi\phi} = u/r, \quad e_{rz} = \tfrac{1}{2}(\partial u/\partial z + \partial s/\partial r) \tag{4}$$

are related to $\sigma_{\alpha\beta}(r, z)$ as follows:

$$\sigma_{\alpha\beta} = 2Ge_{\alpha\beta} + (\lambda e - \beta\theta)\delta_{\alpha\beta}, \quad \sigma_{r\varphi} = \sigma_{z\varphi} = 0, \quad e_{r\varphi} = e_{z\varphi} = 0, \tag{5}$$

where $\theta = \Delta T$ denotes the amount of heating of part 1 (or part 2), and G, λ and β are related to Young's modulus E, Poisson's ratio ν and thermal expansion coefficient α by the following expression:

$$G = E/2(\nu + 1), \quad \lambda = E\nu/(1 + \nu)(1 - 2\nu) \quad \text{and} \quad \beta = \alpha E/(1 - 2\nu).$$

Shear stresses at the boundary between parts 1 and 2 can be determined approximately from the condition of equal displacement at the boundary, assuming that in the vicinity of the interface $u(r, z)$ can be considered independent of r and $s(r, z)$ independent of z, neglecting the acceleration and disregarding temperature gradient within each of the contacting parts:

$$\tau = \frac{\beta_2\theta_2(2G_1 + \lambda_1) - \beta_1\theta_1(2G_2 + \lambda_2)}{2(G_1 + G_2) + \lambda_2 + \lambda_1} \cong \tilde{E} \cdot \frac{3}{2}(\alpha_2\theta_2 - \alpha_1\theta_1), \tag{6}$$

where $\tilde{E} = 2E_1E_2/(E_1 + E_2)$. Stress values are maximum in cases where $\alpha_2 > \alpha_1, \theta_2 > \theta_1$ or $\alpha_2 < \alpha_1, \theta_2 < \theta_1$, (e.g., in the FeS–$Al_2O_3$ complex FeS heats up stronger than Al_2O_3); in such situation, $\tau \sim 30 \times 10^{-6}\tilde{E}\theta_2$. Owing to heat conduction, θ_2 at the boundary becomes nearly equal to θ_1 within $\sim 10^{-8}$ s of the pulse, such that only stresses related to dissimilarity of thermoelastic properties remain significant.

A single pulse with $E \sim 10^7$ V/m and $t_p \sim (3\text{–}10) \times 10^{-9}$ s generates stresses of 0.3–3 MPa at pyrite-gold or pyrite-quartz boundaries in complex samples, which would be insufficient for spallation along an ideal (perfect) boundary between the different components. However, in the presence of initial damages, small as they are, failure can build up in a sample irradiated with a series of pulses. Under repeated loading, a small initiating crack of length $l_0 < L$ develops up to $l = \sum_n^N \Delta l_n$, where $\Delta l_n = \int_0^{t_n} (dl/dt)dt$ is crack length increment due to the nth pulse, and dl/dt is the rate of crack development. This mechanism of disintegration would work efficiently if $l \cong L$.

When estimating dl/dt, we proceeded from the assumption that the disintegration thus induced is essentially of quasi-brittle failure type. On this basis, the current dl/dt value can be determined from given stress $\sigma(t)$ and its related critical length l_c, assuming the fulfilment of the condition of crack development $l > l_c(\sigma(t))$, $\sigma(t) > \sigma_c$:

$$dl/dt = v_{\max}^{cr}\sqrt{(1 - l_c/l)\{1 - (2\sigma_c/\sigma - 1)l_c/l\}}, \tag{7}$$

where $v_{\max}^{cr} \cong (0.4/0.6)\sqrt{E/\rho}$ is maximum crack development rate. The condition of reaching the critical stress value σ_c depending on initial crack length l_0 was taken as a cracking start criterion.

For stress varying between 0.3 and 3 MPa, l_0 proves to be about $(0.2\text{–}2) \times 10^{-3}$ m. Since the linear dimension of the particle $L > l_0$, only particles with $L > (0.2\text{–}2) \times 10^{-3}$ m would be expected to disintegrate. For example, when dealing with pyrite-quartz or pyrite-gold contacts, 15–50 pulses with $E_p \sim 10^7$ V/m and $t_p = (3\text{–}10) \times 10^{-9}$ s suffice to break down a sample of size 3×10^{-3}m, and 100–300 pulses with the same parameters would be required to disintegrate a sample of size 10^{-2} m.

2.3 Non-thermal absorption of electromagnetic energy by precious metal particles

The above described mechanisms of electrical breakdown and pulse-heating do not exhaust the multitude of processes that occur in natural mineral media exposed to high-power electromagnetic pulses. In particular, one of the plausible mechanisms of non-thermal action of HPEMP on mineral complexes is related to the effect of electromagnetic energy absorption by thin metallic films or layers thinner than the characteristic skin layer. This model appears quite realistic as the size of disseminated precious metal particles is very small (0.01–0.1 μm), which may be significantly smaller than the characteristic wavelength of the irradiating pulse and the corresponding skin layer thickness.

Theoretical analysis of this disintegration mechanism was performed using a simple 1-D model (developed in collaboration with IRE RAS, Moscow), which consider an electromagnetic pulse striking a flat metallic layer of thickness h enclosed within a dielectric medium. For $h >> \delta$ ($\delta = c/\sqrt{2\pi\sigma\omega}$ is skin layer thickness for frequency ω, c is light speed in vacuum and σ is electrical conductivity of the metal), reflection coefficient R for a flat electromagnetic wave reflected from a thick metallic layer is given by the Hagen–Rubens formula:

$$R = 1 - \sqrt{2\omega/\pi\sigma} = 1 - 2k\delta. \tag{8}$$

For a gold layer, δ values in the UHF and radio frequency ranges (e.g., $\lambda = 3$ cm and 1 m) are equal to 0.72 and 4.17 μm, respectively, and the corresponding reflection coefficient values are 0.99973 and 0.999956, which means that the electromagnetic radiation is almost totally reflected, and absorption may be considered nil.

For $h \ll \delta$, reflection coefficient may be far from unit, and absorption coefficient values may range up to 50% and even more. In this case, the values of reflection, transmission and absorption coefficient for the electromagnetic energy of HPEMP are determined completely by electrical conductivity σ of the thin metallic layer, which is related to conductivity of the same metal in a massive sample σ_0 by the following formula:

$$\sigma(d) = \sigma_0(d/2l_0)(1 + \ln(l_0/d)), \tag{9}$$

where d is layer thickness obeying the condition $d < l_0$ (l_0 is average electron free path length in a sample of infinite thickness). This particular case is suitable for describing the interaction of HPEMP with precious metal particles disseminated in resistant ores, because for particles of size 10^{-9} to 10^{-7} m skin layer thickness would be about $\sim 10^{-6}$ m. The above described 1-D model gives us reason to state that interaction of high-power electromagnetic radiation with precious metal particles hosted in resistant ores may involve highly efficient non-thermal absorption of electromagnetic energy by metal particles, which, at irradiation power high enough, may result in fast energy release and cause certain changes in host mineral structure. In particular, field strength within the rock (ore) samples may be increased significantly, and micro-breakdown features may develop between metal particles, which would result in micro-cracking to promote disintegration of mineral complexes.

3. CONCLUSIONS

A feature in common to the disintegration mechanisms described above is the limitation on the minimum sample size related to low concentration of the irradiating energy. Increasing this concentration by increasing E_p or t_p would result in unwanted through breakdown of the discharge gap and decrease in heating contrast. The material would disintegrate efficiently provided that the pulse irradiation energy suffices for matter sublimation in the channel and the time it takes for this energy to be released is significantly shorter than the time required for heat transfer and dissipation of heat in surrounding zones. Maintaining the discharge further would result in redistribution of the permanently generated heat in the volume of the sample. As a result of electrothermal breakdown, this may give rise to unwanted effects, such as overheating, sintering of particles, fusion of particle surfaces and closure of as-formed micro-damages, which would subsequently hinder the access of lixiviant solutions to valuable components hosted within the mineral particles. Among other issues, the role of the high-frequency HPEMP component in the increase of local energy release and peculiarities of electric breakdown development in a semiconductor matrix call for further investigation.

ACKNOWLEDGEMENTS

The authors acknowledge the financial supports from the President of the Russian Federation (Grant No. HIII-472.2003.5) and the program of the Geoscience Branch of the Russian Academy of Sciences ‘Nanoparticles in Geospheres of the Earth: Modes of Occurrence, Industrial and Economic Sequels’.

REFERENCES

1. V.A. Chanturiya, Y.V. Gulyaev, V.D. Lunin, I.J. Bunin, V.A. Cherepenin, V.A. Vdovin and A.V. Korzhenevsky (1999), The opening of the refractory gold-containing ores under power electromagnetic impulses. *Reports of Russian Academic Science (Doklady RAN)*, 366, 5, pp. 680–683 (in Russian).
2. V.A. Chanturiya, I.J. Bunin, V.D. Lunin, Y.V. Gulyaev, N.S. Bunina, V.A. Vdovin, P.S. Voronov, A.V. Korzhenevsky and V.A. Cherepenin (2001), High-power electromagnetic pulses used for disintegration and breaking-up of resistant mineral products. *Journal of Mining Science*, 4, pp. 95–106 (in Russian).
3. V.A. Chanturiya, I.J. Bunin, V.D. Lunin, Y.V. Gulyaev and G.V. Sedelnikova (2003), Effective Breaking-Up Technology for Resistant Gold-Containing Ores and Beneficiation

Products. In: *22nd Proceedings of XXII International Mineral Processing Congress*, Vol. 1, pp. 232–241.

4. V.A. Chanturiya, I.J. Bunin and A.T. Kovalev (2004), Disintegration of mineral media exposed to high-power electromagnetic pulses. *Proceedings of Russian Academic Science (Izvestiya RAN). Series Physics*, 68, 5, pp. 630–632 (in Russian).

DOMAIN SWITCHING CRITERIA FOR TETRAGONAL PHASE FERROELECTRICS: A COMPARATIVE STUDY

M.G. Shaikh[1], S. Phanish[2] and Srinivasan M. Sivakumar[1]
[1]*Department of Applied Mechanics, Indian Institute of Technology Madras, Chennai 600 036, India*
[2]*Department of Ocean Engineering, Indian Institute of Technology Madras, Chennai 600 036, India*

Abstract The nonlinearity observed in piezoelectric behaviour of ferroelectrics like PZT needs to be modelled appropriately. Mainly the domain-switching phenomenon is the cause of this nonlinearity, which requires adoption of suitable switching criteria. In this paper, different domain switching criteria that describe the nonlinear behaviour are evaluated. The comparative study shows that there is considerable deviation from the experimental results. This calls for a better understanding of effects of domain parameters on the macroscopic behaviour of the ferroelectric material.

Keywords: ferroelectrics, nonlinear behaviour, domain switching.

1. INTRODUCTION

The ferroelectrics like PZT undergo structural phase transformation from cubic to tetragonal when they are cooled below their Curie temperature. This phase transition is accompanied by the spontaneous changes in the lattice dimensions and electric dipoles are induced due to separation of centres of positive and negative charges. The minimum energy requirement for stability forces the material to form individually oriented regions, which are known as ferroelectric domains. With respect to the cubic phase cell, there are six possible stable configurations of the tetragonal cell. Therefore, domains built by cells of identical polarization, scattered randomly in the bulk ferroelectric ceramics, could be designated as of type 1 through type 6. As a result the ceramics shows no piezoelectric properties though the individual domains are piezoelectric.

G. R. Liu et al. (eds.), Computational Methods, 1615–1619.

However, by application of sufficient high electric field, usually at high temperature, the alignment of the domains is possible, making the material useful as piezoelectric that could be used in various transducers. This process is known as poling.

During poling, the domains snap to the most favoured orientation with respect to the applied electric field, switching either through 90° or 180° with respect to their original spontaneous polarization direction. The pattern of domain orientations at any instant of loading is decided by the microstructure of the material. The process of domain switching, which nucleates somewhere in the material, causes movement of their walls which is influenced by defects in the material. Application of electric field causes domains to switch through either 90° or 180°, whereas mechanical stress can switch them through 90° only. The complexities associated with domain-type evolution with loading make understanding of the nonlinear behaviour difficult. So far many researchers have proposed different domain switching criteria. In order to make a comparative study of these criteria, a nonlinear model is implemented in this paper. For various domain types, the following equations are used for the piezoelectric responses, namely, strain and electric displacement:

$$\begin{aligned} \varepsilon &= \varepsilon^* + \mathrm{M} : \sigma + \mathrm{d} \cdot \mathrm{E} \\ \mathrm{D} &= \mathrm{D}^* + \mathrm{d} : \sigma + \mathrm{k} \cdot \mathrm{E} \end{aligned} \tag{1}$$

The electric displacement is predicted by averaging contributions by the randomly oriented different domain types. The results are compared with the experimental values of electric displacement observed for PZT-51 by Lu et al. [1].

2. COMPARISON OF DOMAIN SWITCHING CRITERIA

Several domain criteria proposed, available in recent literature, are given in Table 1. Most of them are based on a threshold on work done during switching.

Based on each of the above models, a polycrystalline form of the PZT-51 material is simulated assuming random crystallographic orientations of the domains. Each domain is assumed to be independently subjected to the same forcing of stress and electric field. This gives upper bound solution to the macroscopic problem of simulating the polycrystalline PZT-51 material behaviour. The material constants for the PZT-51 material are given in Lu et al. [1].

Two cases are simulated. The first case is that of a specimen subjected to purely electric loading and the second case is that of electromechanical loading with stress kept at a constant compressive value of −20 MPa (Figure 1).

Table 1. Domain switching criteria.

No.	Author	180° switching criteria	90° switching criteria	Basis
1.	Hwang et al. [2]	$E_i \Delta P_i + \sigma_{jk}\Delta\varepsilon_{jk} \geq 2P^\circ E_c$ (Δ = difference between two domains)	Same as 180°	Only dependent on spontaneous polarization and strain
2.	Sun and Jiang [3]	$E_3 = \dfrac{E^c_{180}}{1+\lambda_1\sigma_{33}}$; $\lambda_1 = \dfrac{2d_{33}}{P_S}$	$E_1 = \dfrac{E^c_{90} - \dfrac{\sigma_{33}\gamma_S}{P_S}}{1+\lambda_1\sigma_{33}}$	Dependent on total strain and electric displacement.
3.	Lu et al. [1]	$W^f_{180} = 2P^\circ E_c$ (W = driving force)	$W^f_{90} = \frac{c-a}{a_0}\sigma_c + \frac{1}{2}(M_{1111} - M_{3333})\sigma_c^2$ (a, c, a_0 = unit cell dimensions; M = compliance)	Based on Gibb's free energy and driving force
4.	Hwang and McMeeking [4]	$E_3 = \dfrac{E^c_{180}}{1+\bar{\lambda}_1\sigma_{33}}$; $\lambda_1 = \dfrac{d_{33}}{P_S}$	$E_3 = \dfrac{E^c_{90} - \dfrac{\sigma_{33}\gamma_S}{P_S}}{1+\bar{\lambda}_1 E_3}$	Not very different from Sun and Jiang model except λ_1 parameter
5.	Fotinich and Carman [5]	$E_3 = E^c_{180} - \dfrac{d_{33}\sigma_{33}}{\varepsilon_{33}}$	None	Driving force based on electric displacement (flux)

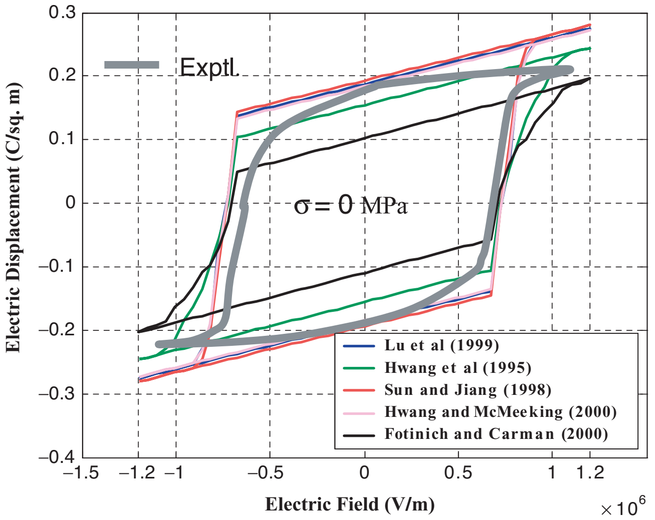


Figure 1. Polarization response under purely electrical loading—Comparison of model and experimental results.

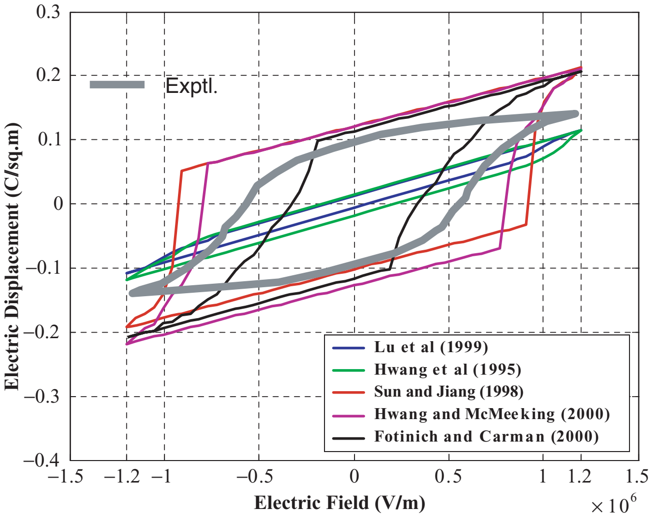


Figure 2. Polarization response under electromechanical loading—Comparison of model and experimental results (constant compressive stress).

Figure 2 shows the results of macroscopic response of polarization in the PZT-51 material under zero compressive stress. The results show that though overall characteristics of the response curve is captured by most of the models, the curved response indicating a gradual switching seen in the experimental curves are not predicted by any of the models studied in this work. The reason could be related to the critical driving force used in the model for switching. Another characteristic that the models do not seem to capture is the flatter response or the saturation response observed in the experiments on a polycrystalline specimen.

As is observed in experiments, the delay in switching phenomena with increase of electrical loading under a compressive stress is seen in simulations in terms of an elongation in the hysteresis loop. However, all the models seem to over-predict the saturation polarization slope. There are considerable variations in the responses of different models for this electrical loading under a constant compressive stress indicating that the 90° switching criterion should to be studied in more detail.

3. CONCLUSION

The above comparison shows that predictions based on theory related only to changes in spontaneous quantities are not sufficient for explanation of the nonlinear behaviour of the ferroelectric ceramics. The predictions based on free energy associated with the various domain types, neglecting any interaction of any type what so ever exists among them and that with defects, also are not enough. An appropriate nonlinear model based on separate criteria for 90° and 180° switching which takes in account the interactions among the various domain types as well as that between domains and defects may be expected to predict the nonlinear piezoelectric response of the ferroelectrics under consideration.

REFERENCES

1. W. Lu, D.N. Fang, C.Q. Li and K.C. Hwang (1999), Nonlinear electric-mechanical behavior and micromechanics modelling of ferroelectric domain evolution. *Acta Materialia*, 47, pp. 2913–2926.
2. S.C. Hwang, C.S. Lynch and R.M. McMeeking (1995), Ferroelectric/ferroelastic interactions and a polarization switching model. *Acta Materiallia*, 43, 5, pp. 2073–2084.
3. C.T. Sun and L.Z. Jiang (1998), Domain switching induced stress induced at the trip of a crack in piezoceramics. In: *Proceedings of the Fourth ESSM and 2nd MIMR Conference, Harrogate*, pp. 715–722.
4. S.C. Hwang and R.M. McMeeking (2000), A finite element model of ferroelectric ferroelastic polycrystals. In: *Proceedings of SPIE, Smart Structures and Materials 2000: Active Materials Behavior and Mechanics, vol. 3992*, pp. 404–417.
5. Y. Fotinich and G.P. Carman (2000), Nonlinear behaviour of polycrystalline piezoceramics. In: *Proceeding of SPIE, Smart Structures and materials 2000; Active Materials Behavior and Mechanics*, 3992, pp. 319–330.

GENERALIZED MAGNETO-THERMOELASTICITY SOLVED BY FEM IN TIME DOMAIN

Xiaogeng Tian and Yapeng Shen
School of Civil Engineering and Mechanics,
Xi'an Jiaotong University, Xi'an, 710049, PR China

Abstract The finite element equations corresponding to two-dimensional generalized magneto-thermoelasticity with G–L theory are derived. Avoiding the complexity of Laplace and Fourier transform and the low precision, the equations are solved directly in time domain. The developed method is used to investigate the generalized magneto-thermoelastic behaviour of a semi-infinite plate subject to thermal shock loading. The results demonstrate that the developed method can faithfully predict the thermoelastic properties and most importantly the delicate second sound effect of heat conduction.

Keywords: generalized magneto-thermoelasticity, finite element method, principle of virtual work, second sound effect.

1. INTRODUCTION

Increasing attention is being devoted to the interaction between thermoelastic field and the electromagnetic field because of its many applications in various branches of science and technology. Ezzat and Othman [1,2] have studied the two-dimensional generalized magneto-thermoelasticity with L–S theory [3] and G–L theory [4]. In solving these problems, both Laplace and Fourier transforms and their inverse counterparts have to be used, and the inverse transform must be carried out numerically due to the complication of corresponding expressions. However, the results cannot capture the generalized thermoelastic properties exactly, e.g., the jump in the temperature will disappear and prediction of the second sound effect is thus not obvious.

G. R. Liu et al. (eds.), Computational Methods, 1621–1627.

The finite element equations for generalized magneto-thermoelasticity in a perfectly conducting medium are derived within the framework of G–L theory, with the equations solved in time domain. The proposed method is then applied to investigate a half space plate under magnetic field subjects to a heat shock loading.

2. FINITE ELEMENT EQUATIONS

In the analysis, we assume that the medium under consideration is a perfect electric conductor. The linearized equations of electromagnetic for moving media are

$$\begin{aligned} \mathrm{curl}\boldsymbol{h} &= \boldsymbol{J} + \varepsilon_0 \dot{\boldsymbol{E}}, \qquad & \mathrm{curl}\boldsymbol{E} &= -\mu_0 \dot{\boldsymbol{h}}, \\ \boldsymbol{E} &= -\mu_0(\dot{\boldsymbol{u}} \times \boldsymbol{H}), \qquad & \mathrm{div}\,\boldsymbol{h} &= 0 \end{aligned} \tag{1}$$

where $\boldsymbol{J}$ is the electric current density, ε_0 and μ_0 are the magnetic and electric permeabilities, respectively. $\boldsymbol{E}$ is the induced electric field, $\boldsymbol{H}$ is the magnetic field, $\boldsymbol{H} = \boldsymbol{H}_0 + \boldsymbol{h}$, $\boldsymbol{H}_0$ is the initial magnetic field and $\boldsymbol{h}$ is the induced magnetic field. Assuming that all causes producing the wave propagation are independent of the variable z and the wave are propagated in the plane of the plate (xy-plane). The initial magnetic field $\boldsymbol{H}_0 = (0, 0, H_0)$. Then the relation (1) yields

$$\boldsymbol{J} = \mathrm{curl}\boldsymbol{h} - \varepsilon_0 \dot{\boldsymbol{E}}, \qquad \boldsymbol{E} = \mu_0 H_0(-\dot{v} \;\; \dot{u} \;\; 0), \qquad \boldsymbol{h} = -H_0(0 \;\; 0 \;\; e) \tag{2}$$

where e is the cubical dilation, $e = \partial u/\partial x + \partial v/\partial y$. According to the Lorentz force $\boldsymbol{F} = \boldsymbol{J} \times \boldsymbol{B}$ and $\boldsymbol{B} = \mu_0 \boldsymbol{H}$, the external body force is

$$\boldsymbol{F} = \mu_0 H_0^2(\partial e/\partial x - \varepsilon_0\mu_0\ddot{u},\; \partial e/\partial x - \varepsilon_0\mu_0\ddot{v},\; 0) \tag{3}$$

The G–L theory of generalized magneto-thermoelasticity consists of the equilibrium equation, the heat transfer equation and Fourier's heat conduction law

$$\sigma_{ji,j} + f_i = \rho\ddot{u}_i, \qquad \rho T\dot{\eta} = -q_{k,k} + \rho\bar{h}, \qquad q_i = -k_{ij}T,_j \tag{4}$$

The linear constitutive equations are adopted, i.e.,

$$\begin{aligned} \sigma_{ij} &= C_{ijkl}\varepsilon_{kl} - \beta_{ij}(T - T_0 + \tau_1\dot{T}), \\ \rho\eta &= \beta_{ij}\varepsilon_{ij} + c_E(T - T_0 + \tau_2\dot{T}) \end{aligned} \tag{5}$$

where σ_{ij} is the stress, f_i is the body force. Only the external body force is considered. ρ is the mass density, u_i is the displacement. T is the temperature, T_0 is the reference temperature, η is the entropy density, q_k is the heat flux, $\bar{h}$ is the heat source density, k_{ij} is the thermal conductivity coefficient. C_{ijkl} is the elastic stiffness, β_{ij} is the thermal constant, and c_E is the specific heat capacity. When τ_1, and τ_2 vanish identically, the G–L theory reduces to the classical thermoelasticity.

The appropriate boundary conditions associated with Equation (4) must be adopted in order to properly formulate a problem. The finite element equations of generalized magneto-thermoelasticity can be readily obtained by following the standard procedure. The principle of virtual work for generalized magneto-thermoelasticity is

$$\begin{aligned}&\int_V \left[\delta\{\varepsilon\}^T\sigma + \delta\{T'\}^T\{q\} - \delta T\ \rho T_0\dot{\eta}\right] dV\\ &= \int_V \delta\{u\}^T(\{F\} - \rho\{\ddot{u}\})\,dV\\ &\quad + \int_{A_\sigma} \delta\{u\}^T\{\bar{T}\}\,dA + \int_{A_q} \delta T\,\bar{Q}dA \end{aligned} \tag{6}$$

Substituting Equation (3), Equation (4) and boundary conditions into Equation (6), the finite element equations corresponding to the generalized magneto-thermoelasticity can be obtained.

$$\begin{aligned}\sum_{e=1}^{ne}\Bigg(&\begin{bmatrix} M_{11}^{(e)} & 0 \\ 0 & M_{22}^{(e)} \end{bmatrix}\begin{Bmatrix} \ddot{u}^{(e)} \\ \ddot{T}^{(e)} \end{Bmatrix} + \begin{bmatrix} 0 & -C_{12}^{(e)} \\ C_{21}^{(e)} & C_{22}^{(e)} \end{bmatrix}\begin{Bmatrix} \dot{\boldsymbol{u}}^{(e)} \\ \dot{\boldsymbol{T}}^{(e)} \end{Bmatrix}\\ &+ \begin{bmatrix} K_{11}^{(e)} & -K_{12}^{(e)} \\ 0 & K_{22}^{(e)} \end{bmatrix}\begin{Bmatrix} u^{(e)} \\ T^{(e)} \end{Bmatrix} = \begin{Bmatrix} F_1^{(e)} \\ -F_2^{(e)} \end{Bmatrix}\Bigg)\end{aligned} \tag{7}$$

where $\{u^{(e)}\}$ and $\{T^{(e)}\}$ are nodal displacement and temperature, ne is the number of elements. For numerical convenience, the non-dimensional quantities are introduced

$$\begin{gathered} x(y) = c_1\xi x^*(y^*), \qquad \theta = (T - T_0)/T_0, \qquad t(\tau_i) = c_1^2\xi t^*(\tau_i\ ^*),\\ Q = Q^*/kT_0c_0^2\eta_0^2 \qquad u_i = c_1\xi u_i^*, \qquad \sigma_{ij} = \sigma_{ij}^*/\mu,\\ \xi = \rho c_E/\kappa, \qquad c_1^2 = (\lambda + 2\mu)/\rho, \qquad i = 1, 2 \end{gathered} \tag{8}$$

The asterisk quantities denote dimensional variables. λ and μ are the Lame constants. The finite element Equation (7) will be solved in non-dimensional space.

Avoiding the larger time consumption and lower precision of transform method, Equation (7) is solved in time domain. In the solution, the element size of the structure, especially at the heat wave front, should be small enough to guarantee the precision of the results due to the large temperature gradient. So the structure should be remeshed during the analysis due to heat conduction.

3. NUMERICAL EXAMPLE

We assume that the plate investigated is made of isotropic material. The physical constants are listed as below.

$$
\begin{aligned}
&\rho = 8945\ \mathrm{kg/m^3}, \qquad \nu = 0.3, \qquad \kappa = 386,\\
&\quad \alpha_t = 1.78e-5\ \ \mathrm{m/K\cdot m}\\
&\lambda = 77.6\ \mathrm{GPa}, \qquad T_0 = 293\ \mathrm{K}, \qquad \mu = 38.6\ \mathrm{GPa},\\
&\quad \tau_1 = \tau_2 = 0.05\\
&c_E = 383.1\ \mathrm{J/kg\cdot K}, \qquad \varepsilon_0 = 10^{-9}/(36\pi)\ \mathrm{F/m},\\
&\quad \mu_0 = 4\pi{}^{*}10^{-7}\ \mathrm{H/m}
\end{aligned}
\tag{9}
$$

The example is a half space plate ($x \geq 0$, $|y| < \infty$, $z = 0$) in magnetic field subject to a heat source of intensity $\theta(y, t) = H(a - |y|)H(t)$ at $x = 0$, a is a constant. The boundary of the plate is traction free. Due to the symmetry of the problem only half of the plate (i.e., $y \geq 0$) is analysed. The x-axis is on the symmetrical edge. The initial magnetic field H_0 is 10e7/(4π), so the magnetic strength is $B = 10\mathrm{T}$.

Figure 1 shows the temperature contours at time $t = 0.04$. The white colour in Figure 1 represents the temperature does not change. We can see that the temperature changed zone is limited in an area and the temperature keeps original out of the zone. The area becomes larger with the passage of time. To exploit the implication of the zone, temperature along the x-axis at different

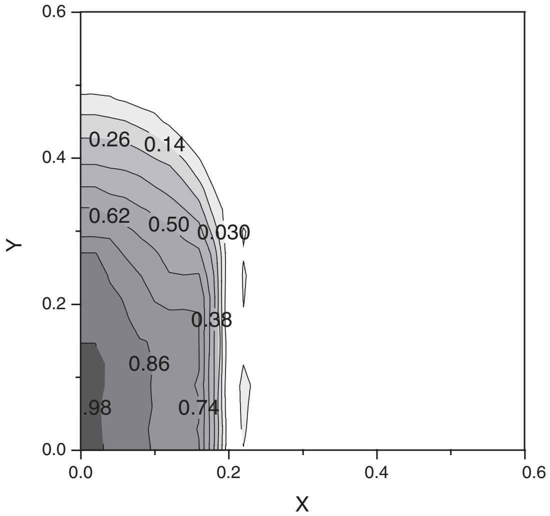


Figure 1. The temperature of the plate.

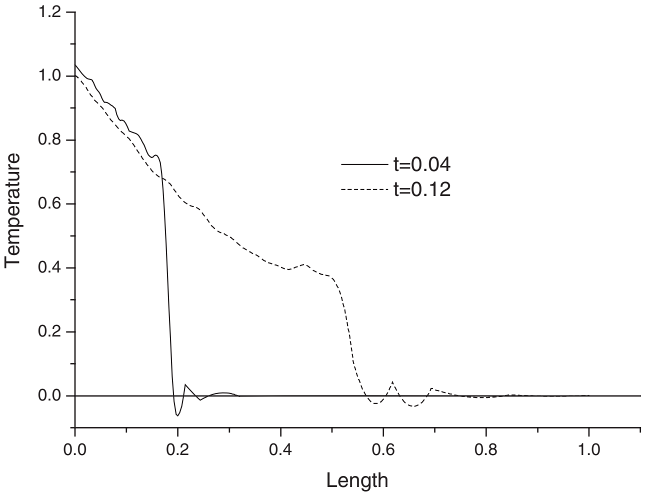


Figure 2. The temperature on the x-axis.

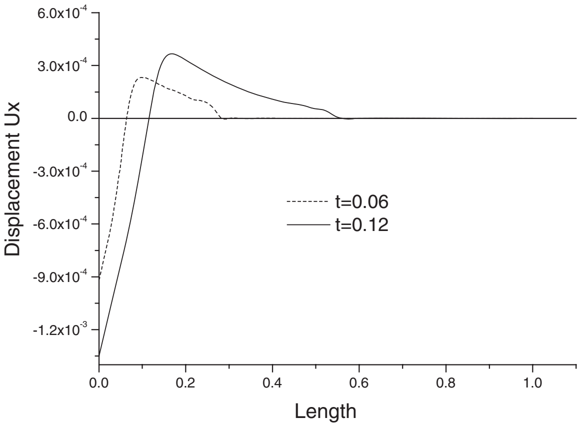


Figure 3. The displacement u_{x} on the x-axis.

time is shown in Figure 2. There is a clear jump in the temperature and the temperature reduces abruptly to zero at the jump. It means the heat conducts in the structure at a finite velocity and the temperature of the structure changes only when the heat wave arrives. It is difficult to be captured by using the transform technology. However, the developed finite element method successfully predicts the second sound effect, showing its superiority over the transform method in two-dimensional generalized magneto-thermoelasticity. The displacement u_{x} on the x-axis at different time is shown in Figure 3. It also shows that the heat conducts at a finite speed.

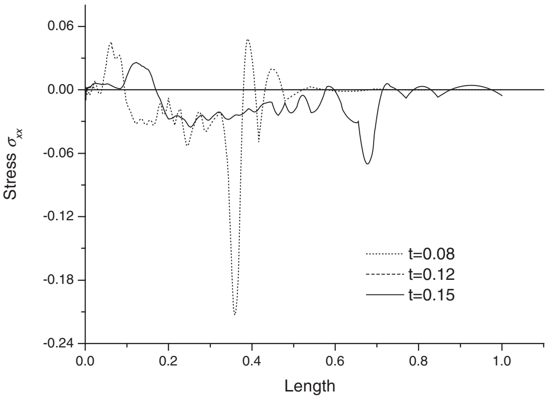


Figure 4. Stress σ_{xx} on the x-axis.

Stress σ_{xx} on the x-axis at different time is shown in Figure 4. The slight oscillation presented on the stress curves is due to numerical noise. We can find that the shape of stress distribution is similar to that of one-dimensional problem obtained theoretically. We also find a peak occurs at the heat wave front and the larger temperature jump, the larger stress peak observed. This may be caused by the temperature discontinue at the heat wave front. The induced magnetic field also has a peak at the heat wave front and the peak becomes smaller with the temperature jump decreasing. It can be predicted that the peak in stress and induced magnetic field will disappear when the temperature jump vanishing.

4. CONCLUSIONS

The finite element equations corresponding to generalized magneto-thermoelasticity with two relaxation time parameters are derived. The numerical results demonstrate that the developed method can faithfully predict the delicate second sound effect in two-dimensional generalized magneto-thermoelastic solids whilst this effect is usually difficult to be modelled by the transform method. The results also show that the stress and induced magnetic field on the plate has a peak at the heat wave front and the magnitude of the peak is related to the jump of the temperature.

REFERENCES

1. M.A. Ezzat and M.I. Othman (2001), Electromagneto-thermoelastic plane waves with thermal relaxation in a medium of perfect conductivity. *Journal of Thermal Stresses*, 21, pp. 411–432.

2. M.A. Ezzat, M.I. Othman and A.A. Smaan (2001), State space approach to two-dimensional electro-magneto-thermoelastic problems with two relaxation times. *International Journal of Engineering Sciences*, 39, pp. 1383–1404.
3. H. Lord and Y. Shulman (1967), A generalized dynamical theory of thermoelasticity. *Journal of Mechanics and Physics of Solids*, 15:299–309.
4. A.E. Green and K.E. Lindsay (1972), Thermoelasticity. *Journal of Elasticity*, 2, pp. 1–7.

MICROSTRUCTURE REPRESENTATION AND SIMULATION TOOLS FOR MICROSTRUCTURE-BASED COMPUTATIONAL MICRO-MECHANICS OF HETEROGENEOUS MATERIALS

A.M. Gokhale, H. Singh and Z. Shan
School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA-30332-0245

Abstract Microstructures of advanced materials such as composites often contain ensembles of microstructural features (particles, fibers, voids, etc.) that have complex morphologies, non-uniform spatial distributions and partially anisotropic angular orientations. These microstructural complexities significantly affect the micromechanical response and mechanical properties of such materials, and provide opportunities and challenges for microstructure sensitive design of materials. In this contribution, we present digital image analysis and microstructure modelling algorithms and tools that can be used to simulate/represent complex geometry material microstructures in the simulations of micro-mechanical response and mechanical behaviour of heterogeneous materials such as composites and multiphase alloys. The methodologies are presented through their implementation in the finite elements (FE)-based simulations on the realistic microstructure models.

Keywords: microstructure-based modelling, finite elements, design, composites.

1. INTRODUCTION

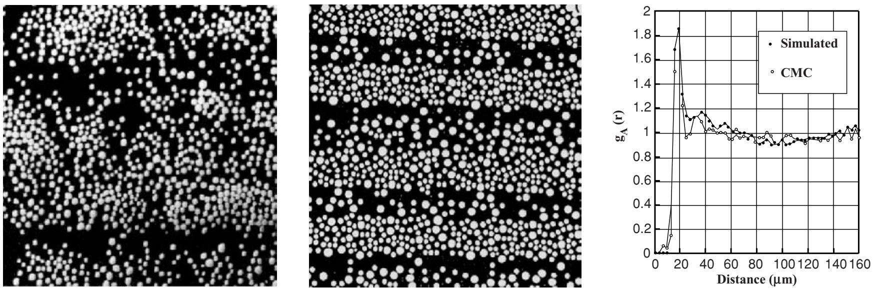
In complex microstructures of engineering alloys and composites, damage evolution and deformation processes are often governed by several types of microstructural features present at different length scales, and microstructural clustering and spatial correlations affect the damage evolution and deformation paths. Nonetheless, to simplify the computations, most micro-mechanics and

G. R. Liu et al. (eds.), Computational Methods, 1629–1633.

damage mechanics-based numerical modelling studies ignore such complex details of the geometry of real material microstructures. On the other hand, to develop the technology of 'materials by design,' it is essential to develop the tools that permit realistic simulations of variations of micro-mechanical response due to changes in specific microstructural parameters. Therefore, there is a need to develop techniques that incorporate (i) complex particle/feature morphologies, (ii) spatial microstructural clustering, (iii) morphological anisotropy, and (iv) multi-scale microstructures in the simulation of micro-mechanical response of heterogeneous materials. Digital image processing and statistical microstructure modelling techniques are now developed to an extent where they can be fruitfully used to incorporate the detailed geometry of multi-scale microstructures in the computational mechanics modelling of micro-mechanical response of elasto-plastic materials. These methodologies are also useful for design of microstructures for optimum micro-mechanical response. Applications of these important enabling tools for modelling of micro-mechanical behaviour of materials are presented in this contribution. The next section of the paper describes computer simulations of realistic microstructures to incorporate spatial clustering, non-randomness and complex particle/void morphologies, which is followed by implementation of the simulated microstructure in the finite elements (FE)-based simulations of micro-mechanical response of the complex microstructure of advanced ceramic matrix composite.

2. COMPUTER SIMULATED REALISTIC MICROSTRUCTURES

Figure 1 (a) shows a low-magnification micrograph of spatially non-uniform microstructure of a ceramic matrix composite (CMC) containing aligned SiC



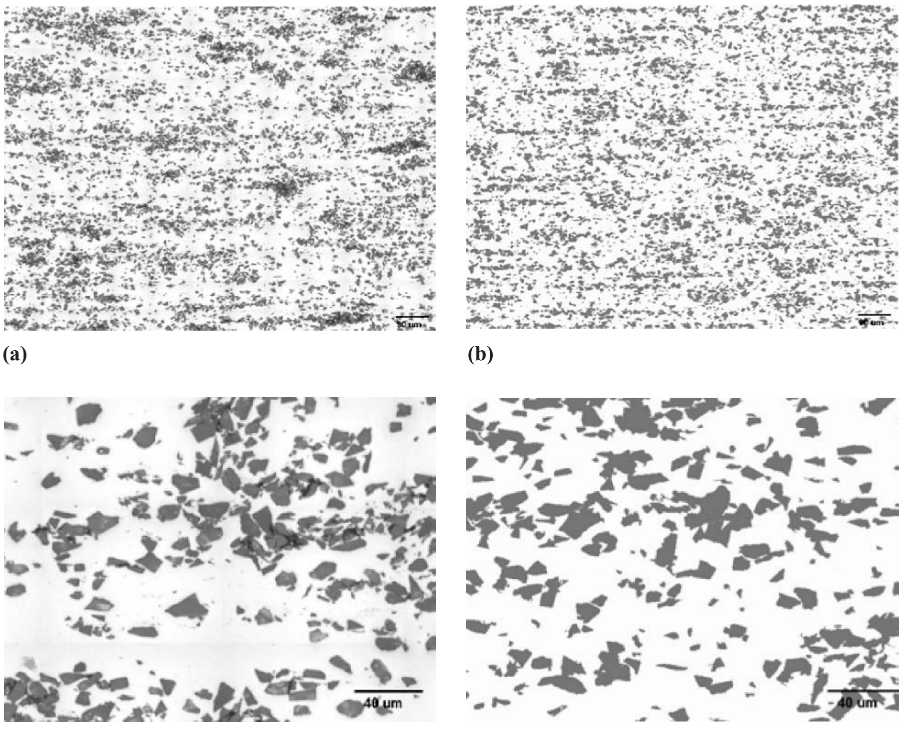
a b c

Figure 1. (a) Real microstructure, (b) Simulation, (c) Radial distributions.

(nicalon) fibres in MAS glass ceramic matrix. The methodology to incorporate the details of spatially clustered fibers in the FE-based simulation via microstructure modelling involves two steps: (1) development of a computer generated microstructure model that is statistically similar to the real microstructure so that it has the same fibre volume fraction, size distribution and statistically representative non-uniform spatial arrangement of fibres and (2) implementation of the computer generated microstructure model in the FE computations of micro-mechanical response. The methodology for development of appropriate microstructure model involves (i) quantitative characterization of volume fraction, size distribution and spatial arrangement of the fibres in the composite, (ii) choice of a generic Monte Carlo type model on the basis of the processing history of the composite and (iii) optimization of the parameters of the computer simulated models till the attributes of the simulated microstructure are in agreement with the corresponding experimental data on the composite. In the present material, the spatial inhomogeneity in the arrangement of fibres is primarily due to the slurry infiltration and hot pressing steps involved in the processing that lead to alternate fibre-poor and fibre-rich regions [1]. Such inhomogeneity can be represented by a model consisting of alternate strips of microstructure having the same size distribution but different number densities of the fibres, where within each strip the fibre centres have a uniform random distribution. In such a simple model, the extent of the inhomogeneity (clustering) can be varied by changing the width and number density of fibres in the alternate fibre-rich and fibre-poor strips. These geometric parameters together with the fibre volume fraction and size distribution specify the simulated microstructure. The optimum model parameters can be arrived at through iterations involving comparison of the statistical distribution functions that quantify spatial arrangement of fibres (for example, radial distribution shown in Figure 1 c) in the simulated microstructure with the corresponding stereological experimental data on the real microstructure till appropriate agreement between the two is attained [2]. For the present CMC, the nearest neighbour and radial distribution functions are utilized to qualify the spatial arrangement of fibre centres [3]. Figure 1 (b) shows the simulated microstructure of the composite obtained in this manner.

Another example of microstructure simulation is depicted in Figure 2, where the real microstructure of a metal matrix composite (MMC) containing SiC particles (not fibres) of irregular shapes in an Al-alloy matrix (Figures 2 (a) and (c)) is simulated (Figures 2 (b) and (d)) via matching their two-point correlation functions [4].

Note that the irregular morphologies of SiC particles in the real microstructure (Figure 2 (c)) are nicely captured by the simulation (Figure 2 (d)).



(a) (b)

(c) (d)

Figure 2. (a) Real microstructure of MMC, (b) Simulated microstructure of MMC, (c) High mag. view of 2a, (d) High mag. view of 2b.

3. IMPLEMENTATION OF SIMULATED MICROSTRUCTURES IN FE COMPUTATIONS

The simulated microstructures that are statistically representative of the corresponding real microstructures can be implemented in the FE schemes using the techniques described in detail elsewhere [2, 5]. This is illustrated in Figure 3, where the simulated microstructure of CMC shown in Figure 1 (b) is implemented in the FE scheme to simulate local distributions of maximum principal stress in the ceramic matrix, which is useful to model damage initiation due to nucleation of micro-cracks in the ceramic matrix of the composite.

ACKNOWLEDGEMENTS

This research was supported through research grants from Air Force Office of Scientific Research (AFOSR grant number F49620-01-0045) and Division

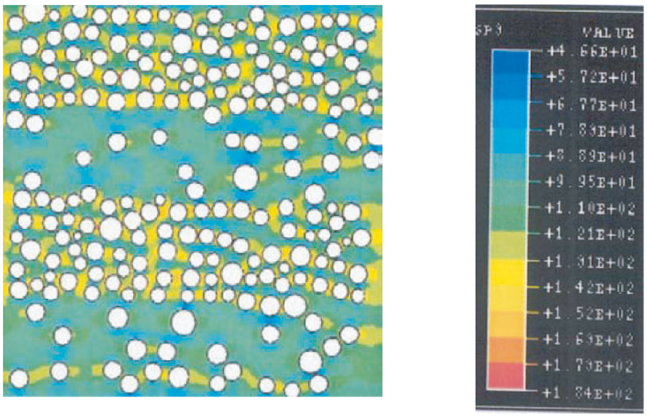


Figure 3. Implementation of computer simulated microstructure in FE scheme to computer distribution of local maximum principal stress.

of Materials Research, U.S. National Science Foundation (DMR-0404668). The financial support is gratefully acknowledged. Any opinions, findings and conclusions or recommendations expressed in this contribution are those of the authors and do not necessarily reflect the views of the funding agencies.

REFERENCES

1. K.K. Chawala (1993), *Ceramic Matrix Composites*, Chapman-Hall, New York, USA.
2. Z. Shan and A.M. Gokhale (2004), Digital image analysis tools for microstructure sensitive design of materials. *International Journal of Plasticity*, 20, pp. 1347–1370.
3. S. Yang, A. Tewari and A.M. Gokhale (1997), Modeling of non-uniform spatial arrangement of fibers in a ceramic matrix composite. *Acta Materiallia*, 45, pp. 3059–3069.
4. A. Tewari, A.M. Gokhale, J.E. Spowart and D.B. Miracle (2004), Quantitative characterization of spatial clustering in three-dimensional microstructures using two-point correlation functions. *Acta Materialia*, 52, pp. 307–331.
5. Z. Shan and A.M. Gokhale (2002), Representative volume element for non-uniform microstructure. *Computational Materials Science*, 24, pp. 361–379.

A UNIFORM EXPRESSION OF INTERMOLECULAR POTENTIAL FUNCTIONS

X. Han[1], G.Q. Xie[2], G.R. Liu[3] and S.Y. Long[2]
[1]*Key Laboratory of Advanced Technology for Vehicle Body Design & Manufacture, Ministry of Education, Hunan University changsha, 410082, P. R. China*
[2]*Department of Engineering Mechanics, Hunan University changsha, 410082, P. R. China*
[3]*Department of Mechanical Engineering, National University of Singapore 9 Engineering Drive 1, Singapore 117576, Singapore*

Abstract In nanomechanics, a systematic approach is required to deal with various expressions of intermolecular potential functions. A concept of universal atomic potential (UAP) is suggested in this paper to unify all the expressions of intermolecular potential functions. The UAP function aims to establish a systematic approach to express the intermolecular potential functions in terms of the UAP form. As examples, two kinds of potential function have been expressed in terms of the given UAP function.

Keywords: molecular dynamics, potential function, uniform atomic expression.

1. INTRODUCTION

Nanomechanics is a rising cross-discipline, it is used to probe the laws of motion and properties of mechanics of substance with 0.1nm$\sim$100nm size. There are main two methods used to simulate physical properties of the nanomaterial and nanostructures, they are Molecular Dynamics (MD) method and Monte Carlo (MC) methods. MD simulation is based on the second Newton law or Hamiltonian principle. The key of MD simulation is to determine the molecular potential function. Normally, the molecular potential functions are experimentally given, however, the characterization of potential parameters is very difficult. Following problems remain: the parameters depend on the substance, thus the potential function of one kind of substance or particle are improper to others; the determined parameters are only composed of two or

G. R. Liu et al. (eds.), Computational Methods, 1635–1639.

three types of atoms; also the parameters derived from the experiment cannot be used to predict the parameters of new substance, etc.

Thus a systematic approach is required to deal with various expressions of intermolecular potential functions. There has been so far no report about a uniform expression of molecular potential functions, this paper attempts to express the molecular potential function uniformly in terms of the suggested UAP form.

2. UAP POTENTIAL FUNCTION

The expression of the UAP function is suggested

$$\phi_{ij}\left(\gamma_{ij}\right) = \sum_k c_{ijk}\left(\frac{1}{\gamma_{ij} - a_k}\right)^k + \sum_l g_{ijl}\left(\gamma_{ij} - b_l\right)^l \tag{1}$$

where, k, l are integers, respectively. γ_{ij} is the distance between particle i and j

$$\gamma_{ij} = \left|\gamma_i - \gamma_j\right| \tag{2}$$

and γ_i is a generalized coordinate vector, it may be linear or angular vector.

The unknown parameters, $c_{ijk}, g_{ijl}, a_k, b_l$, are decided on the basis of the following conditions:

1. the potential function is symmetry;
2. the minimum of the potential function exists;
3. when $\gamma_{ij} = \left|\gamma_i - \gamma_j\right| > \gamma_c$, $\phi_{ij} \approx 0$, here γ_c is the cutting off distance;
4. the function is satisfied with specific conditions of molecular interaction.

Following, as examples of the universal of UAP, the Chelikowsky–Phillips [1–3] and Born-Mayer Huggins (BMH) [4] potential functions are expressed in term of UAP form.

3. CHELIKOWSKY–PHILLIPS POTENTIAL FUNCTION

Chelikowsky and Phillips obtained a covalent interactive classical atomic potential function [1–3]

$$\phi_{ij}\left(r_{ij}\right) = \frac{A\exp\left(-\beta_1 {r_{ij}}^2\right)}{{r_{ij}}^2} - \frac{g_{ij}\exp\left(-\beta_2 {r_{ij}}^2\right)}{r_{ij}} \tag{3}$$

Expanding the following two functions in terms of Taylor series,

$$\exp\left(-\beta_1 r_{ij}{}^2\right) = 1 - \beta_1 r_{ij}{}^2 + \frac{\beta_1^2}{2!} r_{ij}{}^2 - \frac{\beta_1^3}{3!} r_{ij}{}^3 + \cdots + (-1)^k \frac{\beta_1^k}{k!} r_{ij}{}^k$$

$$\exp\left(-\beta_2 r_{ij}{}^2\right) = 1 - \beta_2 r_{ij}{}^2 + \frac{\beta_2^2}{2!} r_{ij}{}^2 - \frac{\beta_2^3}{3!} r_{ij}{}^3 + \cdots + (-1)^k \frac{\beta_2^k}{k!} r_{ij}{}^k$$

Equation (3) can be expressed by

$$\begin{aligned}\phi_{ij}\left(r_{ij}\right) = {} & \frac{A}{r_{ij}^2} - \frac{g_{ij}}{r_{ij}} - A\beta_1 + g_{ij}\beta_2 r_{ij} + \frac{A\beta_1^2}{2!} r_{ij}^2 - \frac{g_{ij}\beta_2^2}{2!} r_{ij}^3 \\ & + \cdots + (-1)^k \frac{A\beta_1}{k!} r_{ij}^{2k-2} + (-1)^{k+1} \frac{A\beta_1}{k!} r_{ij}^{2k-1} \qquad (4)\end{aligned}$$

It can be simplified as

$$\phi_{ij}\left(r_{ij}\right) = \sum_{k=1}^{2} c_{ijk} \left(\frac{1}{r_{ij}}\right)^k + \sum_l g_{ijl} \left(r_{ij}\right)^l \qquad (5)$$

This is the UAP expression in terms of $a_k = 0, c_{ijk} = 0 \ \ (k \neq 1, 2)\,, b_l = 0$.

Figure 1 shows that the $\phi - r$ curve of UAP coincides with that of Chelikowsky–Phillips [1].

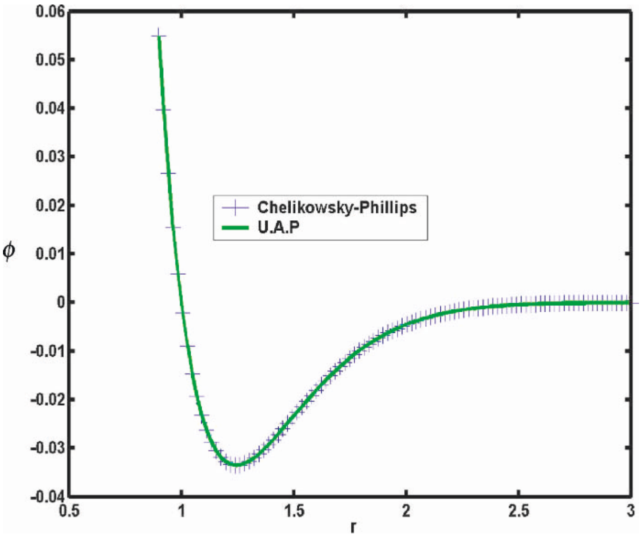


Figure 1. $\phi - r$ curve of C-Phillips vs. UAP.

4. BORN-MAYER HUGGINS (BMH) POTENTIAL FUNCTION

BMH [4] is the atom potential function of alkali metal or alkaline-earth metal

$$\phi_{ij}\left(r_{ij}\right) = \frac{z_i z_j e^2}{r_{ij}} + A_{ij} b \exp\left[\frac{\sigma_i + \sigma_j - r_{ij}}{\rho}\right] - \frac{c_{ij}}{r_{ij}^6} - \frac{D_{ij}}{r_{ij}^8} \tag{6}$$

Expanding the following function on the basis of Taylor series:

$$\exp\left(\frac{\sigma_i + \sigma_j - r_{ij}}{\rho}\right) = \exp\left(\frac{\sigma_i + \sigma_j}{\rho}\right)\left(1 - \frac{r_{ij}}{\rho} + \frac{1}{2!}\frac{r_{ij}{}^2}{\rho^2} - \frac{1}{3!}\frac{r_{ij}^3}{\rho^3} + \cdots + (-1)^k \frac{1}{k!}\frac{\gamma_{ij}^k}{\rho^k}\right) \tag{7}$$

Equation (6) can be expressed by

$$\begin{aligned}\phi_{ij}\left(r_{ij}\right) &= -\frac{c_{ij8}}{r_{ij}{}^8} - \frac{c_{ij6}}{r_{ij}{}^6} + \frac{c_{ij1}}{r_{ij}} - g_{ij1} r_{ij} + g_{ij0} + g_{ij2} r_{ij}{}^2 - g_{ij3r_{ij}}{}^3 + \cdots \\ &\quad + (-1)^l g_{ijl} r_{ij}^l = \sum_k c_{ijk} \frac{1}{\left(r_{ij} - a_k\right)^k} + \sum_l g_{ijl}\left(r_{ij} - b_l\right)^l \end{aligned} \tag{8}$$

This is UAP function in terms of

$$g_{ijl} = (-1)^l \frac{1}{l!} A_{ij} b \exp\left[\frac{\sigma_i + \sigma_j}{\rho}\right] \frac{1}{\rho^l}, c_{ij1} = z_i z_j e^2, c_{ij6} = -c_{ij}, c_{ij8} = -D_{ij}$$

$c_{ijk} = 0, (k \neq 1, 6, 8)$; $a_k = b_l = 0$ Figure 2 shows that the $\phi - r$ curve of UAP coincides with that of BMH and the minimum of UAP and BMH potential functions exist.

5. CONCLUSION

The approach towards a universal expression provides a comprehensive solution strategy for all kinds of the intermolecular potential functions. The examples of the universal expressions of two kinds of intermolecular potential functions have demonstrated the effectiveness. The concept of universal expression is still in the infant stage; however, it offers a new viewpoint for developing the unknown intermolecular potential functions.

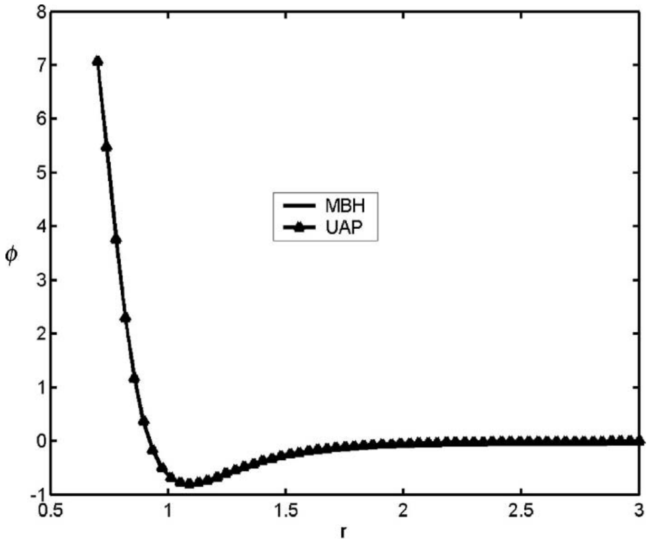


Figure 2. $\phi - r$ curve of MBH vs. UAP.

ACKNOWLEDGEMENT

Authors are grateful for the funds supported by NO.10372031 and NO.10372030 of National Natural Science Foundation of China.

REFERENCES

1. J.R. Chelikowsky and J.C. Phillips (1989), Surface and thermodynamic interact force fields for silicon clusters and bulk phase. *American Physical Society, Physical Review Letters,* 62, 292–295.
2. J.C. Phillips (1973), *Bonds and Bands in Semiconductors*. Academic, New York.
3. J.R. Chelikowsky (1998), Transition from metallic to covalent structures in Silicon clusters. *Physical Review Letter,* 60, 2669.
4. Johnson R.A. and W.D. Wilson (1971), In: Gehlen P.C., et al. eds. *Defect Calculations for Fcc and Bcc Metals Interatomic Potentials and Simulation of Lattice of Lattice Defects*. Plenum, New York, pp. 301–305.

FULL-CHIP SIMULATION OF LSI LITHOGRAPHY MASK USING MULTI-SCALE ANALYSIS

J. Sawamura[1], K. Suzuki[1] and H. Ohtsubo[2]
[1]*The University of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo 113-8656, Japan*
[2]*Hosei University, Japan*

Abstract A novel multi-scale method has been proposed for estimating the in-plane displacements of LSI lithography mask. In order to verify the applicability of proposal method, the mask of simple pattern has been simulated. The simulated results agree with experimental mask feature qualitatively. Moreover, simulation for real device pattern shows full-chip simulation can be performed in practical short time by using this method.

Keywords: lithography mask, multi-scale analysis, full-chip simulation.

1. INTRODUCTION

There is a strong demand for high-density LSI device with minimum pattern sizes below 100 nm. The lithography technique to produce 65/45 nm node devices by using proximity electron lithography (PEL) has been developed [1]. The assessment of overlay error caused by pattern-specific image-placement (IP) error on the mask level is very important for this technique. However, it has been impossible to estimate the IP error in a full-chip scale within the period of time acceptable for practical use because of geometrical complexity of circuit patterns. We have proposed the multi-scale method to the full-chip simulation of mask distortion. In this paper, this method has been demonstrated for contact layer of real device in the 65-nm node. The applicability of this method is discussed in terms of computation time and accuracy.

G. R. Liu et al. (eds.), Computational Methods, 1641–1646.

2. MULTI-SCALE METHOD FOR FULL-CHIP SIMULATION

The finite element method (FEM) enables the simulation of the displacement of local pattern in the part of mask. However, it is impossible to perform the simulation over the full-chip scale, because of the very large number of FEM meshes caused by pattern complexity. Then, we have proposed the multi-scale method for full-chip simulation. The procedure of the proposal method is described as follows:

(1) The entire mask is divided into smaller unit domains with a rectangular shape.
(2) The mask strength distribution induced by geometrical of pattern (e.g., pattern-density) is replaced by the equivalence stiffness matrix for each unit domain. The stiffness matrix is derived from database that relates the pattern-character and the stiffness.
(3) Simulation is performed for the entire mask using these stiffness matrices.

Using the database of stiffness matrices, full-chip simulation can be performed quickly without calculating the stiffness matrices for specific local patterns. Moreover, the accuracy of analysis can be flexibly varied changing the size of local domain.

3. DATABASE OF EQUIVALENCE STIFFNESS MATRIX

For irregular patterns in a real device, the stiffness matrices have to be calculated for all unit domains. However, the database approach need not calculate the detailed FEM analysis of local pattern for estimation of stiffness matrices. Therefore, the full-scale device is analysed in short time. The stiffness matrix—is homogenized elasticity matrix obtained by homogenization assuming the periodic boundary condition. The Young's modulus and Poisson's ratio taken from the material properties of silicon are 160 Gpa and 0.2. Perforation-density of mask pattern is adopted as a parameter to represent the strength induced by pattern feature. Figure 1 shows the database that expresses the relation between the hole-density and the stiffness. In the simulation for real device, the linear approximation as shown in Figure 1 is used for full-chip simulation.

4. SIMULATIONS FOR SIMPLE PATTERN

The applicability of the proposed method is verified by comparing the simulated and experimental data [2]. The simple model used for the test of the

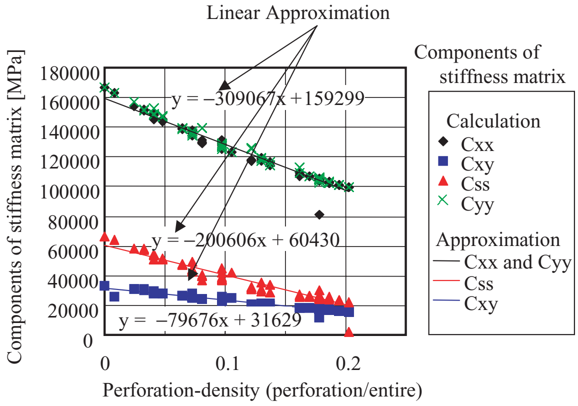


Figure 1. Database of stiffness matrix for contact layer.

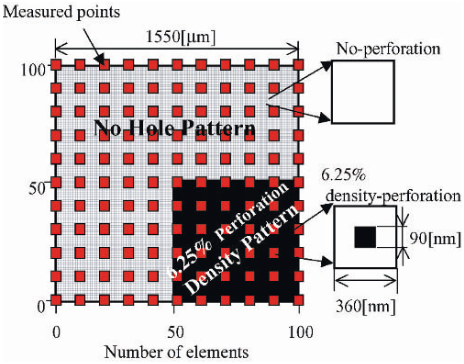


Figure 2. Simple pattern mask.

method is shown in Figure 2. For simple pattern, the matrix is not derived from database of Figure 1. The stiffness matrix for the unit domain with only one hole with the area density of 6.25% is calculated using homogenized method. The stiffness matrix C of no-perforation area and 6.25% density-perforation area are following:

$$\mathbf{C}_{\text{no-perforation}} = \begin{bmatrix} 166667 & 33333 & 0 \\ 33333 & 166667 & 0 \\ 0 & 0 & 66667 \end{bmatrix},$$

$$\mathbf{C}_{6.25\%\text{-perforation}} = \begin{bmatrix} 139246 & 27064 & 0 \\ 27064 & 139246 & 0 \\ 0 & 0 & 49411 \end{bmatrix}$$

Figure 3 shows the boundary conditions. When the analysis is done under this boundary condition, von Mises stress in membranes is estimated at about

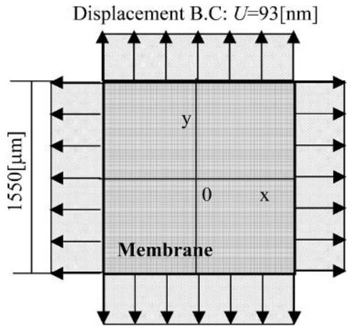


Figure 3. Boundary condition.

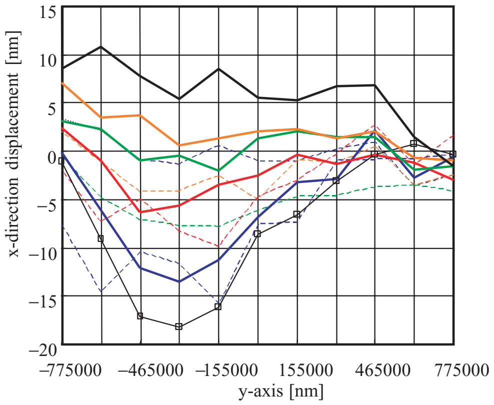


Figure 4. Experimental data.

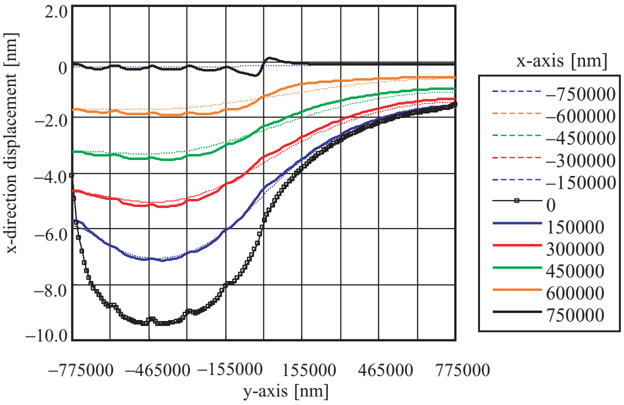


Figure 5. Results of the simulation.

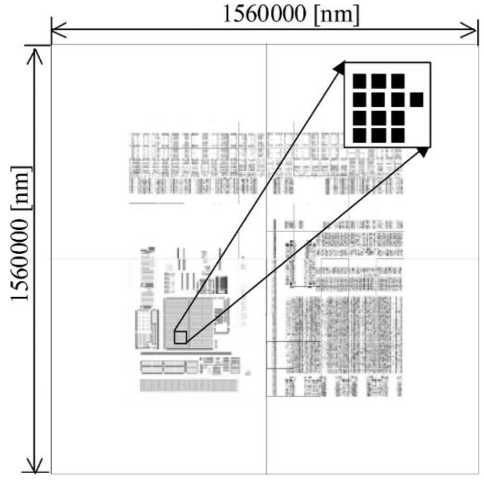


Figure 6. Real device pattern.

30 MPa. The reason for the choice of this boundary is that the inertial stress in membranes of the experiment has been estimated from 20 to 30 MPa. Figures 4 and 5 show the experimental and calculated displacement of membrane, where the vertical and horizontal axes are the x-direction displacement and y-direction coordinate, respectively. Qualitatively, the calculated results agree well with the experimental results and reproduce the deformation of stencil mask under tensions. However, experiment and calculation differ quantitatively. As one of causes of this difference, inappropriate choice of boundary condition related to the experimental uncertainties. Further study is needed to clarify this point.

5. SIMULATION FOR REAL DEVICE

The proposed method and database (Linear approximation in Figure 1) are applied to the simulation for the contact layer of a real device. Figure 6 illustrates the real device pattern for simulation. The size of unit domain is a 1 mm^2. In this simulation, after the calculation of the stiffness for unit domain, the stiffness matrix of larger area is calculated using the stiffness of the unit domain within the area. The reason for 2-step calculation of stiffness matrices is that it is impossible to simulate in the full-chip scale via a 1-step calculation. The boundary condition and material properties are the same as used for the simple mask pattern. Figure 7 shows the distribution of von Mises stress. The calculated stress distribution is reasonable qualitatively. From this simulation, it is shown that the proposed multi-scale method can be applied for the real devices.

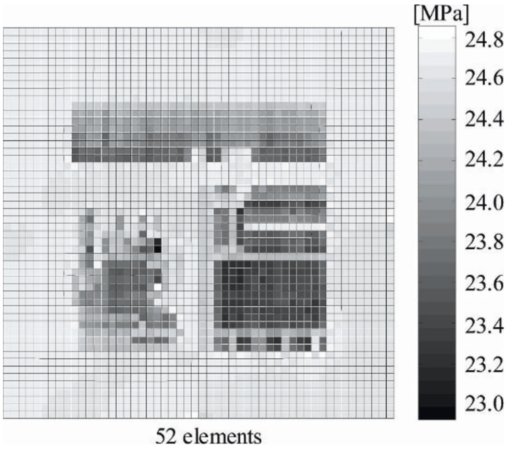


Figure 7. Von Mises stress.

6. CONCLUSIONS

In order to simulate the mask distortion quickly, we proposed the multi-scale method for full-chip simulation. The full-chip contact layer has been calculated by this method. Displacement of simple mask patterns and stress distribution of real patterns have been simulated qualitatively. Using the database of stiffness matrices, we can perform full-chip simulation in the short time. It has been proved that our proposed method can be applied in practical simulation.

REFERENCES

1. S. Omori, K. Iwase, K. Amai, Y. Watanabe, S. Nohama, S. Nohdo, S. Moriya, T. Kitagawa, K. Yotsui, G. Suzuki and A. Tamura (2004), *Proceedings of SPIE*, 5256, 132.
2. K. Yotsui, G. Suzuki and A. Tamura (2003), *Proceedings of SPIE*, 5130, 942.

HIGH-ACCURACY AB INITIO PROGRAMS IN MOLECULAR SCIENCES WITH REFERENCE TO MOLPRO 2000

D.L. Buam[1] and A. Zulfi[2]
[1]*Department of Chemistry, Shillong College, Laitumkhrah, Shillong 793 003, India*
[2]*Grace Systems, Upland Road, Laitumkhrah, Shillong 793 003, India*

Abstract It is a well-known fact that ab initio programs imply a direct quantitative prediction of chemical phenomena from the first principles, thus providing most reliable and accurate results. The term ab initio (Latin) means from first principles and so ab initio quantum chemistry deals with calculations of molecular structure whereby no empirical parameters are included and all integrals are evaluated and thus is independent of any experiment. MOLPRO is an open source ab initio program which is approximately 200,000 lines of code written in Fortran. It is important to note here that an open source program MOLPRO and an open source operating system LINUX was seen as compatible. This paper presents the assessment of the implications of ab initio programs in molecular sciences and the optimized environment for MOLPRO 2000.

Keywords: ab initio methods, assessment, MOLPRO environment.

1. INTRODUCTION

Ab initio quantum chemistry deals with calculations of molecular structure that uses no input other than the Schrödinger equation and fundamental physical constants like Planck's constant, mass and charge of electron, etc. The only information which must be provided are the atomic numbers and positions of the atoms within the system. The basis of these calculations lies in an area of theoretical chemistry called molecular quantum mechanics which relates the molecular properties to the motion and interaction of electrons and nuclei. Thus the key to theoretical chemistry is molecular quantum mechanics. The ab initio methods offer a variety of basis sets of varying complexity where all integrals are evaluated and no empirical parameters are required.

G. R. Liu et al. (eds.), Computational Methods, 1647–1654.

1.1 Theoretical methodology

Molecular orbital theory is an approach to molecular quantum mechanics which uses one-electron functions or orbitals to approximate the full wavefunction. The inclusion of wavefunctions for all possible alternative electron configurations within the framework of a given basis set is termed full configuration interaction (CI). The exact correlation energy can be obtained from a full Configuration Interaction calculation in which all configurations are taken into consideration. Conceptually, we can think of CI calculations as using the variation principle to combine various SCF excited states with the SCF ground state, which lowers its energy. Unfortunately this is not possible for all but small systems. In general, Configuration Interaction is not the practical method of choice for the calculation of correlation energy so a different approach to electron correlation has been introduced, Møller–Plesset perturbation method which is becoming popular in recent years. The simplest electron correlation method to treat electron correlation is Møller–Plesset perturbation theory, which is a special variant of Rayleigh–Schrödinger perturbation theory, with the zeroth-order Hamiltonian [1]. It is a way of introducing electron correlation into a calculation by perturbing the Hamiltonian and calculating the energy to different orders of expansion. The basic idea is that the difference between the Fock operator and the exact Hamiltonian can be considered as a perturbation. Various levels of perturbation theory called MP2, MP3, MP4, etc. can be applied to the problem The Møller–Plesset calculations are not time-consuming and usually give quite accurate geometries and about one-half of the correlation energy. Second-order Møller–Plesset (MP2) perturbation theory is one of the most efficient techniques to include electron correlation.

2. SIGNIFICANCE OF AB INITIO METHODS IN MOLECULAR SCIENCES

The application of quantum chemistry is mainly concerned with the investigation of inter-relations between electronic structure and biological activity, the basic assumption is that all biological processes have a molecular basis. The ultimate goal is in the application to problems which are of interest in research like predicting reactivity and reaction mechanisms, stabilization energies, stable conformations, vibrational frequencies, thermodynamic properties, etc. The ab initio calculations were extensively applied to the study of hydrogen bonded and stacked base pairs in nucleic acids. Such studies would provide the experimentalists with a consistent set of various properties of nucleic acids prior to any experiment being performed or even designed. Now with the advent of high-speed computing systems one can really speculate into the possibility of designing new biological structures which may be potentially useful as therapeutic agents. With these methods one will be able to interpret much better the different physical and chemical properties underlying their biological functions [2–4]. The stabilization energies for AT base pairs calculated by using the second-order Møller–Plesset perturbation theory (MP2) with the 6-31G* basis set yield 11.8 kcal/mol for the Watson Crick pattern and 11.7 kcal/mol [5] for Reverse Watson Crick type. The mechanism and nature of hydrogen bonding were a matter of considerable interest to quantum chemists long before detailed calculations of electronic structure and properties of polyatomic systems became feasible. In our earlier paper [6, 7], we focussed on hydrogen-bonded self-association and hetero-association in nitrogenous bases that could be a basis for constructing the information-bearing macromolecules using the semi-empirical PM3 method. However, these methods are not accurate in their predictions and have been superseded by the ab initio methods. Thus, if one wants to know more than just the structures and other properties that are obtained from potential energy surfaces, one has to resort to a fundamental approach. To name just a few of the well-known ab initio programs are Gaussian 94/98, GAMESS US, GAMESS UK, MOLPRO, SPARTAN, JACGUAR, COLUMBUS, etc.

2.1 Accuracy of ab initio methods

The ab initio programs have a high demand of computational cost, the high computational cost of ab initio calculations implies that they are most profitably used for simulating the detailed electronic structure of a system. High-accuracy computational methods can now be applied to much larger systems than had previously been envisaged, due to increased computer power and algorithmic improvements. This brings the area of biochemistry within the scope of high-accuracy methods previously developed for small molecules. Huge calculations involving large number of basis functions and a large number of electrons are now being more accessible, thanks to the developments of parallel algorithms. Ab initio quantum-chemical calculations with inclusion of electron correlation carried out since 1994 have given a clear picture and insight to the interactions of nucleic acid bases [8]. The progress in theoretical studies on hydrogen bonding has been due to the advent of high-speed computers which allow non-empirical SCF ab initio studies to be carried out on even larger systems which were not feasible earlier. These calculations made it possible to perform the first reliable comparison of the strength of stacked and hydrogen bonded pairs of nucleic acid bases, and to characterize the nature of the interactions of bases. Accurate geometry optimizations on very large molecules have been a long-standing challenge for electronic structure programs, however with the recent ab initio programs on high-performance computers all of this changes. The methods are quite time-consuming and yet give reliable results and represent the most promising means. Though the ab initio methods have a high demand of computer power yet these methods yield accurate results and thus far supersede the semi-empirical methods which have less reliability as their applicability is limited by the requirement of structural parameters. Accurate ab initio calculations can now be performed for much larger molecules than with most other programs. Using recently developed integral-direct local electron correlation methods [9–11], which significantly reduce the increase of the computational cost with molecular size, accurate ab initio calculations can be performed for much larger molecules than with most other programs.

3. MOLPRO PROGRAM

The heart of the MOLPRO program consists of the multi-configuration Self Consistent Field (MCSCF), multi-reference Configuration Interaction (MRCI), and coupled-cluster (CC) routines, and these are accompanied by a full set of supporting features. MCSCF is a means of variationally minimizing the energy of several electron configurations of a given system simultaneously. The methods utilized in the MRCI program, which forms a part of MOLPRO, are almost uniquely capable of dealing with the complexity of the problems, as they are able to efficiently include a large proportion of the correlation energy. Thus emphasis is on highly accurate computations, with extensive treatment of the electron correlation problem, through the multi-reference configuration interaction (MRCI) [12], coupled cluster [13, 14] and associated methods. The simplest electron correlation method to treat electron correlation is Møller–Plesset perturbation theory, which has been mentioned earlier.

3.1 Setting up the MOLPRO environment

MOLPRO is an open source ab initio program which is approximately 200,000 lines of code written in Fortran. It is not like a commercially viable software or application of Microsoft which simply needs installation and ready to run. MOLPRO requires an optimized environment and therefore needs certain specific parameters like (i) processor, (ii) machine architecture, (iii) compatible operating system with the GNU tools/libraries and the compatible

Fortran compiler depending upon the operating system, like f77 or g77, or f90, MIPSpro f90 for SGI systems. Added to this, as any other large program, MOLPRO carries bugs and hence the installation becomes complex. We carried out experimentation with different Operating Systems and machine architectures and found Linux a lean OS, free of fancy. Prime factors of choice are (i) stability, performance and security, (ii) it is one of the fastest operating systems for computational work, (iii) the availability of the GNU tools and libraries along with the g77 compiler, (iv) it is an open source operating system which can be customized to our needs, (v) it is cost effective. Red Hat Linux comes with all the GNU Tools which is Free Foundation Software and gives a complete software development environment. The GNU Tools are distributed primarily as source code, accompanied by the appropriate libraries, header files and make files to compile the tool. Hence, an open source program MOLPRO and an open source operating system LINUX was seen as compatible. MOLPRO is a machine-dependent program and hence it requires a compatible machine architecture for its compilation. We installed Red Hat Linux 6.2 on a Pentium III 550 MHz, 64 MB RAM as this version supports the program and so we loaded the MOLPRO program and configured the program creating the CONFIG and CONFIG_FRONT files which contain the machine-dependent parameters. Then we successfully patched up and compiled MOLPRO and were able to create the MOLPRO executable. An important point here is that when we configure MOLPRO the configure utility after finding the compatible compiler, i.e., Fortran compiler and the right machine architecture prompts for a few questions from the MOLPRO user viz., (a) the licence key and password which must be correctly typed in, (b) the number of atoms and the default selection, (c) the number of valence orbitals and (d) blas options. **NOTE**: At this stage, one must understand that BLAS and LAPACK are subroutine libraries. One must choose from the options of BLAS which are given as 1, 2, 3. On some machines, however, it will be advantageous to use a system-tuned version instead. BLAS = 0 is the default selection and means that the MOLPRO Fortran BLAS routines are used. For example, if you specify 2, the system libraries will be used for level 2 and level 1 BLAS, but MOLPRO's internal routines will be used for level 3 (i.e., matrix–matrix multiplication). In either case, one is prompted for appropriate linker options (e.g., -L/usr/lib –l blas) to access the libraries. It is generally recommended that system BLAS be used if available, but on some workstations MOLPRO's internal matrix routines give comparable performance to system-supplied libraries. A special situation, of course, arises if 64-bit integers are in use since on many platforms the system BLAS and LAPACK libraries only supports 32 integer arguments. **WARNING!** *No system LAPACK library must be used together with BLAS = 4.*

4. RESULTS AND DISCUSSIONS

The tables below represent a comparative study of the application of ab initio and semi-empirical methods to small systems like water molecule and an interesting interpretation of the results is obtained. In this paper, we make use of the ab initio programs MOLPRO and GAMESS and also the semi-empirical program MOPAC 6.0.

Table 1 clearly indicates the vast difference of results between the ab initio and semi-empirical results in regards to simple molecules like water. The energy values recorded in Table 2 represent the sum total of one-electron, two-electron and nuclear repulsion energies at different levels of basis sets using MOLPRO. Here it is noteworthy that the higher basis set level accounts for a much more stable electron and nuclear repulsion energies at different levels of basis sets using MOLPRO energy value as shown by the data. It is also important to note here that the energy values as obtained by the ab initio methods MOLPRO and GAMESS are quite similar (Table 3).

Table 1. Contrast of ab initio and semi-empirical results for water molecule.

	Final energy (kcal/mol)			
System	MOLPRO (ab initio)	GAMESS (ab initio)	PM3 (semi-empirical)	AM1 (semi-empirical)
Water	−76.06458	−75.98374	−53.42645	−59.23816

Table 2. Comparison of results using MOLPRO at different levels of basis set.

				Final energy (kcal/mol)	
Basis set	∠HCN (°)	R_{CN} (Å)	R_{CH} (Å)	Methyl nitrene	Ethyl nitrene
VDZ	110.973	1.4015	1.0987	−93.909260	−132.95359
6-31G*	110.917	1.4028	1.0918	−93.905617	−132.95042
6-311G*	110.897	1.4008	1.0910	−93.921252	−132.96950

Table 3. Comparison of energy values using MOLPRO and GAMESS.

	Final energy (kcal/mol)	
Program	Water	Methyl nitrene
MOLPRO	−76.06458	−93.90561
GAMESS	−75.98374	−94.00396

5. CONCLUSION

Our current approach to the significance of ab initio quantum chemistry programs in molecular sciences urge us in finding out the optimized environment for MOLPRO 2000 which would provide an easy means to the users on how to configure, debug, compile the program. *One can install MOLPRO in a Pentium III 1000 MHz, 128(+) MB RAM, 20(+) GB HDD for running MOLPRO or any other number crunchers for large molecules*. Hence Pentium Architectures 32-bit can become cost-effective solution for these kinds of number crunchers.

6. ACKNOWLEDGEMENTS

The authors are thankful to the writers of the MOLPRO program, P.J. Knowles and H.-J. Werner, and to those who have made contributions to the program like R.D. Amos, A. Bernhardson, A. Berning, P. Celani, D.L. Cooper, H.J.O. Deegan, A.J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, T. Koroma, R. Lindh, A.W. Lloyd, S.J. McNicholas, F.R. Manby, W. Meyer, M.E. Mura, A. Nicklass, P. Palmiei, R. Pitzer, G. Rauhut, M. Schutz, H. Stoll, A.J. Stone, R. Tarroni and T. Thorsteinsson. The authors are also thankful to the writer of GAMESS, Alex A. Granovsky for enabling us to easily and freely access the program.

REFERENCES

1. C. Møller and M.S. Plesset (1934), MP Perturbation Theory. *Physical Review*, 46, p. 618.
2. M. Aida (1988), Characteristics of the Watson-Crick type hydrogen-bonded DNA base pairs. An *ab initio* molecular orbital study. *Journal of Computational Chemistry*, 9, p. 362.
3. P. Hobza, J. Sponer and M. Polasek (1995), H-bonded and stacked DNA base pairs: cytosine dimer. An *ab initio* second order Møller–Plesset study. *Journal of the American Chemical Society*, 117, p. 792.
4. J. Sponer, J. Leszczynski and P. Hobza (1996a), Structures and energies of hydrogen-bonded DNA base pairs. A non empirical study with inclusion of electron correlation. *Journal of Physical Chemistry*, 100, p. 1965.
5. M. Kratochvil, J. Sponer and P. Hobza (2000), Global minimum of the adenine thymine base pair. *Journal of the American Chemical Society*, 122, p. 3495.
6. D.L. Buam and R.H.D. Lyngdoh (2000), Self associative base-pairing in some nitrogenous heterocycles: a PM3 SCF MO study. *Journal of Molecular Structure, THEOCHEM* 505, p. 149.
7. D.L. Buam and R.H.D. Lyngdoh (2002), Information-bearing base-pairs involving hydrogen-bonded nitrogen heterocycles: a theoretical modelling study. *Indian Journal of Chemistry*, Sec B, 41B, 11, p. 2346.

8. J. Sponer, J. Leszczynski and P. Hobza (1996b), Hydrogen bonding and stacking of DNA bases: a review of quantum-chemical *ab initio* studies. *Journal of Biomolecular Structure and Dynamics*, 14, 1, p. 117.
9. C. Hampel and H.-J. Werner (1996), Local correlation methods. *Journal of Chemical Physics*, 104, p. 6286.
10. G. Hetzer, P. Pulay and H.-J. Werner (1998), Local correlation methods. *Chemical Physics Letters*, 290, p. 143.
11. M. Schütz, G. Hetzer and H.-J. Werner (1999), Local correlation methods. *Journal of Chemical Physics*, 111, p. 5691.
12. P.J. Knowles and H.-J. Werner (1988), Multireference CI. *Chemical Physics Letters*, 145, p. 514.
13. P.J. Knowles, C. Hampel and H.-J. Werner (1993), Open shell coupled cluster. *Journal of Chemical Physics*, 99, p. 5219.
14. W.D. Laidig and R.J. Barlett (1984), A multi reference coupled cluster method for molecular applications. *Chemical Physics Letters*, 104, p. 424.

MOLECULAR DYNAMICS SIMULATIONS OF NANOINDENTATION OF POSS MATERIALS

F.L. Zeng and Y. Sun
Department of Astronautic Science and Mechanics, Harbin Institute of Technology,
Harbin 150001, China

Abstract Molecular dynamics (MD) simulations of nanoindentation of polyhedral oligomeric silsesquioxane (POSS) materials $H_8Si_8O_{12}$ (T_8H_8) and POSS polymerized norbornene homopolymer (PNPOSS) at different temperatures were performed. Load-displacement curves were calculated, from which the hardness and elastic moduli at different temperatures were obtained. These results compare well with currently available experimental data and previous theoretical calculations. Surface adhesion and plastic deformation in POSS materials during nanoindention are investigated. The difference between the MD simulations of nanoindentaion and the actual ones are discussed as well.

Keywords: molecular dynamics simulations, nanoindentation, POSS.

1. INTRODUCTION

Polyhedral oligomeric silsesquioxane (POSS), $(RSiO_{1.5})_n$ or T_n, where n is an even number and $R = H$, Cl or a variety of organic groups, is a unique class of materials that has emerged as recently in polymer chemistry. Comparing to other traditional polymer materials, the inimitable characteristic of organic–inorganic nano-hybrid materials POSS is that POSS monomers are combined to the polymer molecules. The POSS monomers lead to serials of excellent properties such as high thermal stability, high tenacity, high intensity and so on. Not only the POSS monomers can form polymers themselves, but also the monomers can be combined to the polymer chains and reinforce the polymer materials. In this paper, mechanical properties of two kinds of POSS materials T_8H_8 (formed by POSS monomer T_8 absolutely) and C6PN (in which the POSS monomers C6POSS ($(RSiO_{1.5})_6$ $R =$ cyclohexyl ring) were combined to

G. R. Liu et al. (eds.), Computational Methods, 1655–1660.

polymerized norbornene homopolymer chains) were studied by the simulations of nanoindentation tests.

2. POLYMER MODEL, INDENTER SHAPE, SIMULATION METHOD AND PROCEDURE

Molecular structures of T_8H_8 and C6PN are shown in Figures 1 and 2, respectively. The C6PN molecule was constructed as the structure in Bharadwaj's work [1] The C6PN molecular chains were built by including 10 mol% of C6POSS and 90 mol% of norbornene. Each chain consists of 10 monomeric units. There is only one C6POSS monomer in the C6PN chain. To avoid the agglomeration of the C6POSS monomer, we selected the POSS monomer attached to the fifth norbornene unit. One such a C6PN included 316 atoms. Two kinds of polymer molecules are included in a layer. The T_8H_8 layer included 200 T_8 molecules (5600 atoms) and the layer size is 62.58 Å × 62.58 Å × 40 Å. The C6PN layer includes 32 C6PN chains (10,112 atoms) and the layer size is 61.12 Å × 61.12 Å × 40 Å. The boundary conditions are periodic in the lateral (x and y) directions, but there is no periodicity along the layer thickness (z) direction. The bottom 2 Å of the layer is held rigid and the surface is free (Figures 1 and 2).

The indenter is a square-based pyramid similar to the Vickers indenter and is constructed by Fe atoms. The indenter material would have to have an elastic modulus much larger than that of the materials being tested. Since we are interested in the material properties of T_8H_8 and C6PN as determined by nanoindentation and not in the tip surface interactions themselves, it is not important how the external load is applied as long as it is applied with the

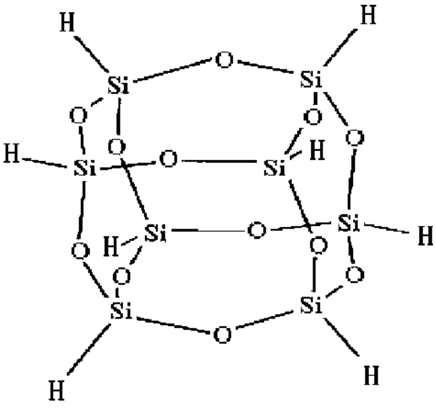


Figure 1. Structure of T_8H_8 molecular.

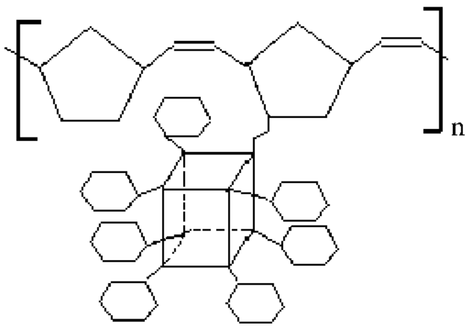


Figure 2. Structure of C6PN chain.

required geometry. During the indentation process, the relative positions of Fe atoms in the indenter were fixed, so the modulus of the indenter can be viewed as infinite. We only need to consider the interactions between the indenter and the materials being tested. The indenter included 285 atoms. The crystal lattice constant of Fe is 2.8664 Å. But after being relaxed, the size of the Fe pyramid changed. The height of the indenter is 15.24 Å. The indenter is shown in Figure 3.

The relationship between the mechanical constants (elastic moduli and hardness) and the load-displacement curves was determined in Sneddon's work [2] for a flat cylindrical punch. The same relation has been shown to hold regardless of indenter shape for a wide range of indenters, provided the area of contact remains constant during the initial unloading This relation is used for most indenter shapes under circumstances where the material indented can be approximated as an elastic continuum.

The PCFF force field established by Maple and Sun [3, 4] was used in order to calculate potential energies between atoms. And the velocity-Verlet algorithm was used to calculate the position and velocity of atoms. The simulation procedure of two kinds of materials at different temperatures are as follows: Firstly, the layer was relaxed at 0 K using a relaxation time of 10 ps, and the

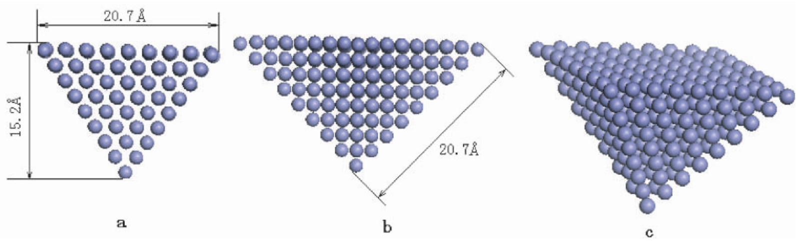


Figure 3. Fe indenter and its size (c) in the 3D view.

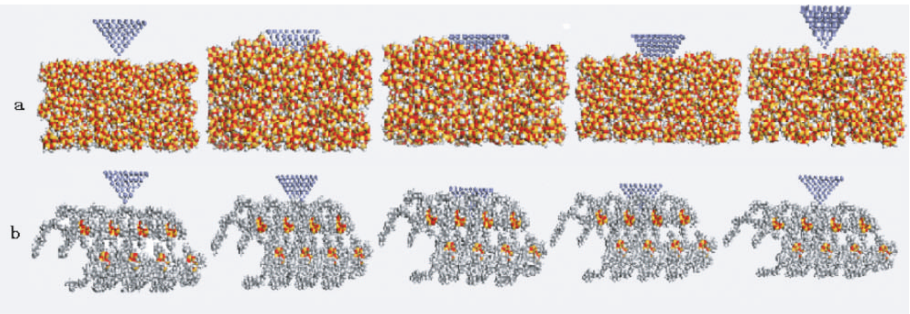


Figure 4. Atomic configurations for some of steps at 300 K (a) T_8H_8 (b) C6PN.

time step was 2 fs, after which the system was heated to 300, 400 and 500 K, respectively, using the Andersen thermostat dynamics over 100 ps and the time step was 1 fs. Then the Fe indenter was brought towards the free surface at a rate of 1 Å per 10,000 time steps, or 10 m/s (the time step was 1 fs). After every 10,000 time steps, the configuration of the layer was saved and local stresses were calculated, The total z-component of the force on the indenter atoms was recorded at the same time. After 100,000 time steps of simulation, the indenter was held 10 Å in the substrate. Then the indenter was withdrawn using the same schedules used in the loading process. From this, a load-displacement curve was obtained for the overall simulation. Figure 4 shows a CPK model of the atomic configurations for different steps at 300 K.

3. RESULTS AND DISCUSSIONS

Figure 5 shows the load-displacement curves for the molecular dynamics (MD) simulations of T_8H_8 (a) and C6PN (b) at different temperatures. From these data, the elastic moduli and hardness of T_8H_8 and C6PN at different

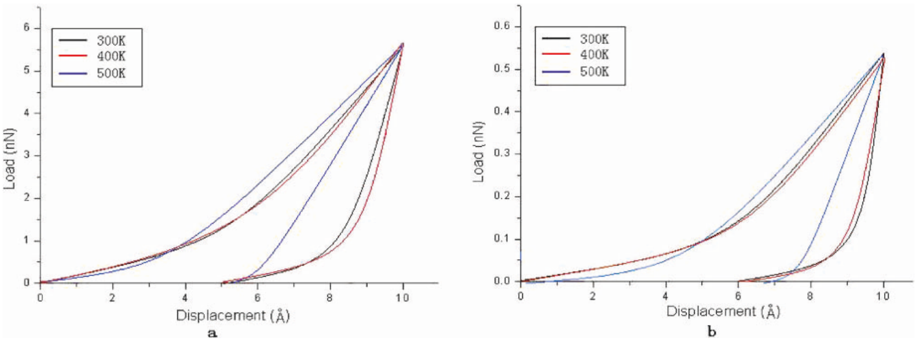


Figure 5. Load-displacement curves for (a) T_8H_8 (b) C6PN at different temperatures.

Table 1. Simulation results and theoretical results (GPa).

Temperature (K)	E (T_8H_8)	E** (T_8H_8)	E (C6PN)	E* (C6PN)	H (T_8H_8)	H (C6PN)
300	10.2	12.05	2.70	2.25	3.97	0.31
400	16.7	15.83	1.67	1.40	3.60	0.32
500	4.82	4.97	0.78	0.69	6.18	0.35

E: Elastic modulus; H: Hardness; *From Bharadwaj's work; **From the method in Bharadwaj's work.

temperatures were obtained. The results are listed in Table 1. The elastic moduli compare well with Bharadwaj's results [1] and using the method introduced in Bharadwaj's work [1]. It is seen from Table 1 that the difference between the present results and that in the existing literature is within 10%. The discrepancy may come from two ways. One is that the size of layers used in this study is too small to account for statistical variation; and the other maybe due to the inadequate relaxation in simulations. Although there are no existing results to compare the hardness data, it has been pointed out by several investigators that hardness is known to increase with decreasing indentation sizes. Since the size of the indenter here is quite small, the hardness should be much higher than the measured values using micro-indentation test. If the attractive forces between the indenter and substrate atoms were considered, the jump-to-contact and surface-adhesion phenomena would be observed. This is the fundamental difference between micro- and nanoindentations.

4. CONCLUSIONS

MD simulations of nanoindentation have been successfully used to compute the mechanical properties of POSS materials. It was found that the pure POSS materials have excellent mechanical properties and the POSS monomers can reinforce the normal homopolymer significantly. The deformation and damage mechanism with external contact can be investigated using this method.

ACKNOWLEDGEMENT

The authors would like to thank the National Natural Foundation (10472028), for the financial support of this project.

REFERENCES

1. R.K. Bharadwaj, R.J. Berry and B.L. Farmer (2000), Molecular dynamics simulation study of Norbornene-POSS polymers. *Polymer*, 41, pp. 7209–7221.

2. I.N. Sneddon (1965), The relaxation between load and penetration in the axisymmetric Boussinesq problem for a punch of arbitrary profile. *International Journal of Engineering Science*, 3, pp. 47–57.
3. J.R. Maple, M.J. Hwang, T.P. Stockfisch, U. Dinur, M. Waldman, C.S. Ewig and A.T. Hagler (1994), Derivation of class II force fields. 1. Methodology and quantum force field for the alkyl functional group and alkane molecules. *Journal of Computer Chemistry*, 15, pp. 162–182.
4. H.Sun, S.J. Mumby, J.R. Maple and A.T. Hagler (1994), An ab initio CFF93 all-atom force field for polycarbonates. *Journal of American Chemistry Society,* 116, pp. 2978–2987.

STRESS CONCENTRATION FOR NANOCAVITIES FROM MOLECULAR SIMULATION

Haitao Zhang
School of Aeronautics and Astronautics, Purdue University, USA

Abstract Size-dependent stress concentration of nanocavities was investigated with EAM (embedded-atomic potentials) based molecular statics simulations. It was found that at nanoscale stress concentration factor of nanocavities decreases as the characteristic size decreases and this trend is not affected by grain orientations.

Keywords: nanocavities, molecular simulation, stress concentration.

1. INTRODUCTION

With the development of nanotechnology, there is a growing interest in developing devices and materials with nanomaterials. For example, microresonators with nanothickness were made of silicon to detect single virus. High frequency resonator (380 MHz), which can be used for many communication applications as frequency references and filters, with size of 2 mm × 200 nm × 100 nm was fabricated from single crystal silicon. These promising applications of nanomaterials have also brought new challenges in analysing mechanical behaviours of nanomaterials. For example, recent research has shown that conventional continuum mechanics fails to predict the mechanical behaviour of nanomaterials [1, 2]. At nanoscale, material elastic constants also show significant size-dependence [1, 3].

At nanoscale, stress and strain significantly affect the physical properties of nanomaterials. For example, strain will affect the optoelectronic properties of quantum dots. Consequently, it is critical to predict the stress and strain accurately at nanoscale.

G. R. Liu et al. (eds.), Computational Methods, 1661–1666.

Recently, Sharma et al. [4] studied the stress concentration of nanocavities with an elasticity theory involving surface stress and found stress concentration factor is size-dependent at nanoscale. In their work, Sharma et al. considered isotropic materials, however, they employed anistropic surface material constants in the discussion and concluded that stress concentration factor can both increase and decrease with the decrease of characteristic nanocavity size. Consequently, it is not clear how the stress concentration factor depends on nanocavity size for either isotropic or anisotropic materials.

In the present paper, I will studied how the stress intensity factor depends on characteristic nanocavity size with EAM (embedded-atomic potentials) molecular statics simulations. The detailed approach used in the simulations are introduced first, the simulation results are then presented and discussed. Finally, conclusions are drawn based on the simulation results.

2. SIMULATION APPROACH

Recognizing nanocavities generally exist in anisotropic matrix, anisotropic problems are considered in this paper. Without loss of generality, face-centred cubic (FCC) single crystal Ni is chosen for study. Schematic diagram illustrating the cross-section of the material system is shown in Figure 1. This represents a material with infinite dimension in the [100], [010] and [001] directions. The cavity, with square cross-section geometry, goes through the whole material body in the [010] direction. Due to the symmetric properties of the material

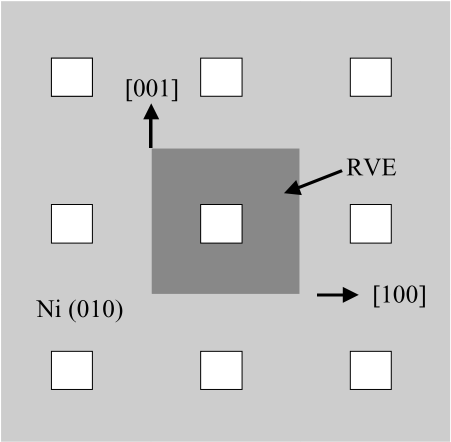


Figure 1. Cross-sectional of the modelling.

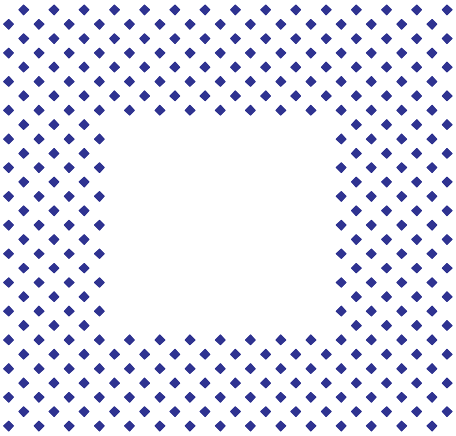


Figure 2. Atomic modelling projected in the [010] direction.

body, a representative volume element is chosen for the simulations (see Figure 1). In the simulations, the length L is chosen as 52.8 Å, the characteristic size of nanocavities ranges from 7.04, 14.08, 21.12 to 28.16 Å. The RVE length in the [010] direction is 17.6 Å. Figure 2 shows a typical molecular modelling of the RVE. Periodic boundary condition is applied in the [100], [010] and [001] directions. The lattice constant of Ni is 3.52 Å.

In this study, an EAM based molecular statics program is used to obtained the those material constants. Embedded-atom method (EAM) potential is a widely used potential for FCC metals. In the EAM potentional, the total energy Φ is composed of an electrostatics pairwise interaction energy between atoms α and β, $V(r_{\alpha\beta})$, and an embedding energy $F(\rho_\alpha)$, which is an energy to embed the atom in the local-electron density created by its near neighbours, i.e.,

$$\Phi = \frac{1}{2}\sum_{\alpha\neq\beta}^{N} V\left(r_{\alpha\beta}\right) + \sum_{\alpha}^{N} F\left(\rho_\alpha\right), \tag{1}$$

$$\rho_\alpha = \sum_{\beta} \phi\left(r_{\alpha\beta}\right). \tag{2}$$

where N is the number of atoms in the system, $r_{\alpha\beta}$ is the distance between atoms α and ρ_α is electron density contributed by other atoms to the site of atom α (the summations are over all the atoms β that interacts with the atom α), and $\phi(r_{\alpha\beta})$ is electron density from atom β to the site of atom α. Empirical functions $\phi(r_{\alpha\beta})$, $V(r_{\alpha\beta})$ and $F(\rho_\alpha)$ are fitted to experimentally measured bulk material properties, such as equilibrium lattice constants, sublimation energies, elastic constants and vacancy formation energy. The equation for the force $\overrightarrow{F_\alpha}$ on atom α is the derivative of Equation (2) and is given as

$$\overrightarrow{F_\alpha} = -\sum_\beta \left[\frac{1}{2} V'_{\alpha\beta} + F'_\alpha \rho'_{\alpha\beta} + F'_\beta \rho'_{\beta\alpha} \right] \overrightarrow{r_{\alpha\beta}}, \tag{3}$$

where F'_α is the derivative of atom α's embedding function F, and $\rho'_{\alpha\beta}$ is the derivative at atom α of the electron density due to atom β. Again, the summations are over all the atoms β that interact with the atom α. The stress used in this study is the viral stress, which is defined as

$$\sigma_{ij}^\alpha = -\frac{1}{2\Omega^\alpha} \left(\sum_\beta F_i^{\alpha\beta} r_j^{\alpha\beta} \right). \tag{4}$$

where Ω^α is the atomic volume of atom α, $F_i^{\alpha\beta}$ is the i-component of the force between atoms α and β obtainable from the derivative of the potential $\Phi(r)$ by Equation (3), and $r_j^{\alpha\beta}$ is the j-component of the separation of atoms α and β.

Let define the stress concentration factor (SCF) as

$$SCF = \frac{\sigma_{11}}{\overline{\sigma_{11}}}. \tag{5}$$

where σ_{11}, $\overline{\sigma_{11}}$ are the local stress at the right bottom corner atom on the front surface and $\overline{\sigma_{11}}$ is the average stress of the whole volume of the simulating slabs. During the loading process, the dimension of the molecular statics cell is uniformly expanded in x_1 direction, while the dimensions in the x_2 and x_3 directions are kept constant.

3. RESULTS AND DISCUSSIONS

The stress concentration factor as a function of characteristic size, a, is depicted in Figure 3. It can be seen that as the characteristic size decrease, the stress concentration factor decreases. A 37% decrease is observed in our simulation. This trend is in contrast with the that obtained by Sharma et al. In their work, Sharma et al. used surface material constants in the [100] crystal orientations and predicted stress concentration factor increase with decreasing characteristic size.

As recently noted, as the surface Young's modulus is smaller than the value of its bulk part in the [100] direction while the surface Young's modulus is larger than the value of its bulk part in the [110] direction. In order to study

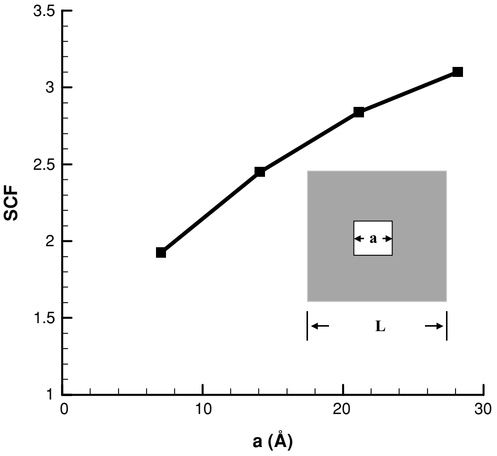


Figure 3. SCF as a function of characteristic.

whether stress concentration factor can increase with decreasing characteristic size in different crystal orientations, simulations are also conduct for FCC Ni with x_1, x_2 and x_3 in the [110], [011] and [001] directions. Since it is difficult to build square nanocavities, we construct rectangular nanocavities with length as L_1 and L_2. Effective characteristic size R is also introduced in such a way that $\boldsymbol{R}^2 = \boldsymbol{a}^*\boldsymbol{b}$. The stress concentration factor as a function of characteristic size, a, is depicted in Figure 4. Again, it can be seen that as the characteristic size decrease, the stress concentration factor decreases.

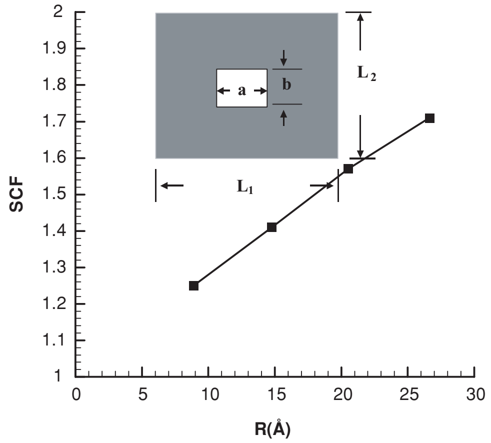


Figure 4. SCF as a function of size (in [100] crystal orientation) characteristic size (with [110] crystal orientation).

4. CONCLUSIONS

Size-dependent stress concentration of nanocavities was investigated with EAM (embedded-atomic potentials) based molecular statics simulation. It was found that at nanoscale stress concentration factor of nanocavities decreases as the characteristic size decreases and this trend is not affected by grain orientations.

REFERENCES

1. H. Zhang and C.T. Sun (2002), Semi-continuum model for plate-like nanomaterials. *AIAA/ASME/ASCE/AHS Structures, Structural Dynamics, and Materials Conference*, 2002-1316, Denver, Colorado, USA, pp. 1–11.
2. H. Zhang and C.T. Sun (2004), Nanoplate model for plate-like nanomaterials. *AIAA Journal*, 42, pp. 2002–2009.
3. C.T. Sun and H. Zhang (2003), Size-dependent elastic moduli of plate-like nanomaterials. *Journal of Applied Physics*, 93, pp. 1212–1218.
4. P. Sharma, S. Ganti and N. Bhate (2003), Effect of surfaces on the size-dependent elastic state of nano inhomogeneities. *Applied Physics Letters*, 82, pp. 535–537.

ATOMISTIC SIMULATION ON THE STIFFENING AND SOFTENING MECHANISM OF NANOWIRES

H.Y. Liang,[1] G.R. Liu[1] and X. Han[2]
[1]*Centre for ACES, Department of Mechanical Engineering, National University of Singapore, 9 Engineering Drive 1, Singapore 117576*
[2]*College of Mechanical and Automotive Engineering, Hunan University, Changsha 410082, P.R. China*

Abstract This paper studies the stiffening and softening of copper nanowire due to surface effect along different crystallographic orientations. Using molecular statics, nanowires of [001], [110] and [111] directions, are found to be elastically stiffer or softer than its bulk counterpart. These opposite trends of stiffening and softening can be attributed to two factors: surface elasticity and surface stress induced non-linear elasticity in the core region. The surfaces of nanowires are found to be softer instead of stiffer than bulk copper.

Keywords: nanowire, stiffening, softening, atomistic simulation.

1. INTRODUCTION

Decreasing characteristic dimensions of materials to nanoscale often leads to size effect: a deviation of elastic properties from its macroscopic counterpart behaviour [1]. With the substantial development of nanotechnology [2, 3], nanomechanical structures, such as nanowire/plate, are no longer hypothetical. Using molecular dynamics simulation, nanowire and nanoplate of EAM tungsten are found to be softer along $\langle 001\rangle$ direction as the cross-section decreases [4]. Besides softening, stiffening is also possible in nanostructures. In an early attempt to study the size dependence of Young's modulus of EAM copper film, Streitz et al. [5] found softening in $\langle 001\rangle$ direction and stiffening in $\langle 110\rangle$ direction as the film thickness decreases. The opposite trends in Young's modulus are reconfirmed by atomistic calculations with EAM potentials for

G. R. Liu et al. (eds.), Computational Methods, 1667–1672.

aluminium [6] and copper [7]. It is well known that the elastic moduli of materials near surface are different from corresponding bulk values. This difference is regarded as the cause for the size effect [1,6,7]. For macroscopic solids, surface elasticity contributes little to the total elastic behaviour, while it turns pivotal at nanoscale where surface-to-volume ratio is extremely large. In addition to surface elasticity, surface stress induced non-linear elasticity could be another factor responsible for size effect. Actually, for free standing nanosized structures, surface stress can induce large contraction strain well beyond linear elastic region [4,8], or even phase transformation [9].

In this paper, atomistic simulations are employed to investigate the physical mechanism of size effect on the Young's modulus of copper nanowire.

2. SIMULATION PROCEDURE

Three types of copper nanowires of square cross-section (Figure 1) are set up with [001], [110] and [111] direction along axis X. For [001] nanowire, the two lateral free surfaces are (100) and (010), respectively. For [110] nanowire, the two lateral free surfaces are (001) and(1$\bar{1}$0), respectively. For [111] nanowire, the two lateral free surfaces are ($\bar{1}\bar{1}$2) and (1$\bar{1}$0), respectively. Different square cross-sections, with width ranging roughly from 1.3 to 15.0 nm, are simulated for each nanowire type. The width of nanowire is defined as the average of length along Y and Z-axes after relaxation. The cross-section of [001] nanowire is exactly square. But for [110] and [111] nanowires, we construct the cross-section as square as possible.

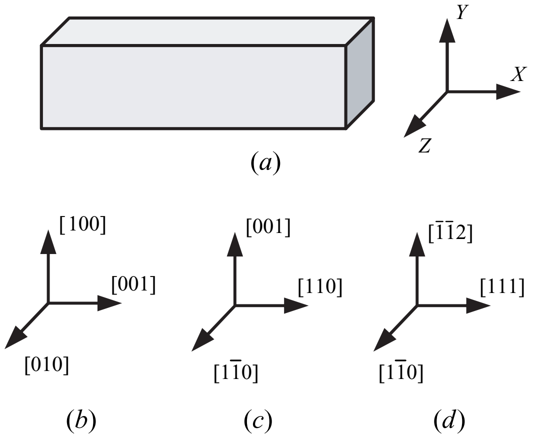


Figure 1. Configuration of nanowires. Axis X is the strain direction with PBC. Y and Z surface are traction free.

For all cases, periodic boundary conditions are applied along axis X, and the two lateral surfaces are traction free. The EAM potential for copper developed by Mishin et al. [10] is adopted. Strains in the range of [−5.00%, 1.00%] with an increment of 0.05% are applied along axis X. Conjugate gradient method is used to find the local energy minimum at each strain step. The Young's modulus is calculated via the second derivative of the total system energy with respect to strain: $E = (1/V)d^2U_{\mathrm{tot}}/d\varepsilon^2$, where U_{tot} is the total system energy, ε the strain. V is the volume of nanowire at the initial state.

3. RESULTS AND DISCUSSION

Due to the presence tensional surface stress, a copper nanowire will reach its energy minimum with a contraction along axis X. Figure 2 shows the strain energy changes with applied strain for two [001] nanowires with width of 3.5 and 2.0 nm, and for the bulk copper. The energy minimum shifts to the right as the width reduces. This trend is the same for [110] and [111] nanowires. Figure 3 demonstrates the contraction strain for nanowires of different width. The contraction behaviour of [110] and [111] nanowire are quite similar, while the contraction of [001] nanowire is much larger. The contraction strain of nanowires results mainly from the competition between the surface stress and Young's modulus in the core area. For example, the larger contraction strain of [001] nanowire could be attributed to the smallest Young's modulus (for bulk copper, $E_{001} < E_{110} < E_{111}$) and moderate surface stress [11].

Figure 4 shows the variation of Young's modulus with the width of [001], [110] and [111] nanowires. [001] and [111] nanowires exhibit softening, i.e.,

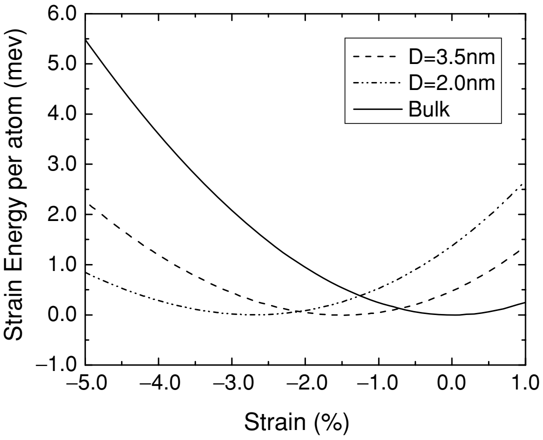


Figure 2. Strain Energy per atom for [001] nanowires and bulk copper.

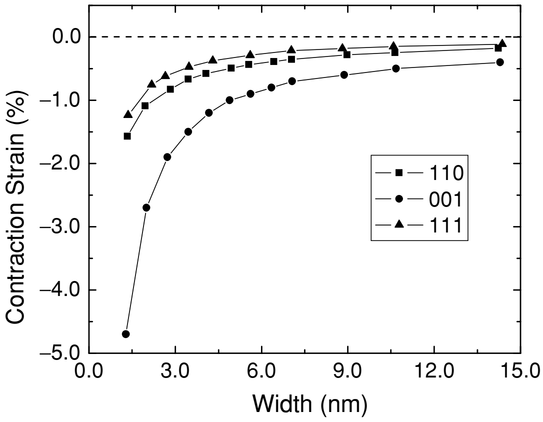


Figure 3. Contraction strain for all nanowires with different width.

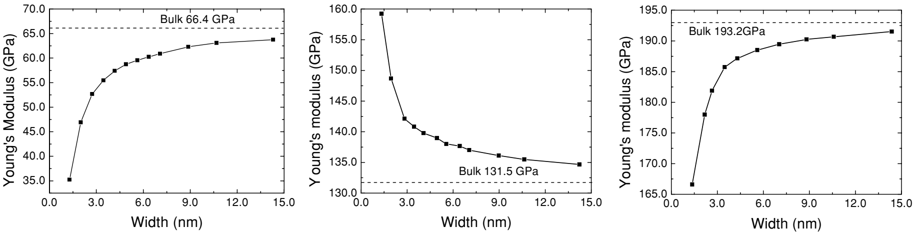


Figure 4. Young's modulus vs. width of [001], [110] and [111] nanowires.

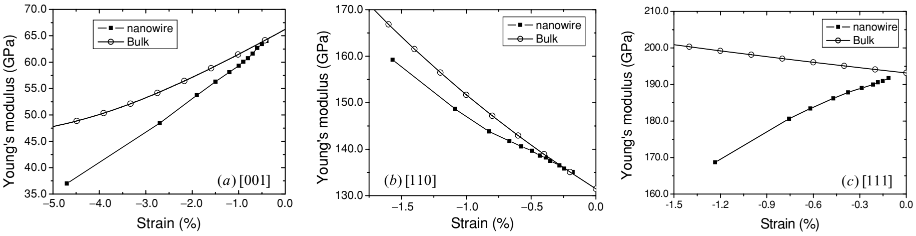


Figure 5. Comparison of Young's modulus between bulk and nanowires.

Young's modulus decreasing with smaller nanowire, while [110] nanowire shows stiffening. The present results are consistent with the findings by Streitz et al. [5] and Zhou and Huang [7]. To understand the mechanism of softening and stiffening, we calculate also the Young's modulus of bulk copper at different levels of uniaxial strain, and the results are plotted in Figure 5 together with those for nanowires for comparison. It is noted that though a high surface-to-volume ratio compared with macro counterpart, surface atom accounts for less than 50% even for the thinnest nanowire in our simulations, suggesting an important role of core region. Hence, the Young's modulus of [001] and [110] nanowires follow the trends of corresponding bulk copper (Figures 5a and b). Surprisingly, the Young's modulus of [111] nanowires goes down despite a weak increase of 3.5% (with strain from 0.0 to -0.015) for bulk copper (Figure 5c).

It is noted from Figure 5 that the Young's modulus of nanowires are always smaller than that of bulk copper, indicating a softer surface other than a stiffer surface. The surface softening is originated from the configurations of surface atoms with lower atomic coordinates. For [110] nanowire, the nearest neighbour atoms form continuous lines along axis X on $(1\bar{1}0)$and (001) surfaces. For [001] nanowire, the nearest neighbour atoms form 45° crossing on (100) and (010) surfaces. For [111] nanowire, due to the less atom interaction on $(1\bar{1}0)$ surface and especially no interaction for the outmost atoms (beyond the cutoff distance of EAM potential) along axis X on $(\bar{1}\bar{1}2)$ surface. It is expected that the surface elasticity are moderate for [001] and [110] nanowires and neglectable for [111] nanowire. Therefore, the surface softening and the non-linear elasticity at core region will determine whether a nanowire shows softening or stiffening.

4. CONCLUSIONS

Due to surface stress induced contraction, nanowires of [001], [110] and [111] directions, are found by molecular statics to be elastically stiffer or softer than its bulk counterpart. These opposite trends of stiffening and softening can be attributed to surface elasticity and surface stress induced non-linear elasticity in the core region. The surfaces of nanowires are always found to be softer instead of stiffer than bulk copper.

ACKNOWLEDGEMENT

This work is supported by the National Science Foundation of China under the grant number 10372031.

REFERENCES

1. S. Cuenot, C. Fretigny, S. Demoustier-Champagne and B. Nysten (2004), Surface tension effect on the mechanical properties of nanomaterials measured by atomic force microscopy. *Physical Review B*, 69, pp. 165410–165417.
2. M.A. Haque and M.T.A. Saif (2004), Deformation mechanisms in free-standing nanoscale thin films: a quantitative in situ transmission electron microscope study. *PNAS*, 101, pp. 6335–6340.

3. Y. Wu, J. Xiang, C. Yang and C.M. Liber (2004), Single-crystal metallic nanowires and metal/semiconductor nanowire heterostructures. *Nature*, 430, pp. 61–64.
4. P. Villain, P. Beauchamp, K.F. Badawi, P. Goudeau and P.-O. Renault (2004), Atomistic calculation of size effects on elastic coefficients in nanometre-sized tungsten layers and wires. *Scripta Materialia*, 50, pp. 1247–1251.
5. F.H. Streitz, K. Sieradzki and R.C. Cammarata (1990), Elastic properties of thin fcc films. *Physical Review B*, 41, pp. 12285–12287.
6. R.E. Miller and V.B. Shenoy (2000), Size-dependent elastic properties of nanosized structural elements. *Nanotechnology*, 11, pp. 139–147.
7. L.G. Zhou and H.C. Huang (2004), Are surfaces elastically softer or stiffer? *Applied Physics Letter*, 84, pp. 1940–1942.
8. R.C. Cammarata (1994), Surface and interface stress effects in thin films. *Progress in Surface Science*, 46, pp. 1–38.
9. J.K. Dai, K. Gall and M.L. Dunn (2003), Surface-stress-induced phase transformation in metal nanowires. *Nature Materials*, 2, pp. 656–660.
10. Y. Mishin, M.J. Mehl, D.A. Papaconstantopoulos, A.F. Voter and J.D. Kress (2001), Structural stability and lattice defects in copper Ab initio, tight-binding and embedded-atom calculations. *Physical Review B*, 63, pp. 224106–224122.
11. M. Schmid, W. Hofer, P. Varga, P. Stoltze, K.W. Jacobsen and J.K. Norskov (1995), Surface stress, surface elasticity, and the size effect in surface segregation. *Physical Review B*, 51, pp. 10937–10946.

THE HIGH STRAIN-RATE SCALE FOR NANOWIRES

H.A. Wu[1] and G.R. Liu[2]
[1]*CAS Key Laboratory of Mechanical Behavior and Design of Materials, University of Science and Technology of China, Hefei, Anhui 230027, China*
[2]*Department of Mechanical Engineering, Center for Advanced Computations in Engineering Science, National University of Singapore, 9 Engineering Drive 1, Singapore 117576, Singapore*

Abstract Three-dimensional molecular dynamics (MD) simulations of mechanical properties of copper nanowire have been carried out to study the size effects on the high strain-rate scale. It is found that the strain-rate scale of a nanosized structure is much higher than common materials. Even strain rate of 10^8/s can be considered as quasi-static loading rate for metal nanowires. The results from three different loading methods agree well to this finding, which proves that this high strain-rate scale can be true for small structures. The origin of such a high strain-rate scale is that structures with a length scale of nanometers can respond to the external loading very quickly.

Keywords: molecular dynamics, nanowire, strain rate, size effect, material property.

1. INTRODUCTION

Mechanical devices are shrinking in dimensions to reduce mass, increase resonant frequency and lower the force constants of these systems to achieve desired purposes. The thinnest metal Au nanowire with only four atomic rows was successfully prepared using electron-beam irradiation in an ultrahigh vacuum electron microscope [1]. The concept of using nanowires as building blocks for self-assembling logic and memory circuits was presented [2]. Metallic nanowires have received considerable attention in recent years, especially numerical simulation study by employing atomistic simulation methods, which can elucidate the details of deformation evolution by motion of atoms [3]. In this paper, our study will focus on the strain-rate scale of nanowires, and MD

G. R. Liu et al. (eds.), Computational Methods, 1673–1678.

simulations will be conducted to model the mechanical behaviour of copper nanowires under external loadings. This work is motivated by the controversies about the high strain-rate scale of nanowires. Our MD simulations will prove that this high strain-rate scale is true, and it is due to the size effect.

2. STRAIN-RATE EFFECT STUDY

We consider a rectangular nanowire cut from bulk single crystal copper, whose size is 2.0 nm × 2.0 nm × 5.4 nm. The initial atomic configuration is positioned at the fcc lattice sites. The X, Y, Z coordinate axes represent the lattice direction of [1 0 0], [0 1 0], [0 0 1], respectively. Periodic boundary condition is applied in the length direction, i.e., Z direction. The surfaces in X and Y directions are free. An embedded atom potential of copper [4], are employed to represent the atomic interactions. The temperature is kept constant at 0.01 K by using a direct velocity scaling method. The Gear algorithm [5], which is used to integrate Newton's equations of motion, uses up to the fifth time derivative of the atom position. After full relaxation of initial configuration, extension strain loadings are applied by uniformly scaling the z coordinates [3], which we call scaling method. The atoms of the outermost layer at each end are only constrained in the z direction during each loading step. The tensile strain step is 0.0001. It is relaxed during sometime in each step. Both the strain step and relaxation time determine the strain rate. The stress in the nanowire under above strain loading is computed by the average of atomic stress.

The different rates of tensile strain range from 10^7/s to 5×10^{10}/s. Figure 1 shows the plot of normalized yielding stress as a function of extension strain

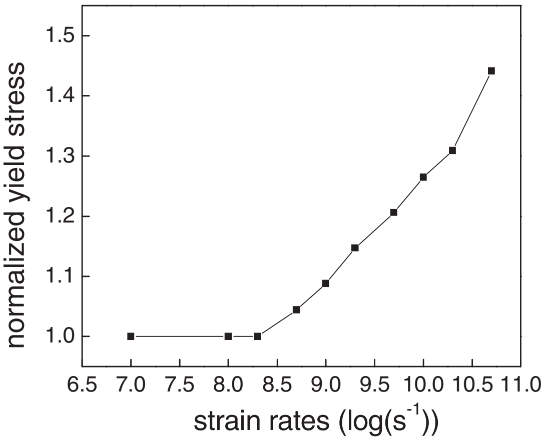


Figure 1. Yield varying with strain-rates.

rate. The strain-rate effect is similar to that in macroscale. However, the strain-rate scale is much different. It is found from the results that the strain rates of 10^8/s or below are just quasi-static loading rate for this nanowire.

This method is generally used in simulating extension deformation by many researchers. It is stable and efficient. However the scaling process seems to have introduced some artificial factors. Maybe we helped the atoms to move to their balanced positions. If so, the high strain rate will be meaningless or false. In our present work, we use another two methods, force method and displacement method, to apply extension loading to verify this high strain-rate scale.

3. FORCE AND DISPLACEMENT LOADING METHODS

For the force method, one end of the nanowire is fixed after free relaxation, and the extension forces are uniformly applied to one layer atoms at the other end. These forces drive the last layer atoms to move along the length direction. The next layer atoms will receive an attractive force due to bigger distances between the last two layer atoms, so they will also move in the same direction. Thus, all atoms except the fixed ones will move along the length direction. The result is that extension deformation occurs. The force loading is also applied step-by-step. During each step, the simulation cell is fully relaxed before next loading step. Because the length of the simulation cell changes with loading and is not predetermined, periodic boundary condition (PBC) cannot be applied in length direction. The simulation cell is a finite length copper nanorod, with the size of 2.0 nm × 2.0 nm × 200.3 nm. Three loading steps are applied, with each step of 3.94 nN. The force is uniformly distributed to 61 atoms of the last layer.

We can calculate that the induced extension stress step in nanowire would be 1.0 GPa. The strain can be easily obtained from the displacement of last layer atoms and the simulation cell length at free relaxation state. Figure 2 shows the plots of strain as a function of time. It is can be found that during the 20 ns of each loading step, resulted strain converge to stable value. The elastic modulus can be obtained by stress divide by strain, about 66 GPa, which agrees well with the scaling method results. In current simulation, the nanorod is fully relaxed during each loading step so that a stable state can be reached. This loading speed can be considered as quasi-static loading. On the other hand, we can calculate the corresponding strain rate from strain and loading step time, which is about 10^6/s.

For the displacement loading method, one end of the nanorod is fixed after free relaxation, and the enforced displacements along the length direction are uniformly applied to one layer atoms at the other end. The next layer atoms will receive an attractive force due to bigger distances between the last two

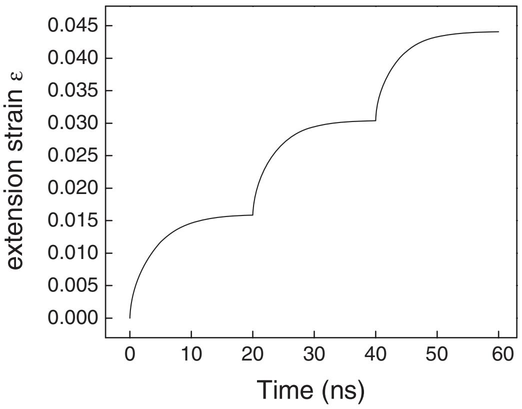


Figure 2. Strain varying with time.

layer atoms, so they will also move in the same direction. Thus, all atoms except the fixed ones will move along the length direction. The displacement loading is also applied step-by-step. During each step, the simulation cell is fully relaxed before next loading step with two layer atoms fixed at either ends. The displacement increment of each step is 0.05 nm, so the strain step is 0.25%. In force loading method, we input stress and output strain to get strain–stress relation. In displacement loading method, we input strain and output stress. To ensure the assumption of quasi-static loading, the nanorod is fully relaxed during each step. The relaxation time is 0.5 ns for above strain step of 0.25%, so the strain rate is 5×10^6/s.

Figure 3 shows the stress plot with time. It can be found that during the elastic stage, the stress converges during each step quickly. This further proves that 5×10^6/s is a comparatively low strain rate for elastic deformation of copper nanowire. The elastic modulus obtained in this simulation is almost the same as those obtained by scaling method and force method. However, the yielding strain and yielding stress are some different. The yielding strain and stress are 0.10 and 7.5 GPa, respectively, from the figure. It is also illustrated that the reactions to external strain rates of plastic deformation and elastic deformation are different. A 0.5 ns is long enough for elastic strain of 0.25%, but not for same plastic strain.

The simulation results show that different loading models have no significant influence on the elastic behaviours of metal nanowire, but have some influence on yielding behaviour. Scaling method of applying strain loading is very stable and efficient, and the strain rate in this method is still meaningful. All results prove that the strain-rate scale of nanostructures is different from that of macro-materials and structures. The former is much higher. This is due

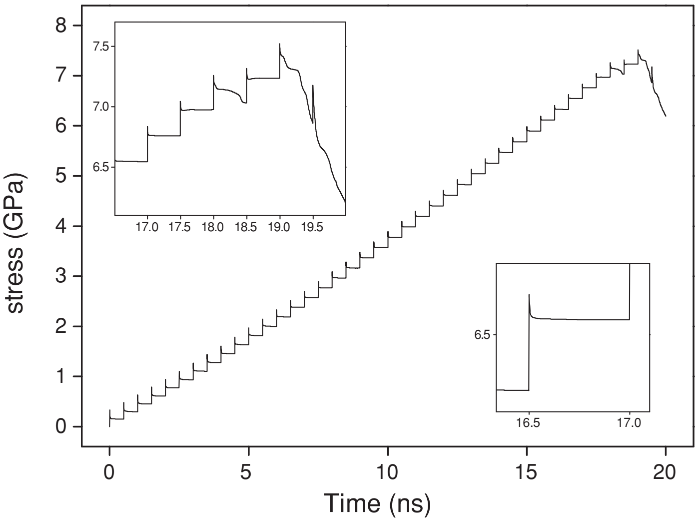


Figure 3. Stress plot as a function of time.

to the much smaller length size scale, which permits the structure to respond external loading very quickly.

4. CONCLUSIONS

Molecular dynamics simulations with three different loading methods, scaling method, force method and displacement method, to apply extension on metal nanowire have been presented in this paper. The results shows that the mechanical behaviours of nanowires are strain-rate dependent. However, the strain-rate scale is different from that of macrostructures. Strain-rate loading of 10^6–10^8/s can be considered as quasi-static loading, which are super high strain rates for macro-materials and structures. This high strain-rate scale is verified by different loading methods. The origin of such high strain-rate scale is that the nanowire can respond external loading very quickly because of its small size of nanometers.

REFERENCES

1. Y. Kondo and K. Takayanagi (1997), Gold nanobridge stabilized by surface structure. *Physical Review Letters*, 79, 18, pp. 3455–3458.
2. N.I. Kovtyukhova and T.E. Mallouk (2002), Nanowires as building blocks for self-assembling logic and memory circuits. *Chemistry-A European Journal*, 8, 19, pp. 4355–4363.

3. T. Kitamura, K. Yashiro and R. Ohtani (1997), Atomic simulation on deformation and fracture of nano-single crystal of nickel in tension. *JSME International Journal Series A-Solid Mechanics and Material Engineering*, 40, 4, pp. 430–435.
4. M. Doyama and Y. Kogure (1999), Embedded atom potentials in fcc and bcc metals. *Computational Materials Science*, 14, 1–4, pp. 80–83.
5. M.P. Allen and D.J. Tildesly (1987), *Computer Simulation of Liquids*. Oxford University Press, New York.

VOXEL BASED RIGID BODY DYNAMICS FOR COMPUTER GRAPHICS

K. Suzuki, J. Kubota and H. Ohtsubo
Department of Environmental Studies, The University of Tokyo, 7-3-1 Hongo,
Bunkyo-ku, Tokyo, Japan 113-8656

Abstract For the rigid body simulation, robust and efficient rigid body dynamic simulation algorithm is developed. The rotation of rigid body is represented using quaternion and volume based collision detection algorithm is developed. The rigid body is subdivided into voxel, and collision detection is carried out with spheres that include voxels. The strategy for determining time step was described. Several examples are shown to demonstrate the efficiency, robustness and applicability to the non-convex objects of the method, and compared with polygon based method. It was shown that by changing the level of sphere subdivision, it was possible to control the accuracy and computational efficiency.

Keywords: rigid body simulation, quaternion, voxel, collision detection.

1. INTRODUCTION

For the simulation of rigid body dynamics, DEM (Distinct Element Method) developed by Cundall [1] and DDA (Discontinuous Deformation Analysis) developed by Shi [2] which considers deformation of body are two major methods. These methods have been applied to the rock fall analysis but most analysis uses two dimensional model, which is quite different from real situations. However, the extension of rigid body dynamics to three dimensions is quite difficult in algorithm and computational load.

On the other hand, in the field of computer graphics almost all computation is carried out in three dimensions. These days more and more physics based simulations are used in creating motions of realistic animations. In this paper, we propose to use the algorithm based on the rigid body dynamics technology that is used in the field of computer graphics, and considers the possibility of combining them in the field of engineering simulations.

G. R. Liu et al. (eds.), Computational Methods, 1679–1687.

Most difficult part of the rigid body dynamics is collision detection. Jimenez et al. [3] gives good survey of the collision detection technique. Traditional collision detection often fails and penetration of objects occurs quite often. Also another problem is collision detection of non-convex objects. For the collision detection of convex objects, GJK (Gilbert–Johnson–Keerthi) algorithm [4] is quite effective. However, for non-convex objects the objects need to be divided into several convex objects to enable collision detection, which cannot be done automatically.

Bounding box and oriented bounding box approach approximate the objects by rectangle or oriented rectangle and are good for very rough approximation but not accurate collision detection. OBBTree [5] is another volume based collision detection technique using oriented bounding box in hierarchical way. Dingliana and O'Sullivan [6] proposed sphere tree based collision detection technique that is good balance of accuracy and efficiency. However, their application is limited to primitive 2D examples. In this paper, we are going to implement the sphere tree based algorithm into three dimensional rigid body dynamics and verify the accuracy, efficiency and robustness in detail.

In Chapter 2 we are going to describe rigid body dynamics using quaternion, which is commonly used in computer graphics to describe three dimensional rotations. In Chapter 3 we are going to propose volume based collision detection technique. Finally we are going to show several examples to demonstrate the efficiency and robustness of the method. The proposed collision detection method can be utilized in either DEM or DDA.

2. RIGID BODY DYNAMICS

The rigid body dynamics in three dimensions is described in the following manner. State variables for single rigid body are defined as follows [7].

$$\boldsymbol{y}(t) = \begin{Bmatrix} \boldsymbol{x}(t) \\ \boldsymbol{q}(t) \\ P(t) \\ L(t) \end{Bmatrix} \quad \frac{d\boldsymbol{y}(t)}{dt} = \begin{Bmatrix} \boldsymbol{v}(t) \\ \omega(t)\boldsymbol{q}(t) \\ m\dot{\boldsymbol{v}}(t) \\ I\dot{\omega}(t) \end{Bmatrix} \tag{1}$$

where $\boldsymbol{x}(t)$ is position vector at centre of gravity, $\boldsymbol{v}(t)$ is velocity, $\omega(t)$ is angular velocity and $\boldsymbol{q}(t)$ is quaternion [8] that describe the rotation of rigid body. $P(t)$ is momentum vector and $L(t)$ is angular momentum vector. From the equation of motion, we have following relations.

$$\begin{aligned} \boldsymbol{F}(t) &= m\dot{\boldsymbol{v}}(t) \\ \tau(t) &= I\dot{\omega}(t) \end{aligned} \tag{2}$$

where $\boldsymbol{F}(t)$ is external force and $\tau(t)$ is external moment, m is mass of rigid body and I is moment of inertia matrix. By integrating this differential equation over time domain, we obtain the state of rigid body at each time step. In this paper, we employed explicit Euler time integration scheme.

3. VOXEL BASED COLLISION DETECTION

In the traditional collision detection methods, the collision between polygons are handled by detecting the intersection between vertex vs. face, and edge vs. edge. This process is quite complicated algorithm and sometimes fails to detect the collision that results in the penetrations of objects. Also different type of collisions results into different evaluation of collision force, which again complicates the simulation algorithm. These difficulties reduce the robustness and computational efficiency. Here, we employed the concept of using sphere tree for collision detection [6].

An object is originally defined in the form of polygon. As preprocess, we define voxel subdivision of the polygon in different level as shown in Figure 1. Sphere that involves each cube is defined on the voxel. We call the voxel that defines single cube for an object as level 0, and subdivide each edge by two to obtain one level higher subdivision. The collisions between objects are judged by the collision of spheres. By increasing level we get more accuracy in shape representation of objects, hence more accuracy in collision detection, with more computational time. By this way, we can control the accuracy and computational efficiency easily.

For the sphere of level 0, a sphere that includes the object is used instead of spheres that include the cube, which gives better accuracy. The algorithm of collision detection is as follows. By obtaining the distance between centres of level 0 spheres of objects and comparing them with the radiuses of them, we can check the possibility of collision between two objects. If there is collision between spheres, we increase the level of subdivision of both objects and check the collision between spheres, until we reach the maximum level of subdivision we defined. By subdividing the only spheres that collide each other as shown in Figure 2, the computational efficiency can be increased dramatically. When the collision is detected, the collision forces need to be computed to obtain the motion after collision. Since the forces acts in quite short time, we consider them as impulse force, and consider to act at the mid-point of the centres of two colliding spheres. When we take the centre of each sphere as $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, the direction of impulse force becomes as follows when no friction is acting, as shown in Figure 3. When friction is acting, the friction force is applied by multiplying coefficient of friction to the normal force. We denote $-$ as before collision and $+$ as after collision, the relative velocity of spheres v_{rel}^- (before

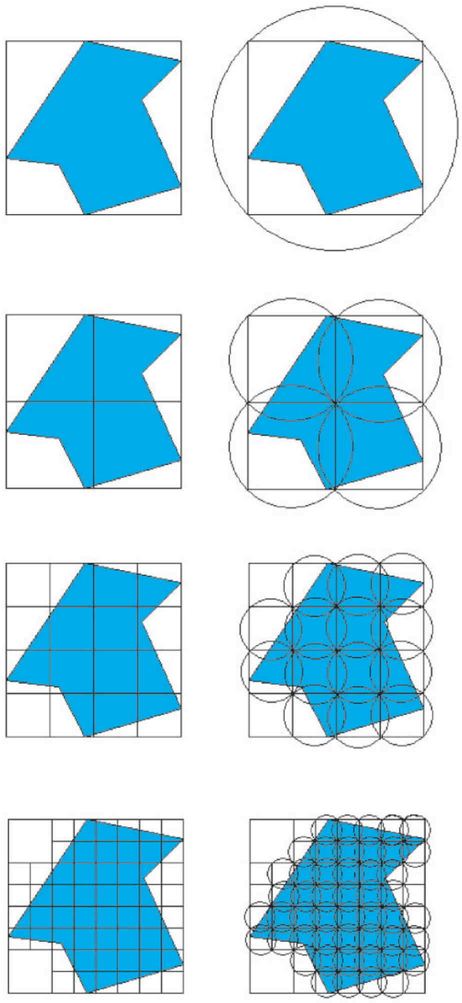


Figure 1. Voxel subdivision.

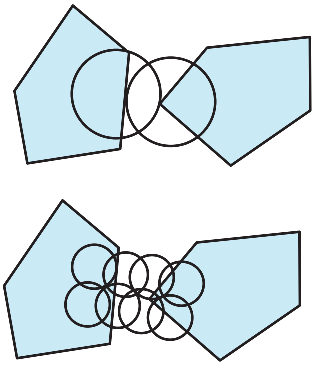


Figure 2. Subdivision of colliding sphere.

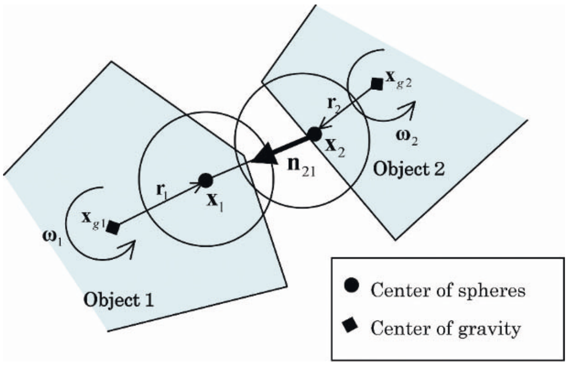


Figure 3. Direction and position of collision force.

collision) and v_{rel}^{+} (after collision) can be calculated as follows, where n is direction of one colliding sphere to another.

$$v_{rel}^{-} = \boldsymbol{n}_{21} \cdot (\dot{\boldsymbol{x}}_1^{-} - \dot{\boldsymbol{x}}_2^{-}) \qquad v_{rel}^{+} = \boldsymbol{n}_{21} \cdot (\dot{\boldsymbol{x}}_1^{+} - \dot{\boldsymbol{x}}_2^{+}) \tag{3}$$

They have following relation using restitution coefficient ε.

$$v_{rel}^{+} = -\varepsilon v_{rel}^{-} \tag{4}$$

The angular velocity $\boldsymbol{\omega}$ can be considered as follows. The velocity of the centre of sphere is described by the position of the centre of gravity $\boldsymbol{x}_g$ and vector r of the centre of sphere from the centre of gravity.

$$\boldsymbol{x}_1^{-} = \boldsymbol{x}_{g1}^{-} + \boldsymbol{\omega}_1^{-} \times \boldsymbol{r}_1 \quad \boldsymbol{x}_1^{+} = \boldsymbol{x}_{g1}^{+} + \boldsymbol{\omega}_1^{+} \times \boldsymbol{r}_1 \tag{5}$$

When we denote the change of the momentum as j and mass of object as M, and moment of inertia matrix as I, we obtain following relations.

$$\boldsymbol{x}_{g1}^{+} = \boldsymbol{x}_{g1}^{-} + \frac{j\boldsymbol{n}_{21}}{M_1} \tag{6}$$

$$\boldsymbol{\omega}_1^{+} = \boldsymbol{\omega}_1^{-} + I_1^{-1}(\boldsymbol{r}_1 \times j\boldsymbol{n}_{21}) \tag{7}$$

By substituting Equations (6) and (7) into Equation (5) we can obtain $\boldsymbol{x}_1^{+}$ from $\boldsymbol{x}_1^{-}$. From Equations (3) and (4), we can obtain j as follows.

$$j = \frac{-(1+\varepsilon)v_{rel}^{-}}{\dfrac{1}{M_1} + \dfrac{1}{M_2} + \boldsymbol{n}_{12}\left\{\left(\boldsymbol{I}_1^{-1}(\boldsymbol{r}_1 \times \boldsymbol{n}_{12})\right) \times \boldsymbol{r}_1 + \left(\boldsymbol{I}_2^{-1}(\boldsymbol{r}_2 \times \boldsymbol{n}_{12})\right) \times \boldsymbol{r}_2\right\}} \tag{8}$$

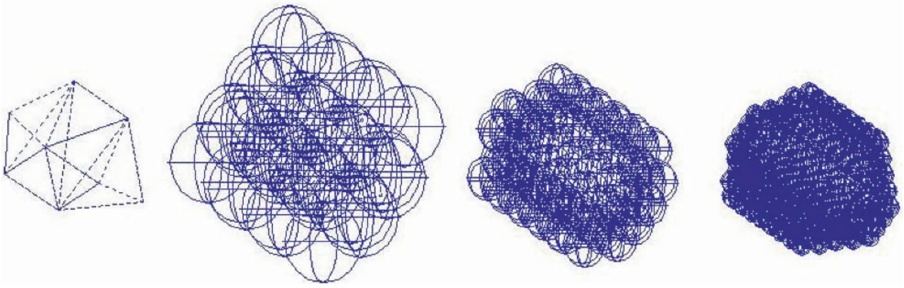


Figure 4. Shape of objects (Example 1).

4. EXAMPLES

4.1 Comparison with polygon base method

For the second example the projection of objects of shape shown in Figure 9 (left) were used for comparison with traditional method of polygon base collision detection. Figure 9 shows from left to right, original shape with 12 polygons, sphere subdivision of level 2, 3, 4 respectively. From time 0 to 4 s 10 objects are projected from height 1 m with horizontal speed 2 m/s and vertical speed 0.5 m/s and simulation during 5 s are carried out. With level 4 Equation (21) gives $\Delta t_{max} = 1/113$ (s), and $\Delta t = 1/120$ (s) was used for each simulation. In traditional method there is no indicator to determine time step so various Δt was used.

In the traditional method Δt need to determine empirically to obtain the results with no penetration. With $\Delta t = 1/1440$ and 1/1920, the penetrations of objects have been observed. With $\Delta t = 1/2400$ a result with no penetration was obtained. In voxel base method no penetration occurred even with $\Delta t = 1/120$ in each level. The solution time for each case is shown in Figure 5. The reason

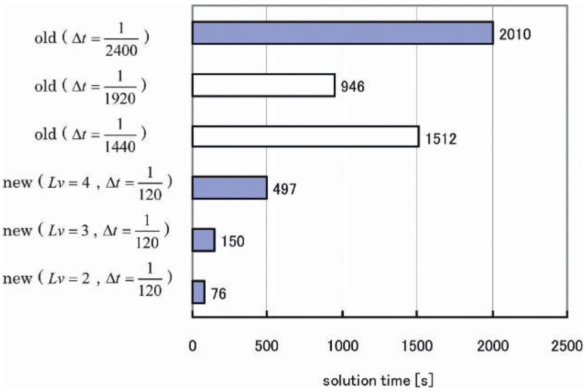


Figure 5. Comparison of solution time.

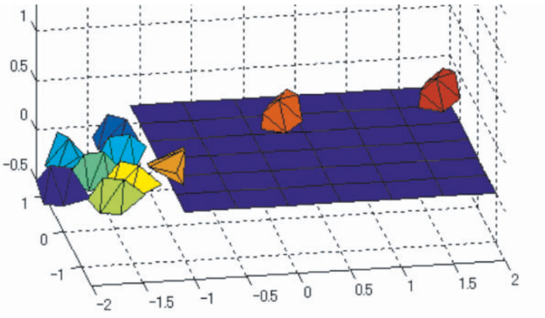


Figure 6. Halfway view.

why it takes more solution time with $\Delta t = 1/1440$ than $1/1920$ is because of penetration, the number of collision detection increases. It was shown that the collision detection algorithm described in this paper is effective in both robustness and computational time. The view in the mid of simulation is shown in Figure 6.

4.2 Non-convex objects

One big advantage of this new method is handling of non-convex objects is easy, while in the polygon base method handling of non-convex objects takes several times more computational time. In this example the cross shape objects shown in Figure 7 is used. The voxel subdivision is shown in Figure 2. 26 objects of this shape were dropped from the height 1.5 m from time 0 to 12 s in even interval and simulation until time 18 s was carried out. Solution time was about 1500 s.

The halfway view of the simulation process is shown in Figure 8.

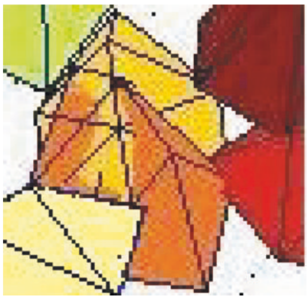


Figure 7. Example of penetration.

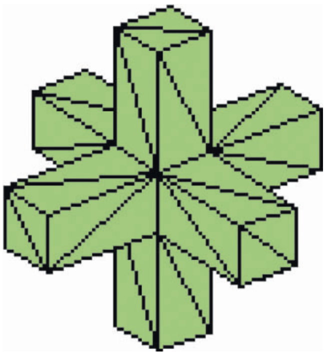


Figure 8. Non-convex objects.

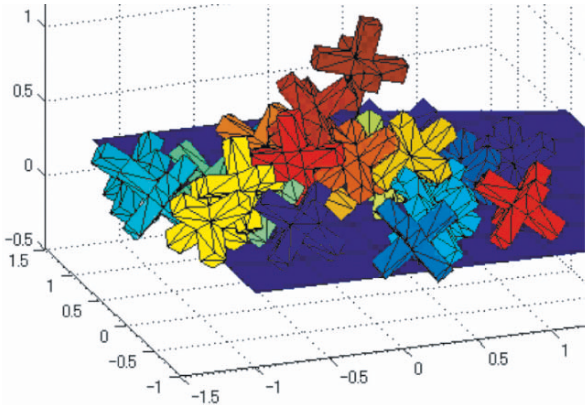


Figure 9. Halfway view of simulation process.

5. CONCLUSIONS

For the simulation of dynamics of discrete system, we have presented the robust and fast algorithm. The rotation of a rigid body is represented using quaternion and collision detection algorithm based on the volume sphere is developed. By changing the level of the subdivision, control of accuracy and efficiency is quite easy. Also the simple algorithm to determine time step to prevent penetration is shown for robust simulation. Several examples are shown for verification, and it was shown that new method is much robust and fast, and it can handle non-convex objects easily.

REFERENCES

1. P.A. Cundall (1979), A discrete numerical model for granular assemblies. *Geotechnique*, 29, 1, pp. 47–65.

2. G.H. Shi (August 1989), Block system modeling by discontinuous deformation analysis. Department of Civil Engineering, University of California, Berkeley.
3. P. Jimenez et al. (2001), 3D collision detection: a survey. *Computers and Graphics* 25, pp. 269–285.
4. E.G. Gilbert, D.W. Johnson and S.S. Keerthi (1988), A fast procedure for computing the distance between complex objects in three dimensional space. *IEEE Journal of Robotics and Automation*, 4, pp. 193–203.
5. S. Gottschalk, M.C. Lin and D. Manocha (1996), OBBTree: a hierarchical structure for rapid interference detection. In: *Proceedings of 23rd Annual Conference on Computer Graphics and Interactive Techniques*.
6. J. Dingliana and C. O'Sullivan (2000), Graceful degradation of collision handling in physically based animation. *Computer Graphics Forum*, 19, 3, pp. 239–248; R. Hill (1965), A self-consistent mechanics of composite materials. *Journal of Mechanics and Physics of Solids*, 13, pp. 213–222.
7. D. Baraff (1999), Physically based modeling. In: *SIGGRAPH '99* (Course Notes).
8. K. Shoemake (1985), Animating rotation with quaternion curves. *ACM SIGGRAPH Computer Graphics*, 19, pp. 245–254.

COMPLEXITY OF CABLE DYNAMICS

Xu Daolin and Guo Yufeng
School of Mechanical Engineering and Production, Nanyang Technological University,
Nanyang Avenue, Singapore 639798

Abstract Non-linear cable dynamics is investigated under a harmonic excitation. A set of non-linear partial differential equations is derived by the extended Hamilton's principle. Global bifurcation cascades are provided to show the multiple solution structures with the variation of forcing amplitude. The ultimate steady state of the cable is sensitively dependent on initial conditions and evolution history of parameters. The coexistence of different solutions may induce a sudden jump in response under small perturbations, a new feature found in non-linear cable dynamics.

Keywords: bifurcation, chaos, multiple solution, sudden jump.

1. INTRODUCTION

Cable with its outstanding flexibility and strength has been widely used in engineering. Exposing to turbulent surrounding mediums (wind or wave), cable experiences violent oscillation that degrades the system performance. Within the framework of linear theory, cable dynamics has been well understood. However, we know a little about non-linear dynamics of cable, due to the difficulty in dealing with strong geometrical non-linearity and intensive coupling effects as well as sensitive dependence on initial conditions and parameters. In the last decade, a remarkable progress has been made forward with a number of literatures [1] documented for non-linear cable dynamics. Some basic non-linear behaviours such as harmonic, sub-harmonic and chaotic motions have been reported in numerical studies [2] and experimental observations [3]. Nevertheless, the richness of non-linear cable dynamics has not been fully explored. Some peculiar phenomena were observed but not well explained. The existing knowledge about this topic is distant from the level of understanding.

G. R. Liu et al. (eds.), Computational Methods, 1689–1694.

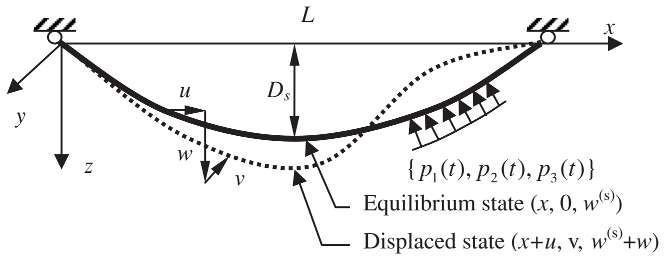


Figure 1. Schematic configuration of a suspended cable.

2. FORMULATION OF GOVERNING EQUATIONS

Figure 1 shows a schematic configuration of a suspend cable. $w^{(s)} = w^{(s)}(x)$ describes the static equilibrium. Under the excitation $p(t)$, the cable is displaced by a shift (u, v, w) from a reference point $(x, 0, w^{(s)})$, where $u = u(x, t)$, $v = v(x, t)$ and $w = w(x, t)$, L is the chord length of the cable. Based on Hamilton's principle, the governing equation of the cable in three dimensions can be expressed by the three partial-differential equations as

$$\begin{cases} m\ddot{u} = p_1 - \mu_1\dot{u} + [(EA\varepsilon + T)(1 + \mu_{,x})]_{,x} \\ m\ddot{v} = p_2 - \mu_2\dot{v} + [(EA\varepsilon + T)\nu_{,x}]_{,x} \\ m\ddot{w} = p_3 + mg - \mu_3\dot{w} + [(EA\varepsilon + T)(w_{,x} + \partial w^{(s)}/\partial x)]_{,x} \end{cases} \tag{1}$$

where m denotes the mass per unit length and the dot indicates the derivative with respect to time t; p_1, p_2, p_3 are the load intensity in x, y, z; μ_1, μ_2, μ_3 are the viscous damping coefficients; E is the elastic modulus and A is the cross-sectional area of the cable; T is the tension of the cable in equilibrium. The boundary conditions are $u(0, t) = v(0, t) = w(0, t) = u(L, t) = v(L, t) = w(L, t) = 0$. Consider an in-plane vibration of the cable with the assumption of $|v'| \ll 1$, $|w'| \ll 1$, $|u'| = O^2(v', w')$, $\dot{u} \ll 1$. Introduce dimensionless quantities $W = \frac{w}{D_s}$, $W^{(s)} = \frac{w^{(s)}}{D_s}$, $\eta = \frac{x}{L}$, $\tau = t/(\frac{L^2}{D_s}\sqrt{\frac{\rho}{E}})$. The in-plane governing equations become

$$\varepsilon = \frac{D_s^2}{L^2}\int_0^1 [W^{(s)}_{,\eta} W_{,\eta} + \frac{1}{2}(W^2_{,\eta})]\, d\eta = \frac{D_s^2}{L^2} e$$

$$W_{,\tau\tau} + \alpha_1 W_{,\tau} - (e + \beta)(W_{,\eta\eta} + W^{(s)}_{,\eta\eta}) = \gamma(1 + f(\eta, \tau)) \tag{2}$$

Using the Galerkin's method, the vibration solution can be expressed by $W(\eta, \tau) = \sum_{i=1}^{\infty} T_i(\tau)\sin(i\pi\eta)$. The governing Equations (2) can be converted into ODE:

$$\ddot{T}_k + \alpha_1 \dot{T}_k + \pi^2 (e + \beta)(k^2 T_k + \delta_{k1}) = \int_0^1 \sin(k\pi\eta)\gamma(1 + f(\eta, \tau))\, d\eta$$

$$e = \frac{1}{4}\pi^2 \left[2T_1 + \sum_1^n i^2 T_i^2 \right] \tag{3}$$

where $\delta_{k1} = \begin{cases} 1, & k = 1 \\ 0, & k \neq 1 \end{cases}$, $k = 1, 2, \ldots, n$. $\alpha_1 = \frac{\mu_2 L^2}{m D_s}\sqrt{\frac{\rho}{E}}$, $\beta = \frac{T L^2}{D_s^2 E A}$, $\gamma = \rho g L^4/(E D_s^3)$, T_k is unknown time function, $f(\eta, \tau) = f_0 \cos(\omega\tau)$ and n is the number of modes used for approximation.

3. SOLUTION STRUCTURE OF NON-LINEAR CABLE DYNAMICS

Table 1 lists the parameter values used in numerical simulations. The sag-to-span ratio here is 1:20. Based on the parameters in Table 1, the quantities in (3) are calculated as $\alpha_1 = 0.04418$, $\beta = 0.0194$, $\gamma = 0.38249$. In simulations, 10 modes are used for numerical solutions, which can provide reasonably good approximation in analysis. The cable is under a periodic excitation $F = f_0 \cos(\omega\tau)$. Many typical non-linear behaviours of harmonic, sub-harmonic and chaotic motions are observed. The harmonic motion is named as a P1 orbit if its period is identical to the period of excitation, P2 orbit if its period is double of the period of excitation. These notations will be used to illustrate the complexity of cable dynamics in bifurcation diagrams.

Bifurcation is referred to as the case that the characteristic of the motion of a system is qualitatively changed as a parameter varies across a specific value. A bifurcation diagram provides a global picture of different types of motions

Table 1. Assignment of parameters.

Variable	Description	Value (SI)
E	Yong's modulus	1.5988×10^{11} N/m^2
A	Cross-sectional area	7.85×10^{-3} m^2
ρ	Cable density	7.8×10^3 kg/m^3
D_s	Static cable sag	5.0 m
L	Length of cable	100 m
T	Static cable tension	1.5×10^5 N
c_1	Damping coefficient	0.001 M (*mass*)

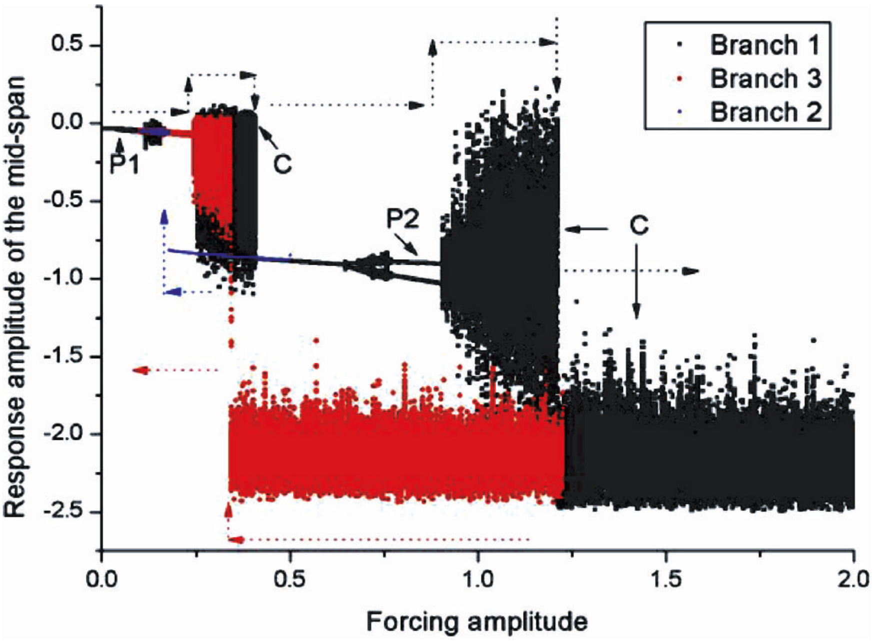


Figure 2. Global bifurcation diagram of the amplitude W_m (mid-span of the cable) against the variation of f_0, at a fixed forcing frequency $\omega = 3.1416$.

associated with a variation of parameters. It enables us to inspect the complete solution structure of cable dynamics and understand overall behaviours of the system.

Figure 2 shows the bifurcation diagram when the forcing amplitude varies in a range of $f_0 = 0.01 \sim 2.00$ with a step of 0.002. The points are sampled from the time evolution of W_m (mid-span of the cable) in every forcing period. Three branches of solution structures are illustrated. Branch 1 starts forwards at $f_0 = 0.01$ and merges with the Branch 2 at $f_0 = 0.4$. Branch 2 consists of P1 orbit ($f_0 = 0.01 \sim 0.65$) and P2 orbit ($f_0 = 0.65 \sim 0.90$) followed by chaotic motion ($f_0 = 0.90 \sim 1.21$). At $f_0 = 1.21$, Branch 2 merges with Branch 3 in which only chaotic motions (marked by C) are observed. Tracking the solution structure backwards from Branch 3 at $f_0 = 1.23$. The red dots show the chaotic solutions against the decrease of f_0. Branch 3 ends at $f_0 = 0.35$ where the solution structure jumps up to Branch 1. By tracking Branch 2 backwards from $f_0 = 0.5$, Branch 2 extends to left and ends at $f_0 = 0.17$, at which the response of the cable jumps to the P1 orbit of Branch 1. Note that two branches of solutions could coexist in some parameter regimes. It implies that the ultimate state of the cable could stay onto either two different solutions sensitively dependent on the initial conditions and the direction of the variation of the parameter f_0.

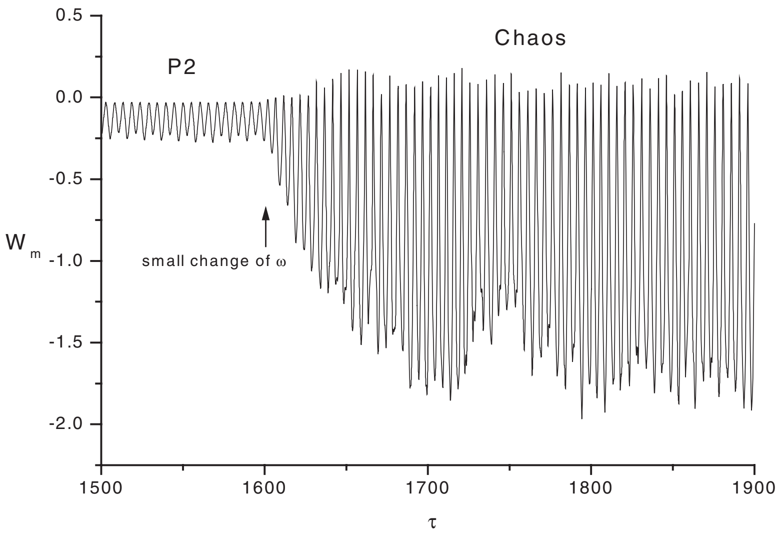


Figure 3. Time history of the motion of W_m of the cable with forcing frequency changed from 1.20 to 1.26 at $\tau = 1600$, $f_0 = 0.1$.

The phenomenon of sudden changes in oscillation amplitude is observed, referred to as a jump event. Figure 3 shows a jump when the forcing frequency ω is perturbed from 1.20 to 1.26 at time $\tau = 1600$. The amplitude of a P2 motion rapidly changes to a chaotic motion with an amplification about 10 times in amplitude. A small change of the input can cause a severe change in the output. This peculiar behaviour is very important to engineering applications. It happens usually without any indication, therefore hardly predictable. The mechanism behind the jump is that there coexist multiple stable solutions (as shown in Figure 2). These solutions are separated by basin boundaries. When a small perturbation applied to the cable, this disturbance may kick the system across the boundary from one solution basin into the other solution basin. Thus, such a transition in dynamical evolution between different solution basins may cause a significant change in response, resulting in a jump event.

4. CONCLUSION

In this paper, the governing equation for a suspended cable is presented. Complicated global bifurcation diagrams are illustrated, under the variation of forcing amplitude. We revealed the jump phenomena in non-linear cable dynamics. The jump event occurs due to a dynamical transition between different solution basins, which could be induced by small disturbances. Jump event could happen without any indications thus hardly predictable. This dynamical feature is important to engineering design because a sudden change in response can damage a system or degrade the system performance.

REFERENCES

1. P. Warnitchai, Y. Fujino and T. Susumpow (1995), A non-linear dynamic model for cables and it's application to a cable-structure system. *Journal of Sound and Vibration*, 187(4), pp. 695–712.
2. Rega (2003)
3. C.G. Koh, Y. Zhang and S.T. Quek (1999), Low-tension cable dynamics: numerical and experimental studies. *Journal of Engineering Mechanics*, pp. 347–354.
4. N. Srinil, S. Rega and S. Chucheepsakul (2003), Three-dimensional non-linear coupling and dynamic tension in the large-amplitude free vibrations of arbitrarily sagged cables. *Journal of Sound and Vibration*.

NUMERICAL SOLUTION OF ROBOT ARM MODEL USING STWS AND RKHEM TECHNIQUES

D. Paul Dhayabaran[1], E.C. Henry Amirtharaj[1], K. Murugesan[2]
and D.J. Evans[3]
[1]*Reader in Mathematics, Bishop Heber College, Tiruchirappalli 620017, Tamilnadu, India*
[2]*Department of Mathematics, National Institute of Technology, Tiruchirappalli 620015, Tamilnadu, India*
[3]*Nottingham Trent University, Nottingham, UK*

Abstract The primary objective of this paper is to study the parameters governing the robot arm model by way of finding the discrete solutions at different values of time 't' for the system of second order equations, which actually represents the dynamics of the arm model of the robot of two degree freedom. In this paper, it is also intended to demonstrate the effectiveness of the proposed numerical methods, Single Term Walsh Series (STWS) technique and extended RK method based on Heronian Mean (RKHeM), in order to find numerical solution for the robot arm model.

Keywords: numerical techniques, STWS, RKHeM, robot arm model.

1. INTRODUCTION

It is indeed true that a good number of researchers have contributed on a variety of aspects in the field of robust control, especially about the dynamics of robotic motion and their governing equations, for the past three decades. Krishnan and Mcclamroch [1] have dealt with the 'Applications of Non-linear Differential—Algebraic Control Systems to Constrained Robot Systems'. Warwick and Pugh [2] have dealt elaborately 'Robot Control Theory and Applications'. Huang and Tseng [3] have discussed 'Asymptotic Observer Design for Constrained Robot Systems'.

G. R. Liu et al. (eds.), Computational Methods, 1695–1699.

In this study, the authors have observed that the robotic motion has been governed by second order linear and non-linear differential equations. Hence a meticulous attempt has been made to study the parameters concerning the control of a robot arm model. In this paper, Single Term Walsh Series (STWS) technique and [5] the extended fourth order RK method based on Heronian mean (RKHeM) have been applied to find the numerical solutions of parameters governing the motion of a robot arm model.

2. ROBOT ARM MODEL

The dynamics of a robot arm is represented as

$$T = A(Q)\ddot{Q} + B(Q, \dot{Q})\dot{Q} + C(Q) \tag{1}$$

where $A(Q)$ = coupled inertia matrix, $B(Q, \dot{Q})$ = matrix of coriolis and centrifugal forces, $C(Q)$ = Gravity matrix, T = Input torques applied at various joints.

For a robot of two degree of freedom, under the assumption of lumped equivalent masses and mass less links, the dynamics are represented by

$$\begin{aligned} T_1 &= D_{11}\ddot{q}_1 + D_{12}\ddot{q}_2 + D_{122}(\dot{q}_2)^2 + D_{112}(\dot{q}_1\dot{q}_2) + D_1 \\ T_2 &= D_{21}\ddot{q}_1 + D_{22}\ddot{q}_2 + D_{211}(\dot{q}_1)^2 + D_2 \end{aligned} \tag{2}$$

For the set point regulation, the state vector is defined as

$$x = (x_1, x_2, x_3, x_4)^T = (q_1 - q_{1d}, \dot{q}_1, q_2 - q_{2d}, \dot{q}_2)^T \tag{3}$$

where q_1 and q_2 are the angles at the joint 1 and joint 2 respectively and q_{1d}, q_{2d} are constants.

In state space representation, Equation (2) can be written as

$$\begin{aligned} \dot{x}_1 = x_2; \quad \dot{x}_2 &= \frac{D_{22}}{d}(D_{122}x_2^2 + D_{112}x_2x_4 + D_1 + T_1) \\ &\quad - \frac{D_{12}}{d}(D_{211}x_4^2 + D_2 + T_2) \\ \dot{x}_3 = x_4; \quad \dot{x}_4 &= \frac{-D_{12}}{d}(D_{122}x_2^2 + D_{112}x_2x_4 + D_1 + T_1) \\ &\quad + \frac{D_{11}}{d}(D_{211}x_{44}^{22} + D_2 + T_2) \end{aligned} \tag{4}$$

The above system of equations is non-linear in nature and it is observed that a synthesis of the control law would be very difficult due to the non-linear and interactive nature of the canonical Equations (4).

3. REDUCTION OF ROBOT DYNAMICS TO A SECOND ORDER LINEAR SYSTEMS

Although the Physical and Mathematical structure of the complete dynamic robot model are analytically coupled and non-linear, it is observed that the transient responses of robot dynamics appear to resemble as transient responses of linear systems. Consequently, each joint of the robot can be characterized as a single-input, single-output system (SISO). The input is the actuator torque (or) force and the output is the joint position. Hence it is determined that the non-linear model Equations (4) of the two-link-robot arm model can be reduced to the following system of linear equations as

$$\begin{aligned} &\dot{x}_1 = x_2, \quad \dot{x}_2 = B_0 T_1 - A_1 x_2 - A_0 x_1 \\ &\dot{x}_3 = x_4, \quad \dot{x}_4 = B_0^2 T_2 - A_1^2 x_4 - A_0^2 x_3 \end{aligned} \tag{5}$$

The above system has been reduced to a system of two linear second order equations as

$$\begin{aligned} \ddot{x}_1 &= -A_1 \dot{x}_1 - A_0 x_1 + B_0 T_1 \\ \ddot{x}_3 &= -A_1^2 \dot{x}_3 - A_0^2 x_3 + B_0^2 T_2, \quad \text{where } x_2 = \dot{x}_1 \text{ and } x_4 = \dot{x}_3 \end{aligned} \tag{6}$$

i.e.,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_3 \end{bmatrix} = \begin{bmatrix} -A_1 & 0 \\ 0 & -A_1^2 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_3 \end{bmatrix} + \begin{bmatrix} -A_0 & 0 \\ 0 & -A_0^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} + \begin{bmatrix} B_0 & 0 \\ 0 & B_0^2 \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} \tag{7}$$

4. EXAMPLE

The values of the robot parameters used were $M_1 = 2$, $M_2 = 5$, $d_1 = d_2 = 1$. The input into the 'Black Box' is the response of $\ddot{y} + 2kn\dot{y} + n^2 y = n^2 y_s$. For $n = 5$ and $k = 1$ the critical damping, the initial conditions and the set points have been taken as

$$q_1(0) = 0,\ q_2(0) = 0,\ \dot{q}_1(0) = 0,\ \dot{q}_2(0) = 0,\ q_{1d} = 1 \text{ and } q_{2d} = 1 \tag{8}$$

The linear model parameters of joint 1 have been found as $A_0 = 0.1730$, $A_1 = -0.2140$, $B_0 = 0.0265$, and that of joint 2 have been determined as $A_0 = 0.0438$, $A_1 = 0.3610$, $B_0 = 0.0967$.

Equation (7) becomes

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_3 \end{bmatrix} = \begin{bmatrix} 0.2140 & 0 \\ 0 & -0.130321 \end{bmatrix}\begin{bmatrix} \dot{x}_1 \\ \dot{x}_3 \end{bmatrix} + \begin{bmatrix} -0.1730 & 0 \\ 0 & -0.00191844 \end{bmatrix}\begin{bmatrix} x_1 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0.0265 & 0 \\ 0 & 0.00935089 \end{bmatrix}\begin{bmatrix} T_1 \\ T_2 \end{bmatrix}, \tag{9}$$

where $x_2 = \dot{x}$ and $x_4 = \dot{x}_3$

Since $(x_1, x_2, x_3, x_4) = (q_1 - q_{1d}, \dot{q}_1, q_2 - q_{2d}, \dot{q}_2)$ and using Equation (9), the initial conditions are $x_1(0) = -1$, $x_3(0) = -1$, $\dot{x}_1(0) = 0$, $\dot{x}_3(0) = 0$.

In order to study the parameters that govern the dynamics of the robot using numerical methods, choose $T_1 = T_2 = 1$ unit. However, one can vary the values of T_1 and T_2 to visualize the effect of the parameters that control the arm model of the robot and the simulation can be done. The exact solution of the system (9) has been determined.

The proposed methods, Paul Dhayabaran [4], Single Term Walsh Series (STWS) technique and extended fourth order Runge-Kutta method based on Heronian Mean (RKHeM), have been applied for the example discussed above to determine the discrete solutions at different time 't' for any length of time. The exact and discrete solutions of the angles q_1 and q_2 at the joints 1 and 2 of the robot arm at different time 't' are given in the Table 1. The absolute error between the exact and discrete solutions of q_1 and q_2 are given in the Table 2.

Table 1. Angles at the joints.

			Discrete solution			
	Exact solution		STWS		RKHeM	
Time	q_1	q_2	q_1	q_2	q_1	q_2
0.00	0.00000007	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
0.50	0.0257576	0.0013785	0.0257576	0.0013785	0.0257564	0.0013764
1.00	1.1057041	0.0053968	0.1057041	0.0053968	0.1057029	0.0053967
1.50	0.2432320	0.0118865	0.2423232	0.0118865	0.2423220	0.0118864
2.00	0.4358009	0.0206887	0.4358009	0.0206887	0.4357999	0.0206886
2.50	0.6837585	0.0316534	0.6837584	0.0316534	0.6837576	0.0316533
3.00	0.9810702	0.0146391	0.9810701	0.0446392	0.9810696	0.0446391
3.50	1.3197869	0.0595124	1.3197868	0.0595124	1.3197867	0.0595123
4.00	1.6891783	0.0761471	1.6891782	0.0761471	1.6891784	0.0761470

Table 2. Error in angles at the joints.

	Error in q_1		Error in q_2	
Time	STWS	RKHeM	STWS	RKHeM
0.0	0.0000E + 00	0.0000E + 00	0.0000	0.0000
0.5	4.6256E − 09	1.1795E − 06	1.7303E − 10	6.6992E − 08
1.0	2.9653E − 09	1.2177E − 06	1.8039E − 10	7.1974E − 08
1.5	6.2724E − 09	1.1616E − 06	3.6119E − 11	7.4573E − 08
2.0	2.3923E − 08	1.0317E − 06	2.4672E − 10	7.6128E − 08
2.5	5.0228E − 08	8.3628E − 07	6.5596E − 10	7.7064E − 08
3.0	8.4720E − 08	5.8317E − 07	1.1803E − 09	7.7576E − 08
3.5	1.2614E − 07	2.8375E − 07	1.8092E − 09	7.7770E − 08
4.0	1.7237E − 07	4.3937E − 08	2.5330E − 09	7.7712E − 08

5. CONCLUSIONS

In this paper, by observing the table of results, it is found that the numerical techniques STWS and RKHeM suit well for determining the discrete solutions of the governing equations of the arm model of the robot of two degree freedom. In particular, this study helps to estimate the variations in the angles at the joints of the robot arm at different fractions of time and enables to identify the robot arm movement.

REFERENCES

1. H. Krishnan and N. Harris Mcclamroch (1994), Tracking in non-linear differential – algebra control systems with applications to constrained robot systems. *Automatica*, 30, 12, pp. 1885–1897.
2. K. Warwick and A. Pugh (1990), *Robot Control – Theory and Applications.* Peter Peregrinus Ltd.
3. H.P. Huang and W.L. Tseng (May 1991), A symptotic observer design for constrained robot systems. *IEE Proceedings Pt – D*, 138, 3, pp. 211–216.
4. D. Paul Dhayabaran (December 2000), A study on second order singular systems using extended RK methods and STWS technique. Ph.D. Thesis. Bharathidasan University, Tamil Nadu, India.
5. D.J. Evans and N. Yaacob (1995), A fourth order Runge-Kutta method based on the Heronian mean formula. *International Journal of Computed Mathematics*, 58, pp. 103–115.

ROTATION OF GRANULAR MATERIAL IN LABORATORY TESTS AND ITS NUMERICAL SIMULATION USING TIJ–COSSERAT CONTINUUM THEORY

K. Sawada, F. Zhang and A. Yashima
Department of Civil Engineering, Gifu University, Gifu 501-1193, Japan.

Abstract It is important to take into consideration the particle rotation when discussing the behaviour of deformation in granular material. A Cosserat continuum theory is suitable for the problems that include rotation of particles in granular materials because the deformation of a ground composed of granular materials is described by both displacements and rotations. In this study, laboratory tests were carried out to investigate the rotation behaviour of granular materials. Then, an elasto-plastic model for sand based on tij-sand model [1] was formulated within Cosserat continuum theory. Furthermore, the model is implemented into a finite element code for the numerical simulation of boundary value problems related to the tests. From a series of laboratory tests and simulations, results are compared and discussed in detail.

Keywords: Cosserat continuum theory, particle rotation, granular materials.

1. INTRODUCTION

In this study, laboratory tests were carried out to observe the deformation of a granular material with particle rotation. Aluminium rods were used to simulate a granular soil mass. From the results of the test, both deformation and rotation of the rods were investigated in details. Furthermore, numerical simulations of the above-mentioned aluminum rod experiments were conducted. In the numerical calculations, an elasto-plastic model for sand based on Cosserat continuum theory was used for finite element analyses. Results of laboratory tests and numerical simulations are compared and discussed in detail.

G. R. Liu et al. (eds.), Computational Methods, 1701–1706.

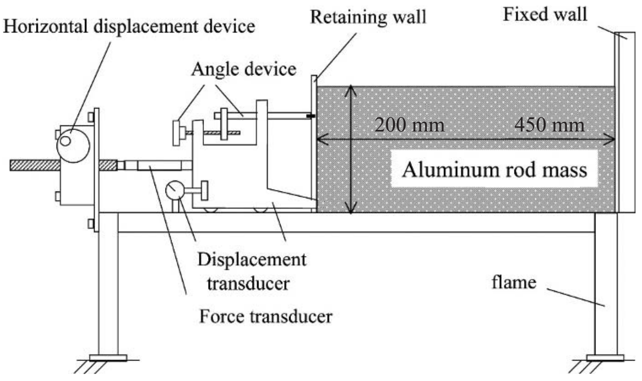


Figure 1. Test apparatus.

2. EXPERIMENTS USING ALUMINIUM ROD MASS

Laboratory tests were carried out to observe the deformation of a granular material with particle rotation. Based on the passive and active earth pressure tests results, it is possible to check the validity of the constitutive model using Cosserat continuum theory, by conducting finite element analyses on the experiments. The aluminium rods used as ground in the experiment were 50.0 mm in length and 1.6 mm or 3.0 mm in diameters. Figure 1 shows the apparatus and its size of passive and active earth pressure tests. The boundary of apparatus and aluminium rods ground are smooth. The assembly of aluminium rods was chosen with a proper weight ratio: 60% of the rods in diameter of 1.6 mm and 40% of the rods in diameter of 3.0 mm, so as to replicate the grain size distribution of Toyoura standard sand. In the tests, given displacements were applied through a retaining wall on the left side of aluminium rods up to a maximum value of 20 mm for active earth pressure tests and 40 mm for passive earth pressure tests. In the tests, both the displacements and the rotations of the aluminum rods were measured. The rotation of the marks are recorded and compared at different times. The selected rods for rotation measurement were chosen where a relatively larger displacement and rotation would take place during loading. Furthermore, a grid was marked on one end of the rods and displacements of the grid can be measured at different times.

The deformations of granular materials are shown from Figures 2 to 7. It is found that a shear band occurred in both tests (see Figures 2 and 5). The angle of shear band is about 35 degree in passive earth pressure condition and about 60 degree in active earth pressure condition. It can be recognized that there is outstanding deformed area for both passive and active earth pressure tests respectively from Figures 3 and 6. In comparison between the passive and active earth pressure condition, the area of deformation under passive earth pressure

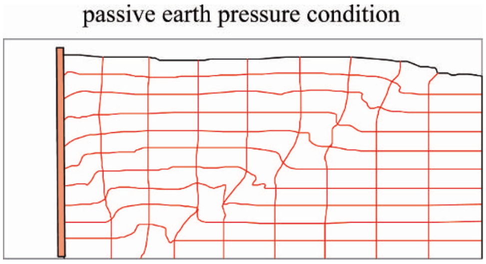


Figure 2. Deformed mesh.

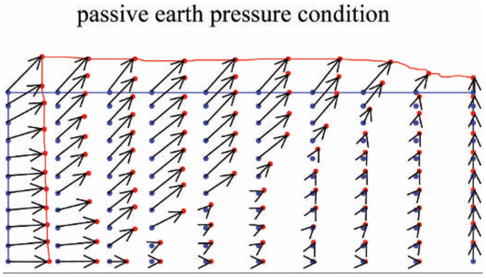


Figure 3. Deformation vector.

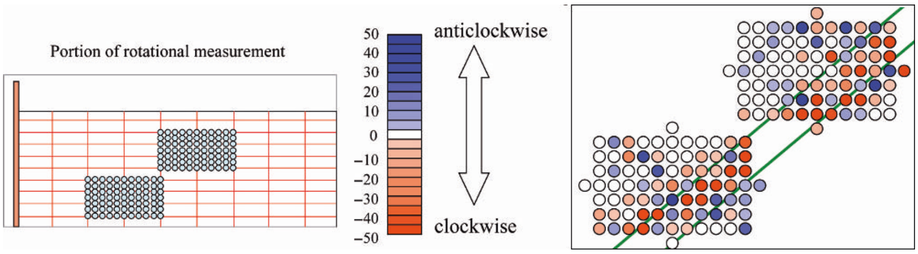


Figure 4. Portion of rotation measurement and the rotation angle in measured area.

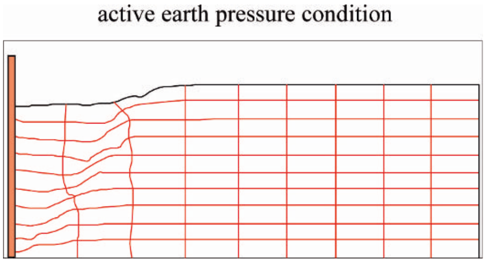


Figure 5. Deformed mesh.

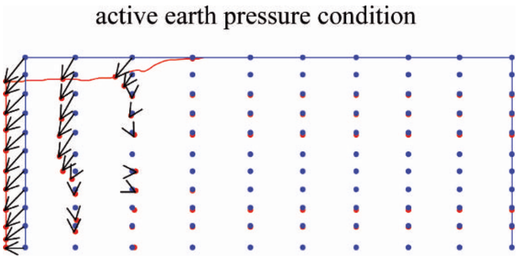


Figure 6. Deformation vector.

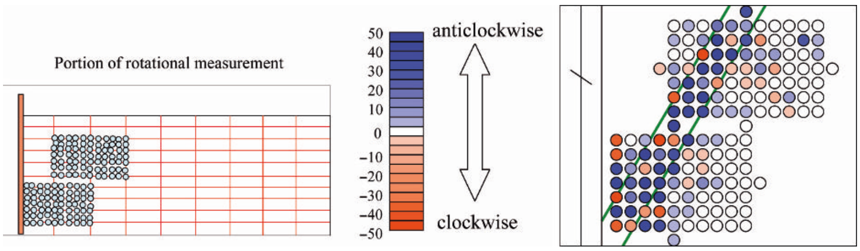


Figure 7. Portion of rotation measurement and the rotation angle in measured area.

condition is larger than active earth pressure condition. From the rotating angle measurement a large mass of clockwise rotation is observed in the passive earth pressure condition (Figure 4). On the other hand, anticlockwise is in large measure for the active earth pressure condition (Figure 7).

3. NUMERICAL ANALYSES

In this section, finite element analyses on the aluminum rod experiment using the Cosserat t_{ij} model is conducted [2]. The boundary conditions are shown in Figure 8. The model ground is consisted with 800 elements that are

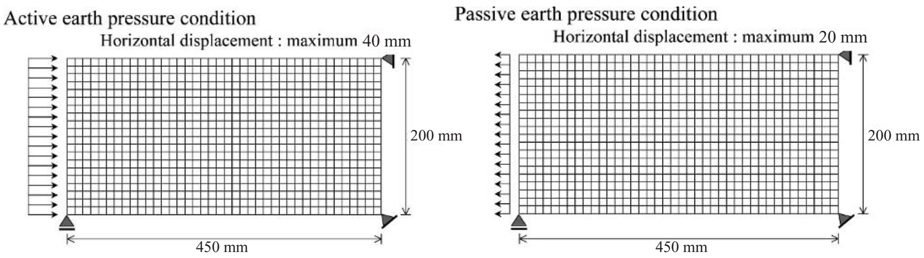
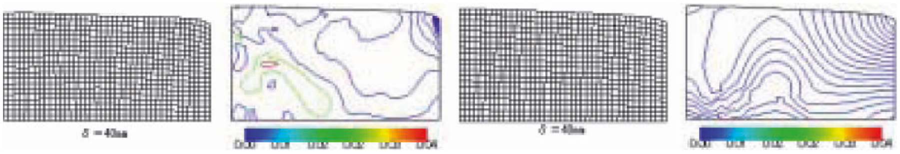


Figure 8. Finite element mesh for both earth pressure conditions.



l_c=0.0 cm Passive earth pressure condition l_c=0.1 cm

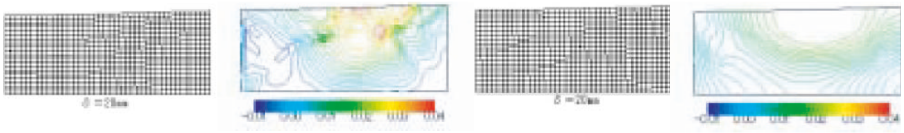
Figure 9. Deformed mesh and distribution of rotational angle (unit = rad).

estimated 20 times 40 elements and the boundary conditions (see Figure 8). For the calculation of Cosserat rotation, the rotation at lower right node of model ground is constraint. In order to evaluate the influence of the characteristic length on the deformation, two different values of Cosserat parameter l_c were used, namely: 0.00 and 0.10 m.

Results when horizontal displacement reached 40 mm under passive earth pressure condition and 20 mm under active conditions are plotted in Figures 9 and 10. No shear band can be observed in both deformed meshes. The influence on the deformation by introducing Cosserat theory is not obvious. The distribution of rotational angle, however, is clearly affected by the value of parameter l_c in both cases. It is known from Figure 9 that the rotational angles of particles are anticlockwise in most part. This tendency is not coincident completely with the experimental results. While in the case of passive earth pressure condition, it is known from Figure 10 that, the rotational angles of particles are clockwise in most part, which is similar to the experimental results. The reason is that, in experiments, rotation angles are measured for the aluminium rods one by one, while in the calculations, the rotation angles are calculated continuously. There is an inherent difference between the way of measurement for rotation angle in numerical calculation and experiments.

4. CONCLUSIONS

In this paper, model tests and numerical analyses were conducted to verify the validity of Cosserat continuum theory. Firstly, laboratory tests were carried out to investigate the behaviours of aluminium-rods mass, especially the



l_c= 0.0 cm Active earth pressure condition l_c= 0.1 cm

Figure 10. Deformed mesh and distribution of rotational angle (unit = rad).

rotation of the rods. Then, an elasto-plastic model is derived based on Cosserat continuum theory with t_{ij} sand model. Using the proposed model, numerical simulations for tests on aluminum-rod mass are conducted. In the numerical simulations, two cases with different Cosserat parameter are used. From both the experimental and analytical tests, rotation of aluminium rods is observed. However, the rotating angle in calculation is smaller than that in the experiments. Also, the rotating direction in calculation is different with experimental results. The calculation results with Cosserat theory cannot describe reasonably the rotation of grains. It still remains a question whether the present calculation scheme can predict the rotation of grains qualitatively. Meanwhile, though the results are not given in detail in the context, further research should be done to take into consideration the strain softening behaviour of sand.

REFERENCES

1. T. Nakai (1989), An isotropic hardening elasto-plastic model for sand considering the stress path dependency in three-dimensional stresses, *Soils and Foundations*, (1), pp. 119–137.
2. K. Sawada, H. Kato, A. Yashima and F. Zhang (2002), Analytical study of grains rotation using tij sand model based on Cosserat continuum theory, *Proceedings of 1st International Workshop on New Frontiers in Computational Geotechnics*, Banff Alberta, Canada, pp. 175–182.

COMPARISON OF CLASSICAL MODELS FOR VISCOELASTICALLY DAMPED SANDWICH BEAMS

H. Hu[1,2], E.M. Daya[1], M. Potier-Ferry[1] and S. Belouettar[2]
[1] *Laboratoire de Physique et Mécanique des Matériaux, UMR CNRS 7554, I.S.G.M.P., Université de Metz, Ile du Saulcy, 57045 Metz Cedex 01, France*
[2] *Laboratoire de Technologies Industrielles, Centre de Recherche Public Henri Tudor, 70, Rue de Luxembourg, L-4002 Esch sur Alzette, Luxembourg*

Abstract The aim of this paper is to analyse classical models for viscoelastically damped sandwich beams. Some are based on the kinematics of Kirchhoff-Love, Mindlin, Reddy or Touratier, the others on the zigzag principle [D.K. Rao, *Journal of Mechanical Engineering Science* 20(5) (1978) 271; E.M. Daya and M. Potier-Ferry, *Revue européenne des éléments finis* 11(1) (2002) 39]. The accuracy of these models is compared in the static field and the dynamic field, for which a finite element based solution is considered as reference. The comparison includes vertical deflection of simply supported sandwich beams under transverse loading, natural frequencies and loss factors of simply supported viscoelastically damped sandwich beams. It is found that the zigzag models are more accurate than the others. The ratio of length to thickness, the ratio of rigidities and the ratio of core and face thickness are chosen as three principal parameters of sandwich beams, and their influences are studied and compared. Finally, the fields of application of these models are defined in terms of these three parameters.

Keywords: comparison, deflection, frequency, loss factor.

1. INTRODUCTION

Vibration control in structures by means of viscoelastic damping material has gained wide acceptance, particularly in the aerospace industry. Significant weight, cost, performance, and reliability payoffs are possible in situations where resonant vibration cannot be avoided. Numerous examples of damping by constrained or unconstrained layers of viscoelastic material have appeared in

G. R. Liu et al. (eds.), Computational Methods, 1707–1713.

the literature. However, the analysis of general structures to predict the damping due to an integral or add-on layered treatment or lumped viscoelastic elements is by no means a routine activity. The key to effective use of the well-developed technology of damping materials will be a parallel improvement in analytical capability. It will be necessary to make reliable and timely predictions of the system damping which can be expected when a given viscoelastic material is incorporated into a given structural configuration. It must also be possible for working engineers to produce these answers in a project environment using widely available analysis tools such as computer codes. It is natural to look to finite element methods for damping predictions of general damped structures, just as they are used for routine calculation of the dynamic properties of undamped systems. Considerable work has already been done in this area, primarily aimed at prediction of damped response of simple structural elements such as sandwich beams, plates, etc. However as with any emerging technology, there are specific questions, both theoretical and practical, which must be investigated and resolved in order to produce a reliable and useful design tool. The damping is introduced by an important transverse shear in the viscoelastic layer. It is due to the difference between in-plane displacements of the elastic layers and also to the low stiffness of the central layer. Numerical simulation of these structures requires, first the use of an adequate model to obtain a reasonable computational cost, second a proper account of the shearing action of the core. Many investigations have been devoted to the static and dynamic analysis of these structures and various approaches and models have been proposed. The purpose of this paper is to compare these proposed models to finite element based solution and to evaluate their efficiency. For these purposes, two problems will be considered: (1) Deformation of simply supported sandwich beams under a constant transverse loading at the centre of beams and (2) free vibration of simply supported viscoelastically damped sandwich beams.

For the first one, Kirchhoff-Love, Reddy [1] and Touratier [2] models are compared with the so-called zigzag models [3, 4], which describe apiecewise continuous displacement field. For the second one, Rao's model is added to the comparison, natural frequencies and loss factors are two elements of comparison.

2. KINEMATICS OF THE MODELS

Consider a three-layer symmetric sandwich beam with a viscoelastic core as Figure 1. Let x be the mid-surface coordinate of the beam and z is the one transverse to the thickness.H_c and E_c are respectively the thickness and Young's modulus of the core. H_f and E_f are the thickness and Young's modulus

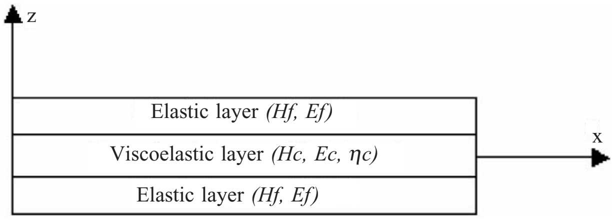


Figure 1. Sandwich beam with viscoelastic middle layer.

of the faces, η_c is the core loss factor. The length and the thickness total are L and H_t.

The following hypotheses, common to many authors, are assumed: (1) All points on a normal to the beam axis have the same transverse displacement $w(x, t)$. (2) The displacement is continuous at the interfaces. (3) All points of the elastic layers on a normal have the same rotations. (4) The core material is homogeneous, isotropic and viscoelastic with a constant loss factor, so the Young's modulus is complex and constant.

The kinematics of non-zigzag models can be written with a general expression:

$$U(x, z, t) = U_0(x, t) - z\frac{\partial w(x, t)}{\partial x} + f(z)\beta(x, t). \tag{1}$$

$U(x, z, t)$ is the longitudinal displacement and $U_0(x, t)$ is the longitudinal displacement of the mid-plane of core. $\beta(x, t)$ is the additional rotation of the normal to the mid-plane, $f(z)$ can be considered as 'shear functions', here expressions considered for $f(z)$ are:

- Model 1: $f(z) = 0$, Kirchhoff-Love kinematical assumptions;
- Model 2: $f(z) = z$, Mindlin kinematical assumptions;
- Model 3: $f(z) = z - 4z^3/(3H_t^2)$, Reddy kinematical assumptions and
- Model 4: $f(z) = H_t \sin(\pi z/H_t)/\pi$, Touratier kinematical assumptions;

For zigzag models, the function of longitudinal displacements are speared by three parts:

$$\begin{cases} U_1(x, z, t) = U_1^0(x,t) - \left(z - \dfrac{H_c + H_f}{2}\right)\dfrac{\partial w(x, t)}{\partial x}, & \dfrac{H_c}{2} < z \leq \dfrac{H_t}{2} \\ U_2(x, z, t) = U_2^0(x,t) - z\dfrac{\partial w(x, t)}{\partial x} + \sum\limits_{i=1}^{n} f_i(z)\beta_i(x, t), & -\dfrac{H_c}{2} \leq z \leq \dfrac{H_c}{2} \\ U_3(x, z, t) = U_3^0(x,t) - \left(z + \dfrac{H_c + H_f}{2}\right)\dfrac{\partial w(x, t)}{\partial x}, & -\dfrac{H_t}{2} \leq z < -\dfrac{H_c}{2} \end{cases} \tag{2}$$

$U_1^0(x, t)$, $U_2^0(x, t)$, $U_3^0(x, t)$ are the axial longitudinal displacements of the mid-plane of the layers. It is clear that Kirchhoff-Love's model is kept for the faces. In the core, three models are considered:

- Model 5: $n = 1$, $f_1(z) = z$, Mindlin's model in the core [3, 4];
- Model 6: $n = 1$, $f_1(z) = z - 4z^3/(3H_c^2)$, Reddy's model in the core;
- Model 7: $n = 2$, $f_1(z) = z$, $f_2(z) = z^3$, Enriched zigzag kinematical assumptions.

3. BASIC EQUATIONS

The sandwich-beam equations can be obtained from Virtual Work Equation:

$$\delta P_{\text{ext}} + \delta P_{\text{int}} = \delta P_{\text{acc}}. \tag{3}$$

As many authors have studied non-zigzag models, here we just derive the basic equations for zigzag models. Considering the continuity of the displacements at the interfaces between the central and the face layers, we can express the face displacement $U_1(x, z, t)$ and $U_3(x, z, t)$ as follows:

$$\begin{aligned} U_1(x,z,t) &= U_2^0(x,t) - z\frac{\partial w(x,t)}{\partial x} + \sum_{j=1}^{n} f_j(z)\beta_j(x,t) \\ U_3(x,z,t) &= U_2^0(x,t) - z\frac{\partial w(x,t)}{\partial x} - \sum_{j=1}^{n} f_j(z)\beta_j(x,t) \end{aligned} \tag{4}$$

With these two expressions, the independent generalized displacements are reduced to $U_2^0(x, z, t)$, $w(x, t)$ and $\beta_i(x, t)$. According to Hooke stress–strain relation and using Equations (3), (6) and (7) are respectively obtained from problems 1 and 2:

$$\int_0^L \left[(M_1 + M_2 + M_3) - \frac{H_c + H_f}{2}(N_1 - N_3)\right] \frac{\partial^2 \delta w}{\partial x^2} dx - \int_0^L F(x)\delta w dx = 0, \tag{5}$$

$$\begin{aligned} &\int_0^L \left[(M_1 + M_2 + M_3) - \frac{H_c + H_f}{2}(N_1 - N_3)\right] \frac{\partial^2 \delta w}{\partial x^2} dx \\ &\quad - \int_0^L (2\rho_f S_f + \rho_c S_c)\frac{\partial^2 w}{\partial t^2}\delta w dx = 0. \end{aligned} \tag{6}$$

Table 1. Material properties and dimensions.

Young's modulus of faces	$E_f = 6.9e10$ Pa
Poisson's ratio	$\upsilon_f = \upsilon_c = 0.3$
Loss factor of core	$\eta_c = 0.3$
Mass density	$\rho_f = 2770$ kg/m^3, $\rho_c = 968$ kg/m^3
Thickness	$H_t = 0.01$ m
Width	$b = 0.1$ m
Example 1	Example 2
$(E_c/E_f = 0.01, H_c/H_f = 10, L/H_t = 50)$	$(E_c/E_f = 3e-4, H_c/H_f = 1, L/H_t = 20)$

For both our two problems, Equation (7) and beam end conditions can be deduced:

$$\int_0^L (N_1 - N_3) \sum_{j=1}^{n} f_j \left(\frac{Hc}{2}\right) \frac{\partial \delta \beta_j}{\partial x} + \int_{S_c} \sigma_{xx} \sum_{j=1}^{n} f_j \frac{\partial \delta \beta_j}{\partial x} dS_c$$

$$+ \int_{S_c} \sigma_{xz} \sum_{j=1}^{n} \frac{\partial f_j}{\partial z} \delta \beta_j dS_c dx = 0. \tag{7}$$

In which M_i is bending moment and N_i is axial force of the ith layer.

4. COMPARISON OF THE RESULTS

By solving these equations, we can compare the results of the two considered problems by two examples. The beam dimensions and material properties are presented in Table 1. A two-dimensional eight nodes quadrilateral element (Q16) is kept for ANSYS 7.1. The results are presented in Tables 2–3 and Figure 2.

Table 2. Deflection comparison of problem 1 ($q = 100$ N/m).

Maximal deflection (m)	Model 1	Model 3	Model 4	Model 5	Model 6	Model 7	ANSYS
Example-1	−1.06E−4	−1.08E−4	−1.09E−4	−1.11E−4	−1.10E−4	−1.11E−4	−1.11E−4
Example-2	−3.01E−6	−3.06E−6	−3.07E−6	−2.05E−5	−1.89E−5	−2.05E−5	−2.06E−5

Table 3. Comparison of problem 2.

Example	Model 1	Model 3	Rao	Model 5	Model 6	Model 7	ANSYS
1 1st frequency	87.5	86.7	85.5	86.1	86.3	86.1	85.9
Loss factor	0	6.5E−3	0.032	0.045	0.040	0.045	
2 1st frequency	627.7	623.2	242.2	242.2	252.8	242.2	241.1
Loss factor	0	5.2E−6	0.44	0.44	0.47	0.44	

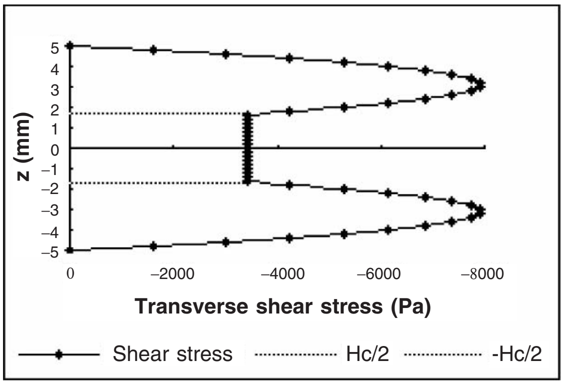


Figure 2. Distribution of shear stress at $x = L/4$ (ANSYS).

5. CONCLUSION

From these tables, we can see that zigzag models are always more accurate than non-zigzag models and the non-zigzag are imprecise when the core is very weak. Models 5 and 7 are very approach with the results of ANSYS. In the case of $E_c/E_f > 0.01$, $L/H_t > 100$ or $|\lg(H_c/H_f)| \geq 3$, all the proposed models give comparatively the same results. Between zigzag models, Rao's model can be used for the modelling of very weak cores. Figure 2 shows that the transverse shear stress is constant in the core. Models 5 and 7 are the more precise models and one can conclude that we do not need to enrich Mindlin's model in the core when we use zigzag principal.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the support of "Ministère de la Culture, de l'Enseignement Supérieur et de la Recherche, Luxembourg" and "Fonds National de la Recherche, Luxembourg".

REFERENCES

1. J.N. Reddy (1984), A simple higher-order theory of laminated composite plate. *Journal of Applied Mechanisms*, 51, pp. 745–752.
2. M. Touratier (1991), An efficient standard plate theory. *International Journal of Engineering Science*, 29, pp. 901–916.
3. D.K. Rao (1978), Frequency and loss factor of sandwich beams under various boundary conditions. *Journal of Mechanical Engineering Science*, 20, 5, pp. 271–282.
4. E.M. Daya and M. Potier-Ferry (2002), A shell element for viscoelastically damped sandwich structures. *Revue européenne des éléments finis*, 11, 1, pp. 39–56.

MODELLING AND ANALYSES OF CRACKS IN FUSELAGE LAP JOINTS WITH A SINGLE-COUNTERSUNK RIVET

G. Shi and G. Li
Structures, Materials and Propulsion Laboratory, Institute for Aerospace Research, National Research Council Canada, 1200 Montreal Road, Building M-14, Ottawa, Ontario, Canada K1A 0R6

Abstract A fracture analysis was carried out to study cracks in fuselage lap joints with a single-countersunk rivet using the finite element and boundary element alternating method. A global–local methodology was developed for crack modelling a complex structure. Stress intensity factor distributions along the crack front at the edge of the rivet hole in the corner of the joint outer sheet were predicted. Parametric studies were performed by varying crack shapes and sizes in the crack analysis. The effects of residual stresses induced by the riveting process on the stress intensity factors were investigated to get a better understanding of the failure mechanism for the lap joints.

Keywords: crack analysis, lap joints with a single-countersunk rivet, FEM–BEM alternating method, global–local approach, residual stress, stress intensity factor.

1. INTRODUCTION

Fatigue crack growth and unstable fracture have the potential to cause catastrophic failures of transport aircraft. Although a great deal of effort has been made to better understand these phenomena, there is still a lack of numerical solutions for predicting the stress intensity factors for cracks in fuselage lap joints with countersunk rivets due to the complex geometry configurations and load conditions. The numerical simulation of crack growth and the calculation of fracture mechanics parameters, such as stress intensity factors, for complex aircraft components subjected to combined external load and residual stresses remain challenging in structural integrity assessment and damage tolerance analysis.

G. R. Liu et al. (eds.), Computational Methods, 1715–1725.

Research is being carried out at the National Research Council Canada on crack growth simulation with an objective to develop computational modelling techniques to simulate the fatigue crack propagation in complex aircraft structures and to estimate the remaining life and residual strength of a component. In the present research work, modelling and analysis were carried out for cracks in fuselage lap joints with a single-countersunk rivet. The stress intensity factors along the crack front of a corner crack in the outer sheet were predicted. For the crack analysis, three models were generated and used: a finite element global model, a local finite element model and a boundary element crack model. A finite element and symmetric Galerkin boundary element alternating method was employed to alternate the finite element and boundary element models to evaluate the stress intensity factors.

2. FEM–BEM ALTERNATING METHOD

In recent decades, various numerical methods, such as finite element methods (FEM), boundary element methods (BEM) and alternating methods, have been developed and employed to obtain stress intensity factor solutions for cracks in three-dimensional structures. The alternating methods, especially the FEM–BEM alternating method, show more advantages in carrying out fracture analysis than other numerical methods, such as a pure FEM or a pure BEM, because they combine advantages from both methods. The finite element method is a robust method for elastic–plastic problems which can easily incorporate various complex boundary conditions. The boundary element method, on the other hand, is more suitable for modelling cracks in infinite or finite bodies. Among the various boundary element methods, the symmetric Galerkin boundary element method provides a more simple, accurate and efficient way of modelling the crack [1]. In recent years, the finite element and symmetric Galerkin boundary element alternating method has been developed [2,3]. In this FEM–BEM alternating method, the solution procedure alternates between the FEM solution for an uncracked structural component of a finite geometry, and the symmetric Galerkin boundary element method solution for cracks in an infinite or finite body.

The basic steps of the FEM–BEM alternating iteration procedure can be described as follows:

1. Using the finite element method, solve the problem in the finite uncracked body with all external boundary conditions. The stress field in the finite body is obtained, which means that the tractions at the location of the crack are determined.

2. Using the symmetric Galerkin boundary element method, solve the problem of a crack in an infinite or finite region. The only loads in this problem are the tractions on the crack surfaces from the FEM solutions.
3. Determine the tractions at the outer boundaries of the FEM model in the finite body from the symmetric Galerkin boundary element method solution.
4. Applying the tractions obtained from the symmetric Galerkin boundary element method in step (3) as the residual forces to the FEM model, re-solve the FEM problem for the finite uncracked body and obtain the tractions at the location of the crack.
5. Repeat steps (2)–(4) until the residual loads are within a preset tolerance.
6. By adding the symmetric Galerkin boundary element method solution to the FEM one, the real solution is obtained.

Once the iteration procedure is completed, the stress and displacement field in the finite body and around the crack can be obtained. Then, the fracture mechanics parameters, such as the stress intensity factors, can be determined.

3. PROBLEM DESCRIPTION

The problem of a joint with a single-countersunk rivet subjected to a uniform tensile loading was considered. The configuration of the joint was consistent with the configuration of the testing specimen used in recent experimental and numerical studies by the authors [4,5]. The joint consisted of two sheets with a thickness of 2.03 mm and one countersunk rivet. The dimensions of the two identical rectangle sheets were 200 × 38.1 × 2.03 mm and the mean rivet hole diameter was 6.51 mm. The rivet shank diameter was 6.35 mm and height was 14.29 mm. Due to symmetry of the joint, only half of the joint was modelled in the numerical simulation, as shown in Figure 1. Symmetric boundary conditions were applied to the joint centre plane along the longitudinal direction and other appropriate boundary conditions were applied. at both ends of the joint. The materials used in the joint were 2024-T3 aluminium alloy bare sheet for the skins and 2117-T4, aluminium alloy for the MS20426AD8-9 rivet. The isotropic hardening behaviour was assumed for both the rivet and, sheet materials at the global FE model analysis stage. Linear elastic behaviour was assumed at the crack analysis stage by neglecting the plasticity effect on the crack characteristics. To simulate cracks, corner cracks with circular and elliptical shapes of different sizes were assumed to be positioned at the countersunk rivet hole in the outer sheet of the joint, as shown in Figure 2. Note that a and c are the characteristic dimensions of the ellipse and t is the

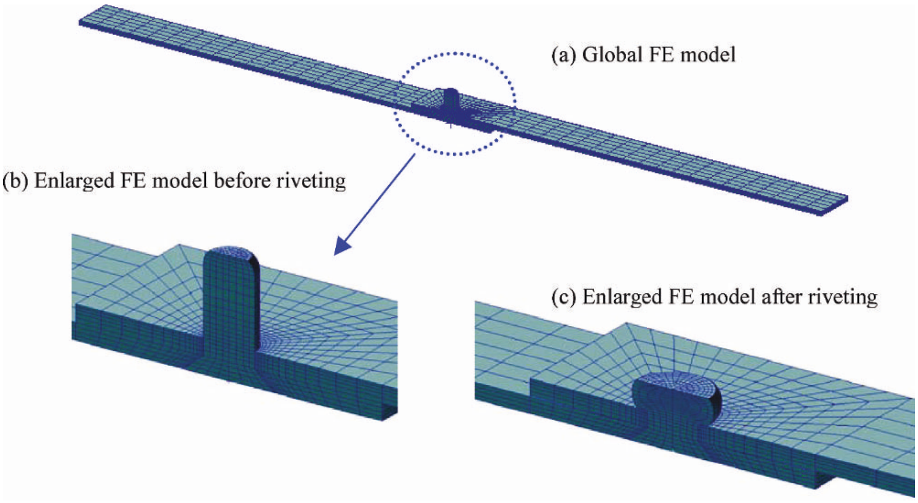


Figure 1. Global FE model.

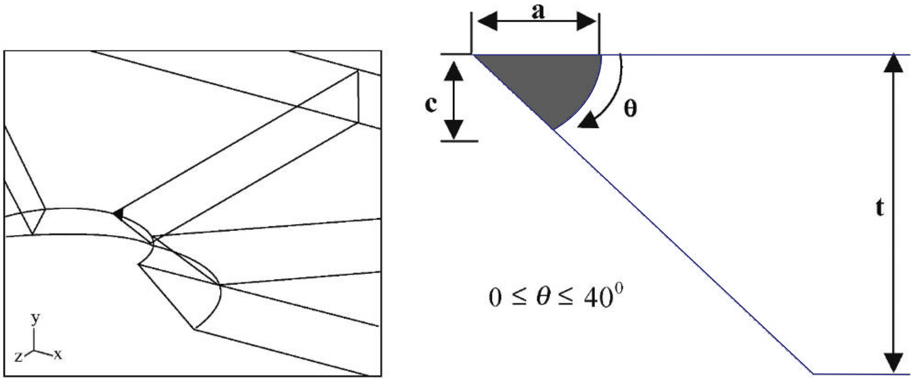


Figure 2. Crack location at the countersunk hole.

thickness of the sheet, θ is a physical angle in this figure from 0° at the faying surface to 40° at the countersunk surface.

4. CRACK MODELLING AND ANALYSIS METHODOLOGY

Fracture analysis for cracks in fuselage lap joints with countersunk rivets have been a challenging issue for many years due to the complex structural configurations, such as the geometry condition (countersunk hole), loading conditions (riveting force and tensile loading) and contact situations between the rivet shank, rivet hole and the faying surface between the two sheets. To carry out a fracture mechanics analysis on such a complex structure, a crack modelling and analysis methodology based on a combination of the FEM–BEM alternating method and global–local approach was developed and implemented to predict the stress intensity factor distributions and the effect that the residual stress induced by the riveting process had on the stress intensity factors.

4.1 Global Finite Element Modelling and Analysis

In a global–local approach the global FE model is usually used to calculate the stress and displacement distributions for the entire model to provide tractions or displacements to a local model as boundary conditions. The finite element model was generated based on the experimental joint dimensions using the commercial FE software package MSC/PATRAN (pre- and post-processor) and MSC/MARC (solver). The global FE model consisted of 5560 8-node reduced integration brick elements associated with 7431 nodes. Three deformable contact bodies, two sheets and one rivet, and two rigid contact bodies, rigid pusher and rigid set, were defined in the model. Four different squeeze forces of 27, 36, 45 and 53 KN were applied to rivet the joints and then the joints were loaded to a remote tensile stress of 98.5 MPa. In the global FE analysis, multiple-load steps were defined. First, a rivet squeeze force was applied to simulate the installation of the rivet; then this squeeze force was released back to zero; after which a remote tensile load was applied to the joint. During the global FE analysis, the residual stresses were calculated during the riveting process simulation and their effects were taken into account in the local stress distributions around the rivet hole during the tensile loading.

4.2 Local Finite Element Model

A local model was used in the global–local approach to calculate the stress and displacement in the finite body without a crack and to alternate with the boundary element solutions for the stress intensity factor evaluation. In the present work, a local model was cut from the global model of the joint in the outer sheet around the rivet hole, as shown in Figure 3. This local FE model consisted of 1568 20-node brick elements associated with 7693 nodes. The traction boundary conditions were transferred from the global model analysis results to the cutting surfaces of the local model.

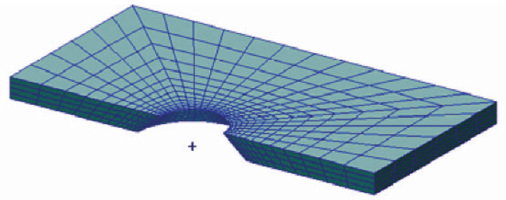


Figure 3. Local FE model (Cutting from the outer sheet).

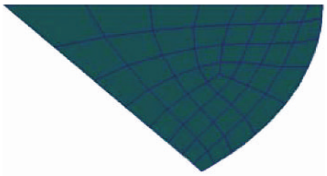


Figure 4. BE crack model (a/c = 1.5, c/t = 0.3).

4.3 BE Crack Model

Boundary element crack models were generated for the fracture mechanics analysis. The crack surface located at the corner of the outer sheet with different shapes and sizes were modelled using quadrilateral 8-node boundary elements and the mid-side nodes were moved to the quarter-side positions for the elements on the crack front, as illustrated in Figure 4.

4.4 Analysis Procedure and Study Cases

Crack analyses were carried out by using the FEM–BEM alternating method based computer code, SAFELIFE, which was developed by Knowledge System Research, LLC. The entire analysis procedure can be summarized as (a) create the global FE models and carry out the stress analysis by applying different squeeze forces during the riveting process simulation and the tensile load during the tension stage; (b) create local FE models and transfer boundary conditions from the global FE model analysis results to the local models; (c) create boundary element crack models with different crack configurations; (d) alternate the local FE model solution and the BE model solution by using the FEM–BEM alternating method until convergence is reached and (e) obtain the stress intensity factors along the crack front. In the present analysis, a total of four global and local models corresponding to four different squeeze forces were generated and analysed, and 12 crack configurations with different crack shapes and sizes were generated.

5. NUMERICAL RESULTS AND DISCUSSIONS

In the fracture analysis, all three fracture modes, K_I, K_{II} and K_{III}, were calculated. However, since it was shown from the analysis that mode I was the dominant failure mode in the lap joint under the current loading condition only mode I results are presented in this paper. In the results, the normalized stress intensity factor is expressed as $F = K_I/K_0$ where $K_0 = \sigma_0(\pi a/Q)^{1/2}$, σ_0 is the remote applied stress, and a is the crack characteristic dimension in the joint transverse direction. Empirical expressions of the shape factor, Q, for an ellipse are given by Raju and Newman [6] as $Q = 1 + 1.464(a/c)^{1.65}$ when a/c $\leq$ 1.0 or $Q = 1 + 1.464(c/a)^{1.65}$ when a/c $>$ 1.0. It should be mentioned that only portions of the numerical results are presented in this paper.

5.1 Effects of Crack Shape and Size on Stress Intensity Factor

For a fixed rivet squeeze force of 36 KN and a circular crack shape (a/c = 1.0), the normalized stress intensity factor distributions along the crack front for cracks of different sizes, c/t = 0.2, 0.3, 0.4 and 0.5, are plotted in Figure 5. This figure shows that the larger the crack size, the smaller the normalized stress intensity factors. In Figure 6, the fracture behaviour for elliptical cracks with a/c = 1.5, 2.0, 3.0, 5.0 and 7.0 show that the normalized stress intensity factors at the faying surface ($\theta = 0°$) are less than those at the countersunk surface ($\theta = 40°$), and the larger crack size corresponds to a smaller normalized stress intensity factor. For another type of elliptical crack with shapes of c/a = 1.5, 2.0 and 3.0 the normalized stress intensity factors at the faying surface are

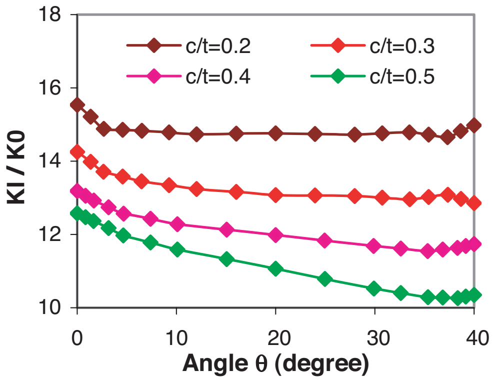


Figure 5. Normalized SIFs for circular cracks (SF = 36 KN).

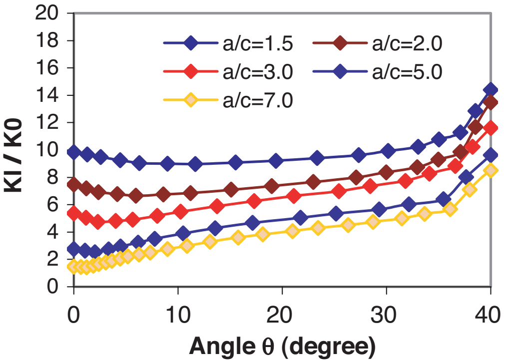


Figure 6. Normalized SIFs for elliptical cracks (SF = 36 KN).

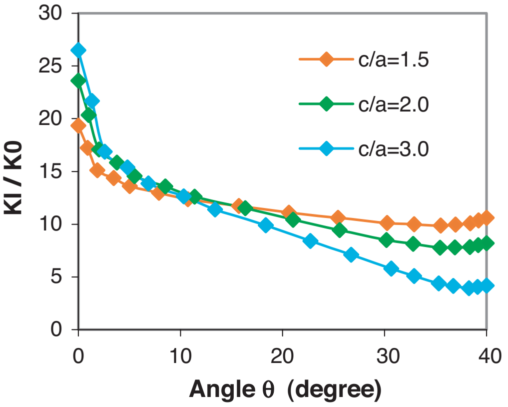


Figure 7. Normalized SIFs for elliptical cracks (SF = 36 KN).

much larger than those on the countersunk surface, and the normalized stress intensity factors near the countersunk surface for c/a = 3.0 are less than c/a = 1.5 and 2.0, as shown in Figure 7.

5.2 Effects of Residual Stress on Stress Intensity Factor

As expected the residual stress induced by the riveting process affected the stress distributions, and thus the crack behaviour in the joints. Figures 8 to 10 show the effects that the residual stress caused by different squeeze forces had on the stress intensity factor solutions in the single riveted lap joint. In Figure 8, the normalized stress intensity factor distributions for the circular crack (a/c = 1.0) along the crack front varying with different squeeze forces (SF = 27, 36,

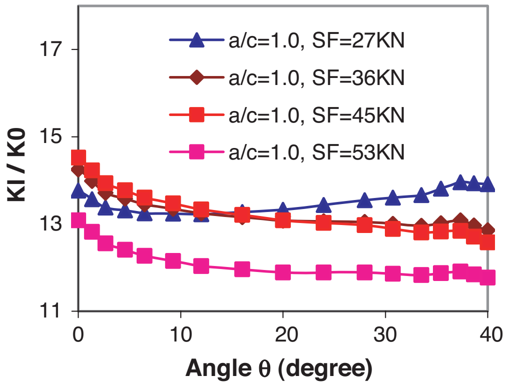


Figure 8. Normalized SIFs for circular cracks (c/t = 0.3).

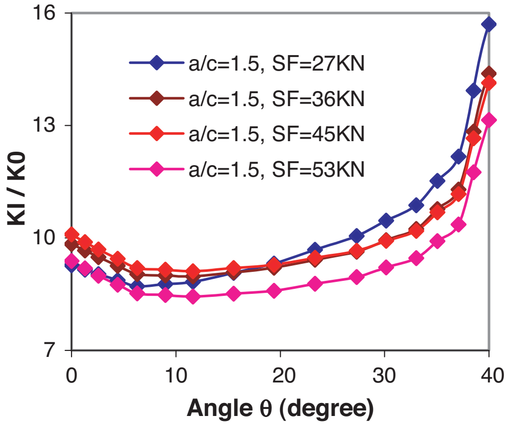


Figure 9. Normalized SIFs for elliptical cracks (c/t = 0.3).

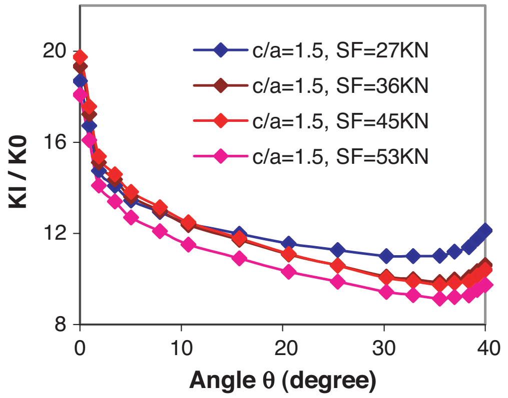


Figure 10. Normalized SIFs for elliptical cracks (c/t = 0.3).

45 and 53 KN) are plotted. It can be seen from this figure that the rivet squeeze force had a significant effect on the stress intensity factor distribution. As the crack front approached the countersunk surface, the normalized stress intensity factor became smaller when the rivet squeeze force was larger. This is because the residual stress reduced the local stress at the edge of the countersunk hole ($\theta = 40°$), which in turn reduced the stress intensity factors. The normalized stress intensity factor distributions under different squeeze forces SF $= 27$, 36, 45 and 53 KN along the crack front are plotted in Figure 9 and 10 for the elliptical cracks with $a/c = 1.5$ and $c/a = 1.5$, respectively. These two figures demonstrate the same trend that the higher rivet squeeze force results in a lower stress intensity factor near the countersunk surface, which will eventually increase the fatigue life of the riveted lap joint.

6. CONCLUSIONS

A fracture analysis was carried out to study cracks in fuselage lap joints with a single-countersunk rivet using the finite element and boundary element alternating method and the developed global–local approach. The stress intensity factor distributions for cracks at the corner of the outer sheet were predicted. The results show that the stress intensity factor distributions significantly depend on the crack shapes and sizes. The residual stress caused by the rivet squeeze force in the riveting process had considerable effect on the stress intensity factors near the countersunk surface. The larger the squeeze force, the smaller the stress intensity factor, which would result in a longer joint fatigue life.

ACKNOWLEDGEMENTS

This work has been carried out under IAR Programme 303 Aerospace Structures, Project 46_QJ0_13, Computational Structures Technology – Crack Growth Simulation. The financial assistance received from DND/DRDC is gratefully acknowledged.

REFERENCES

1. M., Bonnet, G. Maier and C. Polizzoyyo (1998), Review articles – symmetric Galerkin boundary element methods. *Applied Mechanics Reviews*, 51 (11), pp. 669–704.
2. G. P. Nikishokov, J. H. Park and S. N. Atluri (2001), SGBEM–FEM alternating method for analyzing 3D non-planar cracks and their growth in structural components, *Computational Modeling in Engineering and Science*, 2 (3), pp. 401–422.

3. Z. D. Han and S. N. Atluri (2002), SGBEM (for cracked local subdomain) – FEM (for uncracked global structure) alternating method for analyzing 3D surface cracks and their fatigue-growth. *Computer Modeling in Engineering and Sciences*, 3 (6), pp. 717–730.
4. G. Li. and G. Shi. (2004), Investigation of residual stress/strain in fuselage lap joints with a single countersunk rivet: I. Experimental testing, National Research Council Canada, Technical Report, *LTR-SMPL-2004-0130*.
5. G. Li. and G. Shi (2004), Investigation of residual stress/strain in fuselage lap joints with single countersunk rivet: II. Finite element modeling. National Research Council Canada, Technical Report, *LTR-SMPL-2004-0131*.
6. I.S. Raju and J.C. Newman Jr. (1979), Stress-intensity factors for a wide range of semi-elliptical surface cracks in finite-thickness plates, *Engineering Fracture Mechanics*, 11, pp. 817–829.

THREE-DIMENSIONAL SOLUTION OF A DEEP BEAM USING AN EFFICIENT FINITE-DIFFERENCE SCHEME

M. Zubaer Hossain, S. Reaz Ahmed and M. Wahhaj Uddin
Department of Mechanical Engineering, Bangladesh University of Engineering and Technology, Dhaka 1000, Bangladesh

Abstract The application of a new numerical method of solution is described for 3-D analysis of an elastic deep beam. More specifically, an efficient finite-difference scheme has been developed based on a potential function formulation to solve the 3-D deep beam. In the present approach, a new scheme of reduction of unknowns is used to formulate the 3-D beam problem in terms of a single potential function, defined in terms of the three displacement components. Compared to the conventional computational approaches, the present method provides numerical solution of higher accuracy with reduced computational effort. The suitability and reliability of the method has been verified through the comparison of results with those obtained by the usual method of solution.

Keywords: finite-difference method (FDM), stress analysis, 3-D beam, potential function.

1. INTRODUCTION

Analysis of stresses in a material body is basically a 3-D problem. However, still now, in most of the cases, the 3-D problems are approximated to either 1-D or 2-D ones, mainly due to the lack of useful method of solution by which the 3-D problems can be formulated efficiently. Elastic problems of solid mechanics are usually formulated either in terms of deformation or stress parameters. The stress formulation is not suitable for solving mixed-boundary-value problems. The displacement formulation, on the other hand, involves finding three displacement functions simultaneously from three second-order partial differential equations. In fact, 3-D stress analysis of structural problems is mainly handled by FEM. Although the adaptation of FEM relieved us

G. R. Liu et al. (eds.), Computational Methods, 1727–1731.

from the major inability of solving 3-D problems as well as managing the odd boundary shapes, we are constantly aware of the lack of sophistication and doubtful quality of FEM solutions, especially, for the surface stresses [1]. This is mainly because of the manifold increase of computational work and a lot of loss in sophistication in satisfying the boundary conditions, especially in the region where the boundary conditions change from one type to the other. On the other hand, the superiority of FDM in predicting the state of stresses along the boundary has been verified in our previous researches [2–4]and also by Dow et al. [5].

The present paper describes the application of a new method of solution [6] for the analysis of a 3-D both ends fixed deep beam, where the potential function formulation has been used in conjunction with FDM. The present method reduces the problem to finding a single potential function, instead of three functions at the nodal points, and hence, a tremendous saving in computational work is achieved, which, in turn, increases the quality of solution significantly.

2. POTENTIAL FUNCTION FORMULATION OF THE ELASTIC PROBLEM

Considering the equilibrium of a cubic element with its sides parallel to the axes x, y and z, under the action of the continuous functions $\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{xy}, \sigma_{yz}$ and σ_{xz}, the differential equations of equilibrium in terms of the displacement components, u_x, u_y and u_z, are obtained as follows [7]:

$$G'\frac{\partial^2 u_x}{\partial x^2} + G\left[\frac{\partial^2 u_x}{\partial y^2} + \frac{\partial^2 u_x}{\partial z^2}\right] + G''\left[\frac{\partial^2 u_y}{\partial x \partial y} + \frac{\partial^2 u_z}{\partial x \partial z}\right] = 0 \tag{1a}$$

$$G'\frac{\partial^2 u_y}{\partial y^2} + G\left[\frac{\partial^2 u_y}{\partial x^2} + \frac{\partial^2 u_y}{\partial z^2}\right] + G''\left[\frac{\partial^2 u_x}{\partial x \partial y} + \frac{\partial^2 u_z}{\partial y \partial z}\right] = 0 \tag{1b}$$

$$G'\frac{\partial^2 u_z}{\partial z^2} + G\left[\frac{\partial^2 u_z}{\partial y^2} + \frac{\partial^2 u_z}{\partial x^2}\right] + G''\left[\frac{\partial^2 u_x}{\partial z \partial x} + \frac{\partial^2 u_y}{\partial y \partial z}\right] = 0 \tag{1c}$$

Where

$$G'' = \frac{E}{2(1+\mu)(1-2\mu)},\ G' = \frac{E(1-\mu)}{(1+\mu)(1-2\mu)} \text{ and } G = \frac{E}{2(1+\mu)}.$$

In the present approach, the problem is reduced to the determination of a single variable instead of evaluating three functions u_x, u_y, u_z, simultaneously, from the equilibrium Equations (1). A potential function of space variables ψ(x,y,z) is thus defined in terms of the displacement components as [6]

$$u_x = \frac{\partial^2 \psi}{\partial x \partial y} \tag{2a}$$

$$u_y = -\left[2(1-\mu)\frac{\partial^2 \psi}{\partial x^2} + (1-2\mu)\frac{\partial^2 \psi}{\partial y^2} + 2(1-\mu)\frac{\partial^2 \psi}{\partial z^2}\right] \tag{2b}$$

$$u_z = \frac{\partial^2 \psi}{\partial y \partial z} \tag{2c}$$

When the displacement components of Equations (1) are replaced by ψ(x,y,z), Equations (1a) and (1c) are found to be satisfied automatically and the only condition (Equation 1b) that ψ has to satisfy becomes

$$\frac{\partial^4 \psi}{\partial x^4} + \frac{\partial^4 \psi}{\partial y^4} + \frac{\partial^4 \psi}{\partial z^4} + 2\frac{\partial^4 \psi}{\partial x^2 \partial y^2} + 2\frac{\partial^4 \psi}{\partial y^2 \partial z^2} + 2\frac{\partial^4 \psi}{\partial z^2 \partial x^2} = 0 \tag{3}$$

Therefore, a single potential function ψ(x,y,z) has to be evaluated from the governing differential equation of equilibrium (3), satisfying the associated boundary conditions. The displacement boundary conditions are given by Equation (2). Now the six stress components can readily be expressed explicitly in terms of the function ψ(x,y,z), when the displacement components of the stress–displacement relations are replaced by Equation (2).

3. METHOD OF SOLUTION

FDM is used here to transfer the governing differential equation and also the differential equations associated with the boundary conditions into their corresponding algebraic forms. Considering the rectangular shape of the boundary and also the nature of the differential equations involved, a uniform rectangular mesh-network is used all over the region concerned. The discrete values of ψ(x,y,z), at the mesh points are obtained from a system of linear algebraic equations resulting from the discretization of the governing equation and the associated boundary conditions. Finally, all the parameters of interest in the solution of the beam problem can readily be obtained as soon as ψ is known at the mesh points. It is noted that, in the present FDM solution, the number of equations required to be solved is simply one-third of the usual computational approaches.

4. NUMERICAL RESULTS OF THE DEEP BEAM

Figure 1 illustrates the geometry and loading of the both ends fixed deep beam (L/D = 2 and L/B = 2). For obtaining the numerical results, the beam is

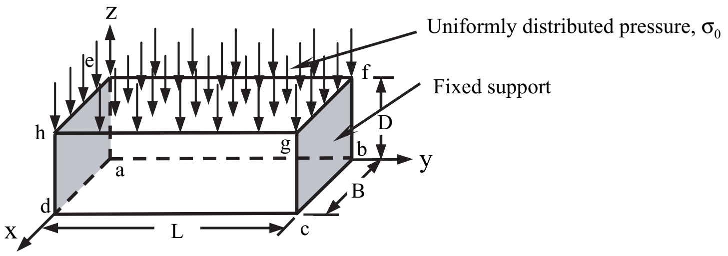


Figure 1. Loading and geometry of the 3-D both ends fixed beam.

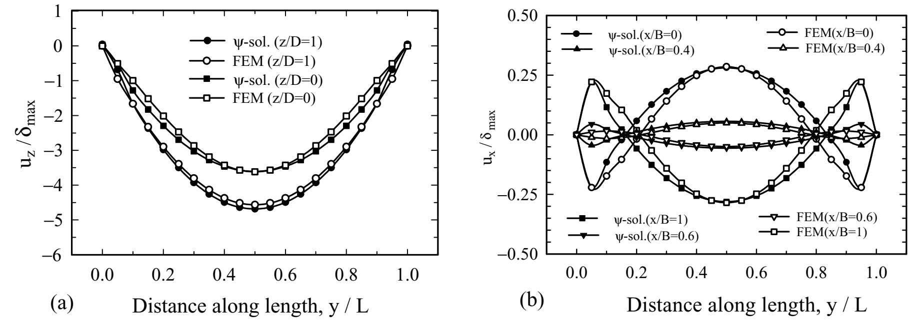


Figure 2. (a) Solutions of u_z at different sections of the vertical plane x/B = 1. (b) Solutions of u_z at different sections of the vertical plane z/D = 0.

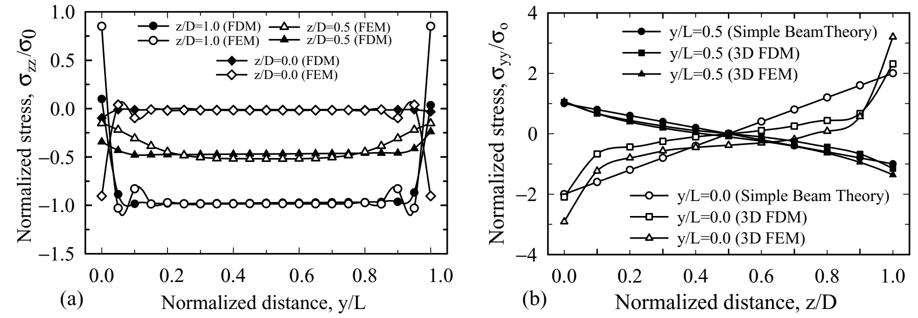


Figure 3. (a) Solutions of σ_{zz} at different sections of the vertical plane x/B = 1. (b) Solutions of σ_{yy} at different sections of the vertical plane x/B = 1.

assumed to be made of ordinary steel ($\mu = 0.3$, E = 209 GPa). The FDM solutions are presented here along with the corresponding FEM solutions obtained by ANSYS code. Figure 2 presents the solutions of u_z and u_x at different sections of two different planes of the beam, where the displacements are normalized w.r.t. the maximum theoretical deflection of the beam, δ_{max}. FDM solutions are found to be in very good agreement with FEM solutions. Figure 3 presents the solutions of two stress components, σ_{zz} and σ_{yy}, at different sections of two different planes of the beam. Two solutions are found to be very close to each other except those around the region of transition. The present FDM solutions conform to all the requirement of the physical characteristics of the beam appropriately, thereby establishing the superiority over the usual method of solution.

5. CONCLUSION

The present Finite-Difference method of solution is verified to be capable of producing accurate and reliable solutions at any critical section of a 3-D deep beam. The drastic reduction in the number of algebraic variables in the present approach improves the accuracy as well as computational efficiency significantly.

REFERENCES

1. T.H. Richards and M.J. Daniels (1987), Enhancing finite element surface stress predictions: a semi-analytic technique for axisymmetric solids. *Journal of Strain Analysis*, 22, 2, pp. 75–86.
2. S.R. Ahmed, A.B.M. Idris and M.W. Uddin (1996), Numerical solution of both ends fixed deep beams. *Computers and Structures,* 61, 1, pp. 21–29.
3. S.R. Ahmed, M.R. Khan, K.M.S. Islam and M.W. Uddin (1998), Investigation of stresses at the fixed end of deep cantilever beams. *Computers and Structures,* 69, pp. 329–338.
4. M.A.S. Akanda, S.R. Ahmed and M.W. Uddin (2001), Stress analysis of gear teeth using displacement potential function and finite-differences. *International Journal for Numerical Methods in Engineering*, 53, 7, pp. 1629–1649.
5. J.O. Dow, M.S. Jones and S.A. Harwood (1990), A new approach to boundary modeling for finite difference applications in solid mechanics. *International Journal for Numerical Methods in Engineering*, 30, pp. 99–113.
6. M.Z. Hossain (2004), A new approach to numerical solution of three-dimensional mixed-boundary-value elastic problems. M.Sc. Thesis, Bangladesh University of Engineering and Technology (BUET).
7. S.P. Timoshenko and J.N. Goodier (1979), *Theory of Elasticity*, 3rd ed. McGraw-Hill Book Company, New York.

COMPUTATIONAL METHOD OF SEA LOADS ON FLOATING STRUCTURES

Yanying Wang
School of Naval Architecture, Dalian University of Technology, 2 Ling-gong Road, Dalian 116024, China

Abstract A computer packet, which is composed of three-dimensional source–sink distribution approach with boundary element numerical method, is applied to determine the interaction of environmental factories and floating structure in the partly non-linear category and in the frequency domain. Some numerical techniques, such as the form of Green function, the treatment of free surface boundary conditions and the calculation of kernel function of Green function, are discussed and all of conclusion remarks are useful for the engineering practice in this chapter.

Keywords: source–sink distribution approach, boundary element method, Green function, environmental loads, waves.

1. INTRODUCTION

The environmental loads with their encountered probabilities and the structural response distribution in the frequency domain induced by wind, waves and current, which are necessary to predict the operative performance and the structural reliability, are calculated by using CFD numerical approach in this chapter. Based on the environmental observational data the design parameters for wind, waves and current can be determined from a computational model of the long-term probability distribution analysis. A computer packet, which is composed of three-dimensional source–sink distribution approach with boundary element numerical method, is applied to determine the interaction of environmental factors and floating structure in the partly non-linear category and in the frequency domain. Thus, the response functions of motion and the long-term probabilities and the frequency characteristics of encountered loads for floating structure can

G. R. Liu et al. (eds.), Computational Methods, 1733–1741.

be calculated. A computational example for a built FPSO is introduced and the results include the prediction of environmental factories in operative sea area, and of motion response functions and dynamic loads with different loading states under stated environmental conditions.

2. MATHEMATICAL MODELS

A series of mathematical models is necessary to predict environmental data in the operative sea area, motion induced by environmental factors for structures in the frequency domain, hydrodynamic pressure and composition of forces as well as their probability distribution on structures and so on.

Based on the stylebook of the long-term distribution of the statistical characteristics for the short-term observation samples of waves in the operative sea area, the Weibull probability function is used to determine the extreme waves encountered by structures with specifically return period [1, 2], so called the design wave:

$$H_D = (H_C - H_0)[17.27 + \ln T_C - \ln T_Z]^{1/\xi} + H_0 \tag{1}$$

in which H_0, H_C, ξ are the parameters of the Weibull probability density function, T_Z is the main cross-up wave period, in s; T_C is the return period for design wave, in year.

The flow field around floating structures is assumed to be the potential flow with incompressible and irrotational fluid, finite water depth (d), rigid bottom parallel to still water surface, progressive wave on the free surface and rigid surface for a floating hull. Therefore, the velocity potential φ pertaining to the fluid region should satisfy the Laplace equation. Based on the linear assumption the total velocity potential can be expressed by sum of the potentials of individual motion for six-freedom degree, i.e.

$$\varphi = (\varphi_0 + \varphi_7 + i\omega\eta_j\varphi_j)e^{i\omega t} \tag{2}$$

The linear velocity potential for incident waves can be expressed as $\Phi_I = Re\left(\varphi_0 e^{i\omega t}\right)$ and the steady velocity potential φ_0 can be written as

$$\varphi_0 = i\frac{gH}{2\omega}\frac{ch[k_0(z+d)]}{ch(k_0 d)}e^{-ik_0(x\cos\beta + y\sin\beta)} \tag{3}$$

where β is angle between progressive direction of wave and x-axis, k_0 is wave number satisfying the dispersion relation of $k_0 th(k_0 d) = \omega^2/g$. The steady motion velocity potentials with unit velocity $\varphi_j (j = 1, 2, \ldots, 6)$ and the diffraction potential φ_7 should satisfy the following boundary conditions:

$$\begin{array}{ll}
[L] & \nabla^2\varphi_j(x,y,z) = 0 \textit{ in flow domain} \\
[F] & (\partial\varphi_j/\partial z) - \nu\varphi_j = 0 \text{ at free surface} \\
[S] & \begin{array}{l} \partial\varphi_j/\partial n = n_j \quad j = 1,2,\ldots,6 \\ \partial\varphi_j/\partial n = -(\partial\varphi_0/\partial n) \quad j = 7 \end{array} \textit{ on body surface} \\
[B] & (\partial\varphi/\partial z)_{z=-d,or-\infty} = 0 \textit{ on sea bad} \\
[R] & \lim\limits_{\rho\to\infty}\sqrt{\rho}[(\partial\varphi/\partial\rho) + ik_0\varphi] = 0 \textit{ at infinite far field}
\end{array} \tag{4}$$

where $\nu = \omega^2/g$, and $n_1 = n_x, n_2 = n_y, n_3 = n_z$, $n_4 = zn_y - yn_z$, $n_5 = xn_z - zn_x$, $n_6 = yn_x - xn_y$, where n_x, n_y, n_z are the direction cosines of the normal to the surface of body at the point (x, y, z).

In this case Haskind sources may be employed, which are distributed on wetted body surface and satisfy free surface and bottom boundary conditions. The velocity potential φ_j can be written as that

$$\varphi_j = \iint_{S_0} \sigma_j(\xi,\eta,\zeta)\cdot G(x,y,z;\xi,\eta,\zeta)dS \tag{5}$$

in which S_0 is average wetted surface of floating body; (ξ, η, ζ) repressed source point, q, on body surface; (x, y, z) repressed field point, p, in the flow field. $\sigma_j(q)$ is Haskind source strength on point q; $G(p, q)$ is Green function which should satisfy the following determination conditions in Equation (4) except body surface condition and considering wave-making effect of body moving speed on free surface condition:

$$\begin{array}{ll}
[L] & \nabla^2 G(p,q) = \delta(p-q) \textit{ in flow domain} \\
[F] & -\nu gG + 2iU\sqrt{\nu g}(\partial G/\partial x) + U^2(\partial^2 G/\partial x^2) + g(\partial G/\partial z) \\
 & = 0 \textit{ at free surface} \\
[B] & (\partial G/\partial z)_{z=-d,or-\infty} = 0 \textit{ on sea bed} \\
[R] & \lim\limits_{\rho\to\infty}\sqrt{\rho}[(\partial G/\partial\rho) + ik_0 G] = 0 \textit{ at infinite far field}
\end{array} \tag{6}$$

where U is forward speed for body in x-direction. Thus, the velocity potential φ_j can be expressed by the distribution of Green function on body surface and its distributing strength σ_j can be determined by the boundary condition of body surface. The form of determination solution for Green function can be also assumed into the following form for finite water depth:

$$G(p,q) = \frac{1}{r_1(p,q)} - \frac{1}{r_1'(p,q)} + \frac{1}{r_2(p,q)} - \frac{1}{r_2'(p,q)} + G^*(p,q) \tag{7}$$

where r and r' can be written as $r_1, r_1' = \sqrt{(x-\xi)^2 + (y-\eta)^2 + (z\mp\zeta)^2}$ and $r_2, r_2' = \sqrt{(x-\xi)^2 + (y-\eta)^2 + (z+2d\pm\zeta)^2}$. Then the Green function in frequency domain for navigating structures can be solved [4] as

$$
\begin{aligned}
G^* = {} & \frac{1}{2\pi}\int_0^\infty \int_{-\pi}^{\pi} \frac{2e^{-2md} sh(m\zeta) ch[m(z+d)]}{ch(md)} e^{im[(x-\xi)\cos\theta+(y-\eta)\sin\theta]} d\theta dm \\
& - \frac{1}{2\pi}\int_0^\infty \int_{-\pi}^{\pi} \frac{2mge^{-md} ch[m(\zeta+d)] ch[m(z+d)]}{ch(md)\left[\left(\sqrt{\nu g}+mU\cos\theta\right)^2 - mgth(md)\right]} \\
& \times [1+th(md)] e^{im[(x-\xi)\cos\theta+(y-\eta)\sin\theta]} d\theta dm
\end{aligned} \tag{8}
$$

By using simple mathematic treatment Equation (8) can be degenerated into the cases for finite water depth and non-navigational body; for infinite water depth and navigational body; as well as for infinite water depth and non-navigational body.

The solution of Equation (5) may be carried out by using an approach similar to Hess–Smith method [3]: p panels (to be triangle or quadrangle) are distributing on body wetted surface S_0, taking a point at each panel (x_i, y_i, z_i) with $(i = 1, 2, \dots, p)$, so called control point. The source strength σ and the direction cosines of the normal to the surface of body (taken as directed outward from the body surface) $\vec{n}_i$ at a panel (x_i, y_i, z_i) are assumed to be constants respectively. Thus,

$$
\sum_{k=1}^{p} \sigma_{jk} \vec{n}_i \iint_k \nabla G(x_i, y_i, z_i; \xi, \eta, \zeta) ds = \begin{cases} n_{ji} & \begin{matrix} j = 1, 2, \dots, 6 \\ i = 1, 2, \dots, p \end{matrix} \\ -\partial\varphi_0(x_i, y_i, z_i)/\partial n & j = 7 \end{cases} \tag{9}
$$

it can be written into the simplification form:

$$
\tilde{P}\tilde{\sigma} = \tilde{Q} \tag{10}
$$

in which the coefficient matrix $\tilde{P}$, called induced matrix, is the square matrix with p orders; $\tilde{Q}$ is the right matrix, and $\tilde{\sigma}$ is the determined matrix. In which $j = 1, 2, \dots, 6$ are to be motion modes and $i = 1, 2, \dots, p$ are to be number of panels in n_{ji} and σ_{ji} respectively. Now the source strength can be solved and both motions and loads of floating body induced by waves can be determined. Then the Bernoulli equation can be used to calculate the hydrodynamic pressure distribution on body surface and the motion equations can be applied to predict the response function for body induced by waves in six-freedom degree.

3. NUMERICAL APPROACHES

It is found that dreadful oscillation and increasing amplitude appear at $\theta = \pm\pi/2$ and $z \to 0$ for the integrand of Green function and integral accuracy is limited by integral step size. An effective technique is that the upper limit of integration, $\pi/2$ is mapped into infinite by using the variable substitution, such as $t = scs\theta$ or $t = tg\theta$ [4]. In order to make identification the following function is selected to be example:

$$I = \int_{\pi/4}^{\pi/2} k_i e^{k_i \omega} (1 + 4\tau \cos\theta)^{-1/2} d\theta \tag{11}$$

in which $i = 1, 2$; $k_1, k_2 = (1/2\cos^2\theta)(1 + 2\tau\cos\theta \pm \sqrt{1 + 4\tau\cos\theta})$; and $\tau = U\omega_e/g$, $\omega = Z + i(x\cos\theta + y\sin\theta)$, $Z < 0$. Then

$$\begin{aligned} I &= \int_{\sqrt{2}}^{\infty} dt \frac{1}{\sqrt{1 + \frac{4\tau}{t}}} \cdot \frac{-1}{t^2\sqrt{1 - \frac{1}{t^2}}} \cdot \frac{t^2}{2}\left(1 + \frac{2\tau}{t} + \sqrt{1 + \frac{4\tau}{t}}\right) e^{\frac{t^2}{2}\left(1+\frac{2\tau}{t}+\sqrt{1+\frac{4\tau}{t}}\right)\omega} \\ &= \int_{\sqrt{2}}^{\infty} dt \frac{-1}{\sqrt{1 + \frac{4\tau}{t}}\sqrt{1 - \frac{1}{t^2}}} \cdot \left(1 + \frac{2\tau}{t} + \sqrt{1 + \frac{4\tau}{t}}\right) e^{\frac{t^2}{2}\left(1+\frac{2\tau}{t}+\sqrt{1+\frac{4\tau}{t}}\right)\omega} \\ &\leq \int_{\sqrt{2}}^{\infty} dt|\cdot| \leq \int_{\sqrt{2}}^{\infty} dt \frac{1}{\sqrt{1 - \frac{1}{t^2}}} \cdot \left(1 + \frac{2\tau}{t} + \sqrt{1 + \frac{4\tau}{t}}\right) e^{\frac{t^2}{2}\left(1+\frac{2\tau}{t}+\sqrt{1+\frac{4\tau}{t}}\right)Z} \end{aligned} \tag{12}$$

$$\begin{aligned} &\lim_{t\to\infty}\left[\frac{1}{\sqrt{1 - \frac{1}{t^2}}} \cdot \left(1 + \frac{2\tau}{t} + \sqrt{1 + \frac{4\tau}{t}}\right) e^{\frac{t^2}{2}\left(1+\frac{2\tau}{t}+\sqrt{1+\frac{4\tau}{t}}\right)Z}\right] \cdot t^2 \\ &= \lim_{t\to\infty} t^2 e^{t^2 z} \overset{\because z<0}{=} 0 \end{aligned} \tag{13}$$

Now taking $i = 1$, $t = 1/\cos\theta$, then "I" can be transformed into Equation (12) in which $|\cdot|$ is the absolute value of integrant for the above formula. From the last formula it can be obtained as Equation (13). Based on Cauchy criterion for infinite improper integral it is well known that the integral $\int_a^\infty f(x)dx$ is convergence absolutely when $\lim\limits_{x\to\infty} x^P|f(x)| = K$, with $0 < K < +\infty$ and $P > 1$. Then

$$\begin{aligned} E(M) = &\int_M^{\infty} dt \frac{1}{\sqrt{1 + \frac{4\tau}{t}}} \cdot \frac{-1}{t^2\sqrt{1 - \frac{1}{t^2}}} \\ &\times \frac{t^2}{2}\left(1 + \frac{2\tau}{t} + \sqrt{1 + \frac{4\tau}{t}}\right) e^{\frac{t^2}{2}\left(1+\frac{2\tau}{t}+\sqrt{1+\frac{4\tau}{t}}\right)\omega} \leq \int_M^{\infty} dt|\cdot| \\ &\leq \int_M^{\infty} dt \frac{1}{2\sqrt{1 + \frac{4\tau}{M}}\sqrt{1 - \frac{1}{M^2}}} \cdot \left(1 + \frac{2\tau}{M} + \sqrt{1 + \frac{4\tau}{M}}\right) e^{\frac{t^2}{2}\left(1+\frac{2\tau}{t}+\sqrt{1+\frac{4\tau}{t}}\right)Z} \\ &\leq A\int_M^{\infty} \frac{1}{2} e^{\frac{t^2}{2}Z} dt \leq A\int_M^{\infty} \frac{1}{2} e^{\frac{Mt}{2}Z} dt = \frac{A}{Mz} e^{\frac{Mt}{2}z}\bigg|_M^{\infty} = -\frac{A}{Mz} e^{\frac{1}{2}M^2 z} \end{aligned} \tag{14}$$

The truncation error of "I" is important part for total error of Green function and the accuracy of Green function can be measured by the analysis of the truncation error of "I"

$$A = [1 + 2\tau/M + (1 + 4\tau/M)^{1/2}][(1 + 4\tau/M)(1 - M^{-2})]^{-1/2} \qquad (15)$$

Let assuming point M which is far away from the origin, then the expectation of truncation error of "I" at M, $E(M)$ can be written as Equation (14). In this equation, A is a finite value as expressing from Equation (15). It can be seen that A will trend to 1 when M trends infinite. It can be found that the expectation of error at M is decreasing speedy when the distance from origin to M is increasing and is also depended on the distance of Z, i.e. the sum of vertical coordinates of source and field points. When both source and field points are close to the free surface the expectation of error at E is decreasing slowly.

4. COMPUTATIONAL EXAMPLE

To be a computational example a brief result for the FPSO is provided in this section. The principal dimensions of hull and the environmental parameters in the operative sea state are listed in Tables 1 and 2 respectively. The response functions of motion induced by the combination of wind, wave and current and the wave moment distributed in frequency domain with different wave directions and loading states, and the wave moment both for significant and maximum values with the exceed probability are calculated in this computation.

Table 1. FPSO vessel principal particulars.

Parameters	Unit	Dimension
Length over all	m	285.00
Depth	m	26.60
Breadth	m	58.00
Draught with full loads	m	16.6
Displacement at full load draught	ton	25,3440

Table 2. Environment parameters.

Item	Unit	Dimension
Wind velocity	m/s	28.0
Wave height	m	10.2
Wave period	s	13.8
Water depth	m	119.5
Current velocity	m/s	0.9

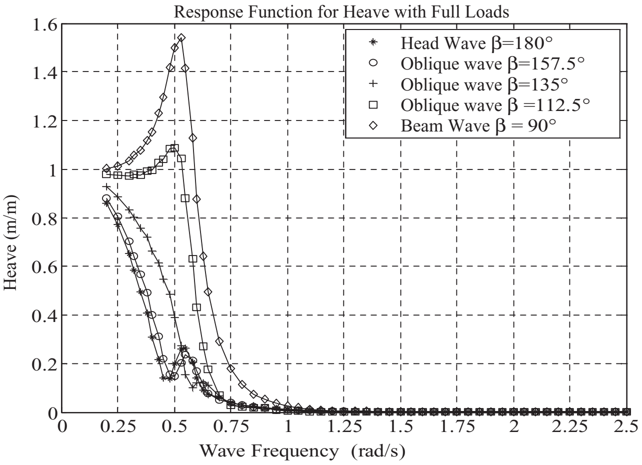


Figure 1. Response function for heave with full loads.

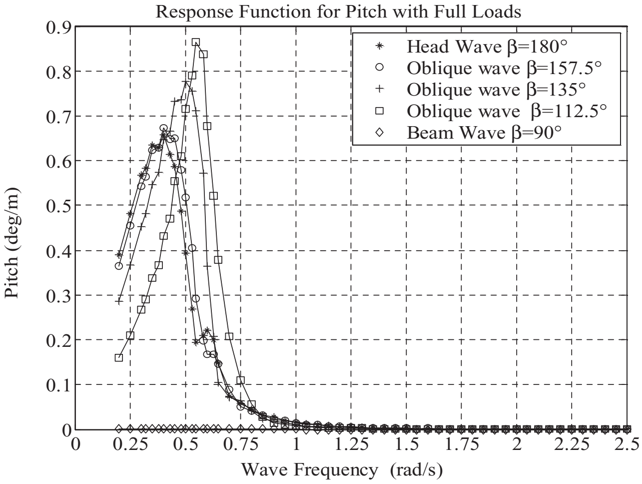


Figure 2. Response function for pitch with full loads.

Some of computational results are listed in the above figures in which the response functions for heave and pitch motion with full loading and different direction of incident waves are shown in Figures 1 and 2; the distribution of hydrodynamic pressure on the hull surface with full loading and different direction of incident waves are given in Figures 3 and 4; the sagging moment distribution in the frequency domain and the wave moment values both for significant and for maximum value with various exceed probability are given in Figures 5 and 6 respectively.

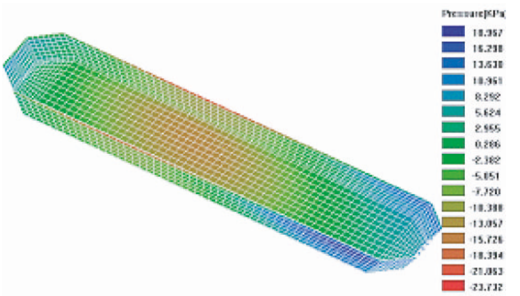


Figure 3. Hydrodynamic pressure on the surface with full loads and oblique wave.

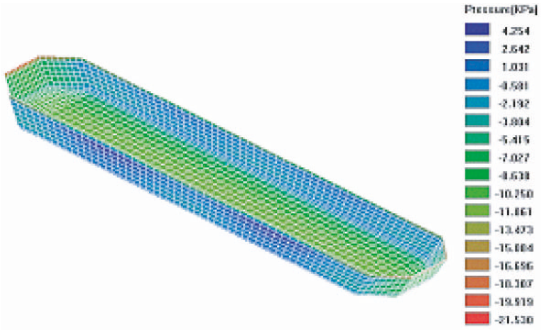


Figure 4. Hydrodynamic pressure on the surface with full loads and beam wave.

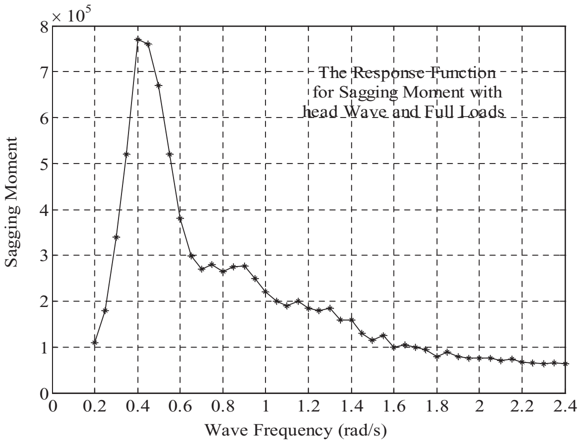


Figure 5. The sagging moment distributed in frequency domain.

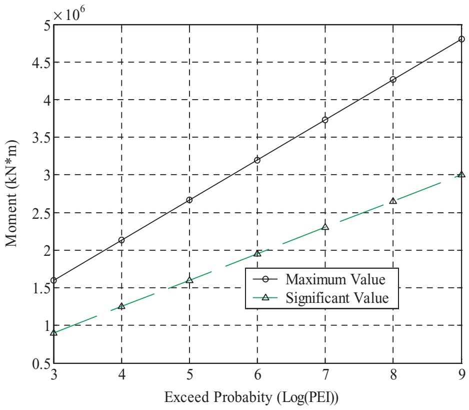


Figure 6. The wave moment values with the exceed probability.

5. CONCLUSION REMARKS

The solution of Laplace boundary problem can be used to analyse the sea loads and motion encountered by floating structures under sea environmental factors. The source–sink distribution approach, the basic form of Green function, and the treatment of the kernel function for the integral Green function in this computation are efficient and the programmed computer packet is useful to engineering practice. This method has been applied to the analysis of mooring system and to be preprocessor for structure analysis by use of FEM computer system.

The project of the calculation of sea loads and motion for floating structures is continuing now and the emphases of investigation are put in the effect of nonlinear factories and the analysis in the time-domain.

REFERENCES

1. Det Norske Veritas (DNV) (September 1987), Strength analysis of main structures of column stabilized units (semi-submersible platforms). Classification Notes, Note No. 31.4.
2. Y.Y. Wang (1988), Investigation of design wave parameters for Chinese coastal areas. *China Ocean Engineering* 2, 4 pp. 71–78.
3. J.L. Hess and A.M.O. Smith (1964), Calculation of non-lifting potential flow about arbitrary three-dimensional bodies. *Journal of Ship Research*, 8, 2, pp. 22–44.
4. Y.Y. Wang, K. Qian and R.C. Zhu (2000), Principle solutions of Green function for radiation problem of floating structures in waves. *Journal of Dalian University of Technology*, 40, 3, pp. 338–340.

NEW FORMULAS FOR DESIGN OF SOCKETS USED IN CABLE STRUCTURES

Y.F. Luo and D.Y. He
College of Civil Engineering, Tongji University, 1239 Siping Road, Shanghai 200092, China

Abstract The empirical formulas are commonly employed for design of the sockets used in cable structures. By using these formulas, accurate stress distribution in the socket may not be obtained. In response to the demand of the FAST project and the socket-manufacturer, new theoretical design formulas are given in this paper and further revised on basis of the comparison of numerical results by theoretical formulas, FEM and empirical formulas. It is shown that the new formulas are more accurate than existing empirical formulas and more reliable when used in engineering design.

Keywords: socket, design formula, conical thick shell, stress.

1. INTRODUCTION

Anchor sockets are the most important joint elements in tension and cable structures. The design formulas of the sockets are still empirical now. None of accurate theoretical design formulas appears recently. The increasing application and rapid development of the structure engineering need more accurate and practical design formulas for socket.

In response to the demand of the FAST project and the manufacturer of the anchor systems, new theoretical design formulas for sockets are given in the paper. Through the comparison of the numerical results by new theoretical formulas, FEM and empirical formulas, the theoretical formulas are revised for practical use. The numerical results show that the revised formulas are more accurate and can be used in engineering design.

G. R. Liu et al. (eds.), Computational Methods, 1743–1747.

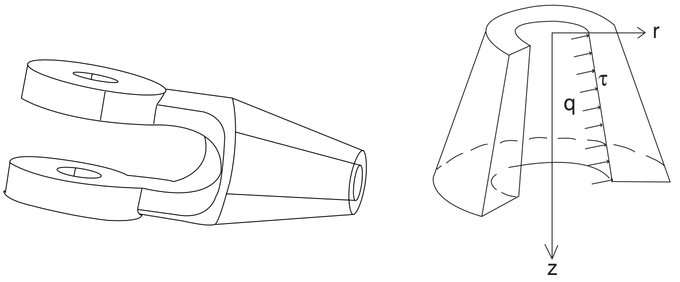


Figure 1. The anchor socket and mathematical model.

2. THE THEORETICAL FORMULAS OF THE CONICAL THICK SHELLS

After simplification of a socket (Figure 1), the new theoretical formulas of stress distribution in the socket can be derived using 3-dimensional elastic theory. Introducing Love's displacement function, $\zeta(r, z)$, the stresses in the socket are given in following (Ref. [1]).

$$
\begin{aligned}
\sigma_z &= \frac{Q}{\pi(r_2^2 - r_1^2)} + \frac{q_r r_1^2 r_2^2 A_z}{(r_2^2 - r_1^2)A} \\
\sigma_r &= \frac{q_r r_1^2 A_r}{r^2 A} \\
\sigma_\theta &= -\frac{q_r r_1^2 A_\theta}{r^2 A} \\
\tau_{rz} &= \frac{q_r z A_\tau}{r A} \qquad (1)
\end{aligned}
$$

where, r_1 and r_2 are the inner and out boundaries of the socket; z is the height. The details of parameter Q, q_r, A, A_r, A_z, A_θ and A_τ can be found in Ref [1].

It is shown in Equation (1) that the stresses in the socket are not only the function of the radius r, but also the height z.

3. THE NUMERICAL RESULTS OF FEM

3.1 Parameters for analysis and simplification of the model

The engineering model provided by manufacturer for numerical analysis is I109 (Ref. [2]). The socket is made of cast steel. Its yielding stress is $f_y = 310$ MPa. The maximum strength is $f_b = 570$ MPa. The elasticity model is $E = 2.06 * 10^5$ MPa and Poison's ratio is $\mu = 0.3$. The friction factors are taken as $tg\phi = 0.2$, 0.3 and 0.4 respectively. The theoretical model is shown in Figure 1. The boundary in two ears is hinged, that is restrained in z direction, free in r and θ. Two load cases, 100% and 167% of the tension capacity of the corresponding cable, are applied on the socket.

3.2 Numerical results of FEM

Taking $tg\phi = 0.2$, the socket is in elastic when the load reaches 100%. The plastic strain appears in the small end of the socket when the load reaches 120%. The plastic strain appears in all inner surface of the socket when the load reaches 167%. Taking $tg\phi = 0.3$, the socket is in elastic when the load reaches 100% and 120%. The plastic strain appears in the small end of the socket when the load reaches 167%. Taking $tg\ \phi = 0.4$, the socket is in elastic when the load reaches 100%, 120% and 167%.

4. THE RESULTS OF EMPIRICAL FORMULAS AND COMPARISON

The empirical formulas collected and the corresponding numerical results are listed in Table 1. The results of the derived theoretical formulas and FEM

Table 1. Stresses of the socket I109 and comparison (100% of tension capacity).

References	Formulas	Stress	Friction factor tgϕ						Error to FEM (%)
			0	0.1	0.2	0.3	0.4	0.5	
Nears	$P_r = T/2\pi tg(\phi + \beta)$	σ_θ	451	216	141	104	82	68	25~33
Pujiang	$\sigma_r = \dfrac{P\cos\phi}{A_e\sin(\theta+\phi)} = \dfrac{p_z\cos\phi}{\sin(\theta+\phi)}$	σ_r	324	157	103	77	62	51	0~25
	$\sigma_\theta = \sigma_c \dfrac{D_s^2 + D_i^2}{D_s^2 - D_i^2}$	σ_θ	595	288	190	142	113	94	
Wei	$H = p/(tg\alpha + \mu_2)\cos\alpha$	σ_θ	2432	1177	833	597	495	411	300~400
Chen	$p = \dfrac{Y}{B}\left((1+B)\left(\dfrac{r}{R_1}\right)^{2B} - 1\right)$	σ_r		116	104	86	64	57	15~40
Thick	$\sigma_r = \dfrac{a^2 p}{b^2 - a^2}\left(1 - \dfrac{b^2}{r^2}\right)$	σ_r	323	156	103	77	61	51	0~25
cylinder	$\sigma_\theta = \dfrac{a^2 p}{b^2 - a^2}\left(1 + \dfrac{b^2}{r^2}\right)$	σ_θ	593	287	189	141	113	94	
New	$z = h/2$	σ_r	323	156	103	77	61	51	0~25
Formulas		σ_θ	653	316	208	155	124	103	
	$z = h/2$	σ_r			82	61	50		
FEM		σ_θ			209	140	110		

are also included in the table for comparison. It can be concluded from the results listed in the table:

The stress given by Nears [3] is the average circumferential stress σ_θ of the socket inner surface. The stresses given by Pujiang [2] are the stresses σ_θ, σ_r of inner surface in half height of the socket. The stress given by Chen [4] is also the stress σ_r of inner surface in half height of the socket. All these stresses are the stresses in particular points of the socket. They cannot describe the exact stress distribution in the socket.

The formulas obtained by circular thick shell theory [5] are only the function of the radius r, but not concern with the shell height z.

The derived formulas are the function of conical radius r and height z. They give accurate stress distribution of the socket. The stress distribution is better close to FEM analysis.

5. THE REVISED THEORETICAL FORMULAS

Compared the numerical results of the theoretical formulas and of FEM, it can be found that both results are better close to each other, but there are some errors in the small end of the socket. After comparison, new-revised theoretical formulas are given:

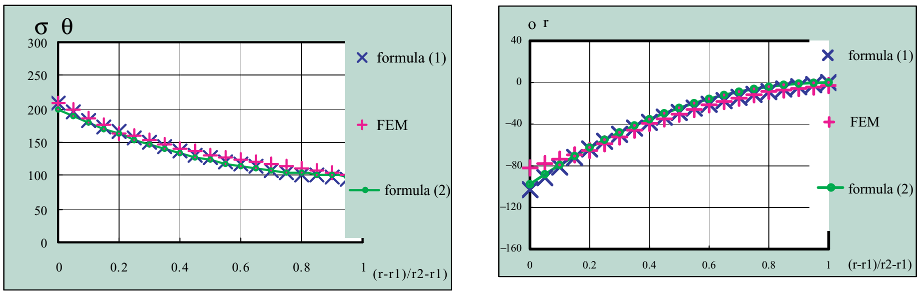
$$\begin{aligned} \sigma_r &= \omega q_r \left(\frac{r_b - r}{r_b - r_a}\right)^2 \\ \sigma_\theta &= \eta q_r \left(1 + \left(\frac{r_b - r}{r_b - r_a}\right)^2\right) \\ \omega &= 0.95, \eta = 0.4\left(1 - \frac{z}{h}\right)^2 + 0.86 \end{aligned} \tag{2}$$

where r_a, r_b are the inner and out surface radii in half socket; q_r is the radial component of the load.

The comparison of the numerical results between the new formulas and FEM are shown in Figure 2. It is shown that two curves go closer to each other.

6. CONCLUSIONS

The empirical formulas cannot give the exact stress of the conical anchor socket. The stresses calculated by circular thick shell theory are only the stresses concerned with the radius, but not with the height. Then σ_θ is obvious changed with the socket height z. The numerical results show that the increase of the



σ_θ (z=h/2) σ_r

Figure 2. The stresses σ_θ and σ_r of I109 socket ($tg\phi = 0.2$, 100% of tension capacity).

friction reduces the stresses of the socket. This is exactly shown in new theoretical formulas.

The new-revised theoretical formulas can exactly describe the stress distribution in the socket. Their results are better close to the results of FEM. The revised formulas can be used for engineering design.

REFERENCES

1. D.Y. He (2004), Analysis and design method of the anchor socket used in high strength cable joints. *Dissertation for Master degree*, Tongji University, Shanghai, China, p. 3.
2. *Hot Casting Anchor Using in Cable Structures* (2001), Shanghai Pujiang Cable Company, Q/IMAA03.
3. J. Nears (1993), Trans. by L.S. Yao, *Cable Bridge*, Publication of People's Jiaotong, p. 3.
4. S.C. Chen and Q.R. Ye (1988), *Plastic Shaping Principal of Metals*, Promotion University of Mechanical Engineers.
5. B.Y. Xu and X.S. Liu (1995), *Applied Elastic Mechanics*, Publication of Tsinghua University.

A NOVEL SUBCYCLING ALGORITHM FOR COMPUTER SIMULATION OF CRASHWORTHINESS

H. Gao, G.Y. Li, X. Han and Z.H. Zhong
Key Laboratory of Advanced Technology for Vehicle Design & Manufactory of M.O.E,
Hunan University, Changsha, China 410082

Abstract It can decrease significantly the computation time in transient structural analysis to use subcycling algorithms that permit multiple time steps in an explicit integration. In this paper, several subcycling algorithms are tested and it is shown that constant velocity subcycling algorithm possesses the best stability properties. A new subcycling algorithm that introduces damping to velocities is proposed. The numerical examples show that the proposed algorithm is both accurate and computationally efficient.

Keywords: finite element, subcycle, multi time step.

1. INTRODUCTION

The maximum time step is limited by the maximum natural frequency of finite element in transient structural analysis. Some elements may require a much smaller time step than the rest and the global time step of a structure model decreases significantly, if only one time step is adopted. So the computation for solving the problem increases dramatically. It is effective to use subcycling algorithms that permit multiple time steps in explicit integration. The subcycling algorithm was proposed by Belytschko et al. [1]. In this method, according to time step, the structure model is arranged into several regions and each region uses a different stabile time step. Within a major time step, in order to calculate internal forces at adjacent nodes that use a smaller time step, displacements or velocities at a nodal interface between different time steps must be obtained. Some studies by Neal and Belytschko [2] and Belytschko and Lu [3] assume a constant velocity over the major time step or a constant

G. R. Liu et al. (eds.), Computational Methods, 1749–1753.

acceleration. While the former has received success in non-linear analysis, accuracy problems have been reported. Although few algorithms assuming of constant acceleration within a major cycle has been proven to be stable, the work by Daniel [4] and Smolinski et al. [5] have shown that they can receive better accuracy than that of constant velocity. Both of algorithms are tested, so as to investigate the possibility using for crashworthiness analysis. The examples show that the assumption of constant acceleration bring unstable, while constant velocity algorithms arouse errors for the simulation of a car collision. In this paper, a new constant velocity subcycling algorithm is proposed by introducing damping to velocities and its accuracy is investigated.

2. SUBCYCLING ALGORITHMS

In an explicit subcycling algorithm, the elements are arranged into element groups. Each element group is integrated with a different time step. As mentioned by Smolinski et al. [6], each group is subject to the following restrictions:

a) The largest step must be an integer multiple of all time steps.
b) If any node is shared by elements in two different integration groups, the time steps in these groups must be integer multiples of each other.

In an explicit finite element analysis, a velocity strain formulation is often used for all element calculations. In constant velocity subcycling algorithms, if an element is connected to a node with a larger time step, it uses the same nodal velocity for all intermediate time steps within a major time step, while in constant acceleration subcycling algorithms, it uses the same nodal acceleration.

The constant velocity subcycling algorithms can be described as followed. Considering a major cycle consists of $n$subcycles, the start of the current major cycle is labeled as state 0, the end of the first subcycle as state 1 and the end of the major cycle as state n, respectively. The middle of subcycle i is labeled as $i + 1/2$ and so on.

a) Common update of the large time step region (region L) and the small time step region (region S) for state 0:

$$\begin{cases} v_S^{1/2} = v_S^{-1/2} + \Delta t_S a_S^0 \\ v_L^{1/2} = v_L^{-1/2} + \Delta t_L a_L^0 \\ x_S^1 = x_S^0 + \Delta t_S v_S^{1/2} \\ x_L^1 = x_L^0 + \Delta t_S v_L^{1/2} \end{cases} \tag{1}$$

b) Subcycle update i, internal forces are calculated only in region S.

$$\begin{cases} v_S^{i+1/2} = v_S^{i+1/2} + \Delta t_S a_S^i \\ v_L^{i+1/2} = v_L^{1/2} \\ x_S^{i+1} = x_S^i + \Delta t_S v_S^{i+1/2} \\ x_L^{i+1} = x_L^i + \Delta t_S v_L^{1/2} \end{cases} \tag{2}$$

Note that region S and L use different time step Δt_S, $\Delta t_L$and the velocities in region L $v_L^{i+1/2}$is held constant in (2).

The constant acceleration subcycling algorithm holds the mid-major cycle acceleration constant. Equations (1) and (2) can be modified to a constant acceleration subcycling algorithm, by replacing Δt_L by Δt_S in (1) and $v_L^{i+1/2} = v_L^{1/2}$ by $v_L^{i+1/2} = v_L^{i-1/2} + \Delta t_S a_L^0$ in (2). A subcycling algorithm that averages acceleration is proposed by Daniel [4]. In this method, the acceleration to be held constant in region L is altered from that computed at the start of the major cycle a_L^0, to a_L^j for major cycle j in Equation (2).

$$(a_C^j + a_C^{j-1})/2.0 = a_L^0 \tag{3}$$

and with the initial $a_C^{j-1} = a_L^0$.

2.1 Testing of subcycling algorithms

Subcycling algorithms except the constant velocity one prove to be unstable in many complicated cases. Can the constant velocity subcycling algorithm provide good accuracy in a car collision analysis? It is shown by Figure 1 that the calculation results from the constant velocity subcycling algorithm are different from the reference one by using single one time step.

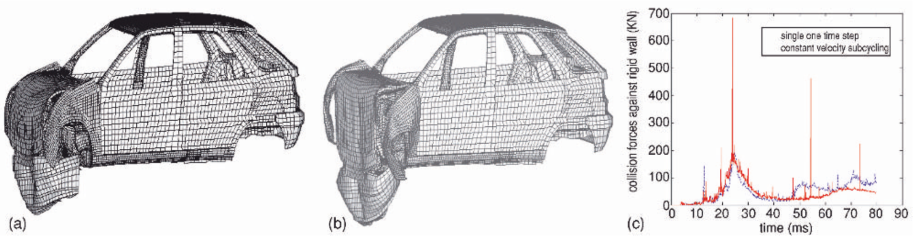


Figure 1. (a) Result from one time step. (b) Result from constant velocity subcycling algorithm. (c) Comparison of collision forces against rigid wall.

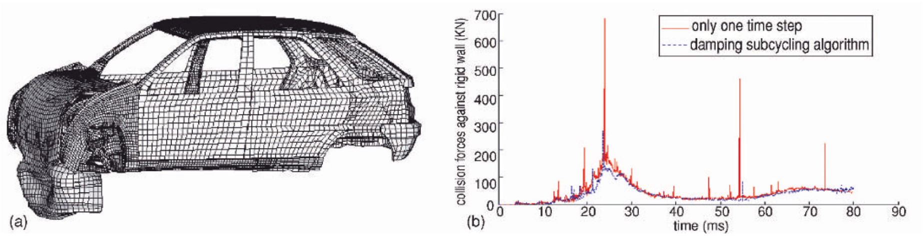


Figure 2. (a) Result from damping method. (b) Comparison of collision forces against rigid wall.

2.2 A subcycling algorithm with damping

To decrease errors by using the constant velocity subcycling algorithm in a car collision analysis, damping is added. This can be done by replacing $v_L^{i+1/2}$ in (2) by

$$v_L^{i+1/2} = v_L^{i-1/2} + \alpha \Delta t_S a_L^0 \tag{4}$$

where α is the damping factor, restricts to 0.01–0.015.

As shown as Figure 2, the new method is proven to give good accuracy for the computer simulation of crashworthiness. The time for solving the problem cost by using dump subcycling algorithm is 10 h 22 min, comparing to 26 h and 21 min by using the only one time step.

3. CONCLUSION

Several subcycling algorithms have been tested and it is shown that the constant velocity algorithm exist better accuracy than others, but brings errors to complicated system analysis. A new subcycling algorithm that introduces dumping is proposed and found to possess good accuracy while sharing high efficiency.

ACKNOWLEDGEMENTS

This work is supported by the National 973 Program of China under the grant number 2004CB719402 and the National Science Foundation of China under the grant number 10372031.

REFERENCES

1. T. Belytschko, H.J. Yen and R. Mullen (1979), Mixed methods in time integration. *Computer Methods in Applied Mechanics and Engineering*, 17/18, pp. 259–275.
2. M.O. Neal and T. Belytschko (1989), Explicit-explicit subcycling with non-integer time step ratios for structural dynamic systems. *Computers and Structures*, 31, pp. 871–880.
3. T. Belytschko and Y.Y. Lu (1992), An explicit multi-time step integration for parabolic and hyperbolic systems. In: *New Methods in Transient Analysis*, PVP-Vol. 246/AMD-Vol, 143; ASME, New York, pp. 25–39.
4. W.J.T. Daniel (1997), Analysis and implementation of a new constant acceleration subcycling algorithm. *International Journal for Numerical Methods in Engineering*, 40, pp. 2841–2855.
5. P. Smolinski, S. Sleith and T. Belytschko (1996), Stability of an explicit multi-time step integration algorithm for linear structural dynamics equations, *Computational Mechanics*, 18, pp. 236–244.
6. P. Smolinski, T. Belytschko and M. Neal (1988), Multi-time-step integration using nodal partitioning. *International Journal for Numerical Methods in Engineering*, 26, pp. 349–359.

NUMERICAL ALGORITHM FOR DETERMINING HOPF BIFURCATION POINT OF NONLINEAR SYSTEM

Dexin Li[1], Yanjun Lu[1] and Lijun Jiang[2]
[1]*School of machinery and precision instrument engineering, Xi'an University of Technology, Xi'an, 710048, P. R. China*
[2]*Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613*

Abstract This chapter presents a new iterative method based on the existence conditions of Hopf bifurcation. By means of proposed method, the problem on how to determine Hopf bifurcation point in an autonomous nonlinear system with bifurcation parameters is converted to how to solve a set of nonlinear algebraic equations. This method can determine both Hopf bifurcation point and the pure imaginary eigen-value pair of Jacobian matrix of a system at the same time. Thus, it can avoid some weakness in some existing methods, such as repeatedly solving the eigen-value whenever the chosen parameters are changed.

Keywords: Hopf bifurcation, nonlinear dynamical system, iterative method.

1. INTRODUCTION

Hopf bifurcation method has been used for dynamic system as a common tool to solve the periodic orbit problem. All of the eigen-values of the first power system should be known before resolving Hopf bifurcation problem of a dynamic system by analytic method. Hopf bifurcation theorem can be directly applied to the low-dimension system (less than 4-dimensions), whereas it is difficult for high-dimension system because it is very hard to obtain all the values of the Nth ($N \geq 5$) power algebraic equation with parameters.

A lot of work have been done by researchers to solve Hopf bifurcation problem for a dynamic system. Wencheng [1] and Jiaqi [2] have proposed direct algebraic criterion of Hopf bifurcation existing problem by using Hurwitz determinant. Based on nonlinear model theory, an analytic method of Hopf

G. R. Liu et al. (eds.), Computational Methods, 1755–1760.

bifurcation was presented by Weiping [3]. Jike [4] converted the eigen-polynomial problem crossing the image axe by using a transforming Jacobian matrix. According to the control theorem of the electronic system, Hamdan and Hamdan [5] presented Hopf bifurcation method for sub-synchronous resonance.

In this chapter, the condition causing Hopf bifurcation is used to construct the numerical algorithm directly according to the theorem of Hopf bifurcation.

2. METHOD ANALYSIS

Considering an ordinary autonomous nonlinear system as follows.

$$\dot{x} = f(x, \alpha), \quad x \in R^n, \quad \alpha \in R \tag{1}$$

When $f(x_0(\alpha), \alpha) = 0$, Jacobian matrix of system (1) is

$$J(\alpha) = D_x f(x, \alpha) = \left[\frac{\partial f_i}{\partial x_j} \right], \quad i, j = 1, 2, \ldots, n \tag{2}$$

Let the eigen-polynomial of $J(\alpha)$ of system (1) is

$$P_n(\lambda) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_{n-1}\lambda + a_n \tag{3}$$

Where $P_n(\lambda)$ is the nth power polynomial of λ.

It can be seen from the classical Hopf bifurcation theory that the necessary existence condition of Hopf bifurcation in system (1) is that the eigenequation $P_n(\lambda) = 0$ has a couple of plural conjugate pure image roots, namely

$$\lambda_{1,2} = \pm i\sqrt{\omega} \tag{4}$$

So $P_n(\lambda)$ can be decomposed at Hopf bifurcation point as follows.

$$P_n(\lambda) = (\lambda^2 + \omega)P_{n-2}(\lambda) \tag{5}$$

Where $P_{n-2}(\lambda)$ is the (n–2)th power polynomial.

For an estimate of ω, the polynomial (3) is transformed as follows.

$$P_n(\lambda) = (\lambda^2 + \omega)(\lambda^{n-2} + b_1\lambda^{n-3} + \cdots + b_{n-3}\lambda + b_{n-2}) + F_1\lambda + F_2 \tag{6}$$

Expanding Equation 6 and simplifying it, the following expression can be obtained.

$$\begin{aligned} P_n(\lambda) = \lambda^n + b_1\lambda^{n-1} + (b_2 + \omega)\lambda^{n-2} + (b_3 + \omega b_1)\lambda^{n-3} \cdots \\ + (b_{n-2} + \omega b_{n-4})\lambda^2 + (F_1 + \omega b_{n-3})\lambda + (F_2 + \omega b_{n-2}) \end{aligned} \tag{7}$$

Comparing Equation 7 with Equation 3, the following recursive relationship of the coefficients a_i, b_i, F_1 and F_2 can be obtained.

$$\left.\begin{aligned} &b_{-1} = 0 \\ &b_0 = 1 \\ &b_i = a_i - \omega b_{i-2},\ i = 1, 2, \ldots, n-2 \\ &F_1 = a_{n-1} - \omega b_{n-3} \\ &F_2 = a_n - \omega b_{n-2} \end{aligned}\right\} \tag{8}$$

The coefficients a_i, b_i, F_1 and F_2 are the function of x, α and ω.

The existence condition of Hopf bifurcation in system (1) is that the eigenequation $P_n(\lambda) = 0$ has a couple of pure image roots at Hopf bifurcation point, which also means that the eigen-polynomial $P_n(\lambda)$ can be decomposed in the form of (5).

In Equation 6, if F_1 and F_2 are all zero, namely

$$\left.\begin{aligned} F_1 = F_1(x, \alpha, \omega) = 0 \\ F_2 = F_2(x, \alpha, \omega) = 0 \end{aligned}\right\} \tag{9}$$

then Equation 6 will be transformed into Equation 5. And Equation 5 is tenable only at Hopf bifurcation point. So, if some x, α and ω which can satisfy Equation 9 can be find out, then α is Hopf bifurcation point in system (1) and $\pm i\sqrt{\omega}$ are the couple of pure image roots of the eigenequation of Jacobian matrix at Hopf bifurcation point.

According to the analyses above, the method of determining Hopf bifurcation point is constructed by using the conditions that F_1 and F_2 are both equal to zero and Hopf bifurcation point must be the balanced point of the system (1). Combining Equation 1 and Equation 9, a set of (n+2) order nonlinear equation is formed as follows.

$$\left.\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2, \ldots, x_n, \alpha) = 0 \\ \dot{x}_2 &= f_2(x_1, x_2, \ldots, x_n, \alpha) = 0 \\ &\ \vdots \\ \dot{x}_n &= f_n(x_1, x_2, \ldots, x_n, \alpha) = 0 \\ F_1 &= F_1(x_1, x_2, \ldots, x_n, \alpha, \omega) = 0 \\ F_2 &= F_2(x_1, x_2, \ldots, x_n, \alpha, \omega) = 0 \end{aligned}\right\} \tag{10}$$

In practical application, the expressions of $f_1, f_2, \ldots, f_n$ are very complicated and the eigen-polynomial of Jacobian matrix cannot be analytically expressed. So, the Newton–Raphson iterative method is used to solve Equation 10.

The specific procedure to form F_1, F_2 is shown as follows.

(a) Select an initial iterative vector $(x_1^0, x_2^0, \ldots, x_n^0, \alpha^0, \omega^0)$ and a small increment of the iterative vector $(\Delta x_1^0, \Delta x_2^0, \ldots, \Delta x_n^0, \Delta\alpha^0, \Delta\omega^0)$ and the corresponding calculation precision value ε. Expand Equation 10 to Taylor series near $x_1^0, x_2^0, \ldots x_n^0, \alpha^0, \omega^0$, and only remain the linear part as follows.

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} & \frac{\partial f_1}{\partial \alpha} & 0 \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} & \frac{\partial f_2}{\partial \alpha} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} & \frac{\partial f_2}{\partial \alpha} & 0 \\ \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_n} & \frac{\partial F_1}{\partial \alpha} & \frac{\partial F_1}{\partial \omega} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_n} & \frac{\partial F_2}{\partial \alpha} & \frac{\partial F_2}{\partial \omega} \end{bmatrix} \begin{Bmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \\ \Delta\alpha \\ \Delta\omega \end{Bmatrix} = - \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \\ F_1 \\ F_1 \end{bmatrix} \tag{11}$$

(b) Calculate $f_1^x, f_2^x, \ldots, f_n^x$ in Equation 1 at $(x_1^0, x_2^0, \ldots x_n^0, \alpha^0, \omega^0)$.
(c) Let $x_i^1 = x_i^0 + \Delta x_i\ (i = 1, 2, \ldots, n), \alpha^1 = \alpha^0 + \Delta\alpha$ and $\omega^1 = \omega^0 + \Delta\omega$, then calculate $f_1^{x+\Delta x_i}, f_2^{x+\Delta x_i}, \ldots, f_n^{x+\Delta x_i}\ (i = 1, 2, \ldots, n)$ in Equation 1.
(d) Calculate all of $\frac{\partial f_i}{\partial x_j}\ (i, j = 1, 2, \ldots, n)$ and all of $\frac{\partial f_i}{\partial \alpha}\ (i = 1, 2, \ldots, n)$.
(e) Calculate the coefficients $a_1, a_2, \ldots, a_n$ in the polynomial (3) according to the formula (12) beneath at $(x_1^0, x_2^0, \ldots x_n^0, \alpha^0, \omega^0)$.

$$a_i = (-1)^i \sum_{1 \le k_1 < k_2 < \cdots < k_i \le n} \begin{vmatrix} J_{k_1k_1} & J_{k_1k_2} & \cdots & J_{k_1k_i} \\ J_{k_2k_1} & J_{k_2k_2} & \cdots & J_{k_2k_i} \\ \vdots & \vdots & \ddots & \vdots \\ J_{k_ik_1} & J_{k_ik_2} & \cdots & J_{k_ik_i} \end{vmatrix}, \quad i = 1, 2, \ldots, n \tag{12}$$

(f) Calculate F_1^x and F_2^x following the formula (8).
(g) Let $x_i^1 = x_i^0 + \Delta x_i\ (i = 1, 2, \ldots, n), \alpha^1 = \alpha^0 + \Delta\alpha$ and $\omega^1 = \omega^0 + \Delta\omega$, then calculate $a_1, a_2, \ldots, a_n, F_1^{x+\Delta x_i}, F_2^{x+\Delta x_i}\ (i = 1, 2, \ldots, n), F_j^{x+\Delta\alpha}$ and $F_j^{x+\Delta\omega}\ (i = 1, 2, \ldots, n)$.
(h) Calculate all of $\frac{\partial F_i}{\partial x_j}\ (i = 1, 2, j = 1, 2, \ldots, n)$, $\frac{\partial F_i}{\partial \alpha}$ and $\frac{\partial F_i}{\partial \omega}\ (i = 1, 2)$.
(i) The values of $\Delta x_1^0, \Delta x_2^0, \ldots, \Delta x_n^0, \Delta\alpha^0, \Delta\omega^0$ can be obtained by solving Equation 11.
(j) Verify the precision, stop if the request precision is satisfied, otherwise Let $x_i^1 = x_i^0 + \Delta x_i\ (i = 1, 2, \ldots, n), \alpha^1 = \alpha^0 + \Delta\alpha$ and $\omega^1 = \omega^0 + \Delta\omega$, then go bank to (b).

3. CALCULATION EXAMPLE

The motion equations of the rotor system supported by the bearing with nonlinear oil film in state space [6] are as follows.

$$\begin{cases} \dot{x}_1 = x_3 \\ \dot{x}_2 = x_4 \\ \dot{x}_3 = \frac{1}{\omega^2} - \frac{s}{\omega}(f_r \cos(\varphi) + f_t \sin(\omega)) + e\cos(t) \\ \dot{x}_4 = -\frac{s}{\omega}(f_r \sin(\varphi) - f_t \cos(\omega)) + e\sin(t) \end{cases} \tag{13}$$

When $e = 0$, Equation 13 represents the motion equation of the balanced rotor system, which is an autonomous system. According to the results of reference [10], the motion form of the system is a stable movement at the balanced point when the rotational speed is low. While the rotational speed is higher than some limits (Hopf bifurcation point), the motion track is a stable movement of the limit circle.

When $e = 0$ and $s = 1.4$, select an initial iterative vector $(x, y, \dot{x}, \dot{y}) = (0.2, 0.1, 0.5, 0.5)$, an initial iterative velocity $\omega_H^0 = 0.8$ and initial $\lambda = 0.5$, then according to the method above, the bifurcation point of the system can be obtained, that is $(x, y, \dot{x}, \dot{y}) = (0.12434682, 0.41090873, 0, 0)$, the bifurcation parameter $\omega_H = 1.05543263$, the couple of the conjugate pure image roots of Jacobian matrix $\lambda_{1,2} = \pm 0.57361864i$.

4. CONCLUSIONS

A numerical algorithm is constructed for the Hopf bifurcation of an autonomous nonlinear system. The problem is converted to solving a set of nonlinear equations firstly. The bifurcation point and the couple of pure image roots of eigen-polynomial of Jacobian matrix at bifurcation point in the system can be both solved at the same time. It is crucial to construct a set of nonlinear equations that includes bifurcation parameters and pure image roots of the eigen-matrix at the bifurcation point in the system. A semi-analytic method is adopted during the constructing process. The expensive calculation problem with many existing methods, which needs to calculate the eigen-values and judges whether its real part is zero or not whenever the parameters cha, is solved in this method.

ACKNOWLEDGEMENT

This work was supported by *the National Science Foundation of China* (Item No. 10172067) and *the Doctor Foundation of Xi'an University of Technology* (Item No. 202-20316)

REFERENCES

1. Wencheng Chen and Guoliang Chen (1992), An algorithm criterion for Hopf bifurcation. *Acta Mathematicae Application Sinica*, 15, 2, pp. 251–259.(In Chinese)
2. Jiaqi Shen and Zhujun Jing (1995), A new detecting method for conditions for existence of Hopf bifurcation. *Acta. Mathematics Applicatae Sinica*, 11, 1, 79–93.(In Chinese)
3. Weiping Fu, Jianxue Xu and Xinhua Zhang (1997), Method of nonlinear normal mode and its application in analysis of Hopf bifurcation in electromechanically coupled subsynchronous torsional oscillating system. *Chinese Journal of Applied Mechanics*, 17, 3, pp. 175–178.
4. Jike Wu and kun Zhou (1993), Numerical computation for high dimension Hopf bifurcation. *Acta Scientiarum Naturalium of Universitatis Pekinensis*, 29, 5, pp. 574–582. (In Chinese)
5. A.M.A. Hamdan and H.M.A. Hamdan (1993), Use of SVD for power system stabilizer signal selection. *Electric machines and power system*, 21, 2, pp. 229–240.
6. Dexin Li (2003), Research on determining the periodic solution of nonlinear system and the dynamic behaviors of rotor-bearing system [Dissertation of PHD], Xi’an: Xi’an Jiaotong University. (In Chinese)

STUDY ON NONLINEAR DYNAMIC BEHAVIOURS AND STABILITY OF A FLEXIBLE ROTOR SYSTEM WITH HYDRODYNAMIC SLIDING BEARING SUPPORTS

Yanjun Lu[1], Yongfang Zhang[2], Heng Liu[1], Lie Yu[1], Dexin Li[3], Lijun Jiang[4]
[1]*Theory of Lubrication and Bearing Institute, Xi'an Jiaotong University, 710049, P. R. China*
[2]*School of Electronic and Information, Northwest Polytechnical University, 710072, P. R. China*
[3]*School of Machinery and Precision Instrument Engineering, Xi'an University of Technology, 710048, P. R. China*
[4]*Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613, Singapore*

Abstract This paper investigates the nonlinear dynamic behaviours of a flexible rotor system with elliptical bearing supports. Based on the variational constraint approach, nonlinear oil film forces and their Jacobians are calculated simultaneously. According to the local nonlinearity of the flexible rotor system, a modal reduction technique based on a free-interface component mode synthesis technique is utilized to reduce linear degrees-of-freedom of the system. Nonlinear unbalance periodic responses of the system are obtained by using PNF (Poincaré-Newton-Floquet) method and the bifurcation point of system can be calculated by using the combination of the predictor–corrector mechanism and the PNF method. The local stability and bifurcation behaviours of periodic motions are analysed by the Floquet theory. The numerical examples show that the schemes of this study not only save computing efforts but also have good precision.

Keywords: nonlinear dynamics, journal bearing-rotor system, finite element method, bifurcation, stability.

1. INTRODUCTION

The nonlinear oil film forces act on few nodal points of rotor individually, thus bearing-rotor system is a local nonlinear system. A modal reduction method for reducing a high-order dynamical system with local nonlinearity

G. R. Liu et al. (eds.), Computational Methods, 1761–1765.

was shown by Zheng and Hasebe [1]. Based on modal synthesis technique with fixed-interface, a method for reducing linear degrees-of-freedom of a dynamic system with local nonlinearity was presented by Zhang et al. [2]. In this paper, a modified modal synthesis technique with free-interface is proposed to reduce linear degrees-of-freedom of a flexible rotor system. After the reduction of degrees-of-freedom, the system keeps the nonlinearity. Elliptical variational equations with Reynolds boundary are solved by finite element method. The unbalance responses bifurcation point of the system are obtained by PNF (Poincaré-Newton-Floquet) method and a method consisting of predictor–corrector mechanism and PNF method.

2. THE BEARING-FLEXIBLE ROTOR SYSTEM EQUATIONS OF MOTION

A typical bearing-rotor system with local nonlinearity is shown in Figure 1. Using the finite element method, flexible shaft equations of lateral motion can be written as

$$\boldsymbol{M}^s\ddot{\boldsymbol{X}}^s + \boldsymbol{G}^s\dot{\boldsymbol{X}}^s + \boldsymbol{K}^s\boldsymbol{X}^s = \boldsymbol{Q}^s + \boldsymbol{f}^s(\boldsymbol{X}^s, \dot{\boldsymbol{X}}^s) \tag{1}$$

where $\boldsymbol{M}^s \in R^{n\times n}$,$\boldsymbol{G}^s \in R^{n\times n}$,$\boldsymbol{K}^s \in R^{n\times n}$, $\boldsymbol{Q}^s \in R^n$ and $\boldsymbol{f}^s(\boldsymbol{X}^s, \dot{\boldsymbol{X}}^s) \in R^n$ are the mass matrix, gyroscope matrix, stiffness matrix, external force vector and nonlinear force vector, respectively. For a shaft with p nodal points, the displacement vector is of the form

$$\boldsymbol{X}^s = [x_1, y_1, \varphi_1, \psi_1, x_2, y_2, \varphi_2, \psi_2, \cdots, x_p, y_p, \varphi_p, \psi_p]^T \tag{2}$$

where $x_j, y_j, \varphi_j, \psi_j (j = 1, 2, \cdots, p)$ are the lateral translations and rotation angles of the jth nodal point along the horizontal and vertical direction, respectively. For rotor system supported by m bearings, the forces of bearings are of the local features as follows

$$\boldsymbol{f}^s(\boldsymbol{X}^s, \dot{\boldsymbol{X}}^s) = [\boldsymbol{f}^s_b(\boldsymbol{X}^s_b, \dot{\boldsymbol{X}}^s_b), \boldsymbol{0}]^T \tag{3}$$

where $\boldsymbol{X}^s_b \in R^{4m}$,$\boldsymbol{f}^s_b(\boldsymbol{X}^s_b, \dot{\boldsymbol{X}}^s_b) \in R^{4m}$ can be written as

$$\boldsymbol{X}^s_b = [x_1, y_1, \varphi_1, \psi_1, \cdots, x_m, y_m, \varphi_m, \psi_m]^T \tag{4}$$

$$\boldsymbol{f}^s_b(\boldsymbol{X}^s_b, \dot{\boldsymbol{X}}^s_b) = [f_{xj}, f_{yj}, 0, 0, \cdots, f_{xm}, f_{ym}, 0, 0]^T \tag{5}$$

Figure 1. Sketch of a bearing-rotor system.

where f_{xj} and f_{yj} are the horizontal and vertical oil film forces acting on the jth point of the shaft. To simplify notations, the order of the vector components is rearranged and Equation (1) can be partitioned as

$$\begin{bmatrix} \boldsymbol{M}_{bb}^{s} & \boldsymbol{M}_{bc}^{s} \\ \boldsymbol{M}_{cb}^{s} & \boldsymbol{M}_{cc}^{s} \end{bmatrix} \begin{Bmatrix} \ddot{\boldsymbol{X}}_{b}^{s} \\ \ddot{\boldsymbol{X}}_{c}^{s} \end{Bmatrix} + \begin{bmatrix} \boldsymbol{G}_{bb}^{s} & \boldsymbol{G}_{bc}^{s} \\ -\boldsymbol{G}_{bc}^{s^{T}} & \boldsymbol{G}_{cc}^{s} \end{bmatrix} \begin{Bmatrix} \dot{\boldsymbol{X}}_{b}^{s} \\ \dot{\boldsymbol{X}}_{c}^{s} \end{Bmatrix} + \begin{bmatrix} \boldsymbol{K}_{bb}^{s} & \boldsymbol{K}_{bc}^{s} \\ \boldsymbol{K}_{cb}^{s} & \boldsymbol{K}_{cc}^{s} \end{bmatrix} \begin{Bmatrix} \boldsymbol{X}_{b}^{s} \\ \boldsymbol{X}_{c}^{s} \end{Bmatrix}$$
$$= \begin{Bmatrix} \boldsymbol{Q}_{b}^{s} \\ \boldsymbol{Q}_{c}^{s} \end{Bmatrix} + \begin{Bmatrix} \boldsymbol{f}_{b}^{s}(\boldsymbol{X}_{b}^{s}, \dot{\boldsymbol{X}}_{b}^{s}) \\ \boldsymbol{0} \end{Bmatrix} \tag{6}$$

where $\boldsymbol{X}_{b}^{s}$ and $\boldsymbol{X}_{c}^{s}$ are nonlinear degrees-of-freedom and the linear degrees-of-freedom, respectively. $\boldsymbol{X}^{s}$ is written as a linear combination of n_d columns

$$\boldsymbol{X}^{s} = \boldsymbol{T}_{1}\boldsymbol{p}', \boldsymbol{T}_{1} = [\psi_{b}, \psi_{c}], \begin{Bmatrix} \boldsymbol{p}_{b}^{s} \\ \boldsymbol{p}_{k}^{s} \end{Bmatrix} = \begin{bmatrix} \psi_{bb}^{-1} & -\psi_{bb}^{-1}\psi_{bk} \\ \boldsymbol{0} & \boldsymbol{I}_{kk} \end{bmatrix} \begin{Bmatrix} \boldsymbol{X}_{b}^{s} \\ \boldsymbol{p}_{k}^{s} \end{Bmatrix} = \mathbf{p}' = \boldsymbol{T}_{2}\boldsymbol{p} \tag{7}$$

where the columns of matrix $\psi_c \in R^{n \times n_k}$ with the kept elastic eigenmodes are the mass normalized solutions $(\boldsymbol{K}^{s} - \omega_j^2 \boldsymbol{M}^{s})\psi_j = \boldsymbol{0}\,(j = 1, \cdots n_k)$ of the undamped eigenproblem for $\omega_k \in (0, \omega_{cut})$. The columns of the matrix $\psi_b \in R^{n \times n_b}$ with the residual flexibility modes is calculated as shown in the work by Craig [3]. The reduced equations become

$$\boldsymbol{T}^{T}\boldsymbol{M}^{s}\boldsymbol{T}\dot{\boldsymbol{p}} + \boldsymbol{T}^{T}\boldsymbol{G}^{s}\boldsymbol{T}\dot{\boldsymbol{p}} + \boldsymbol{T}^{T}\boldsymbol{K}^{s}\boldsymbol{T}\boldsymbol{p} = \boldsymbol{T}^{T}\boldsymbol{Q}^{s} + \boldsymbol{T}^{T}\boldsymbol{f}^{s} \tag{8}$$

where $\boldsymbol{T} = \boldsymbol{T}_1\ \boldsymbol{T}_{2.}$ From Equations (7) and (8), it is evident that the unbalance forces of shaft and nonlinear effects definitely are remained. The reduced dynamic system is given by

$$\boldsymbol{M}\ddot{\boldsymbol{q}} + \boldsymbol{G}\dot{\boldsymbol{q}} + \boldsymbol{K}\boldsymbol{q} = \boldsymbol{Q} \tag{9}$$

When state variables $\boldsymbol{X} = (\boldsymbol{q}, \dot{\boldsymbol{q}})^{T}$ are introduced, the corresponding system equations in state space are

$$\dot{\boldsymbol{X}} = \begin{Bmatrix} \dot{\boldsymbol{q}} \\ \boldsymbol{M}^{-1}(\boldsymbol{Q} - \boldsymbol{G}\dot{\boldsymbol{q}} - \boldsymbol{K}\boldsymbol{q}) \end{Bmatrix} \tag{10}$$

3. CALCULATIONS OF NONLINEAR FORCES AND JACOBIANS OF HYDRODYNAMIC BEARING

Since the Reynolds boundary problem arising in fluid lubrication is equivalent to variational inequalities as shown in the work by Kinderlenhrer and Stanpacchia [4], nonlinear oil film forces and their Jacobians are calculated simultaneously as shown in the work by Lu et al. [5].

4. A PERIODIC SOLUTIONS METHOD CONSISTING OF PREDICTOR-CORRECTOR MECHANISM AND POINCARÉ-NEWTON-FLOQUET

Poincaré-Newton-Floquet (PNF) method covers solution of a two-point boundary value problem and stability analysis of Floquet bifurcation theory. A periodic solutions method consisting of predictor-corrector mechanism and Poincaré-Newton-Floquet Floquet theory as shown in the work by Lu et al. [5].

5. NUMERICAL EXAMPLES AND RESULTS

The unbalance response of a bearing-rotor system with a rigid disk (noted D1) and two elliptical bearing supports (noted B1, B2, pad arc: 150 deg, dynamic viscosity of oil: 0.0287 Pas, width-to-diameter ratio: B/d = 0.8). The shaft (diameter 0.45 m, length 8.4 m, Young's modulus 2.0×10^{11} N/m^2, shear modulus 7.6923×10^{10} N/m^2, mass density 7800 kg/m^3) is equally discretized into six finite elements. Mass eccentricity of the shaft ($e_x^s = e_y^s = 1\,\mu\text{m}$) and the disk ($e_x^{D1} = 0$, $e_y^{D1} = 1\,\mu\text{m}$) has the same rotating phase angle. Eight eigenmodes are taken in the examples. The influence of modal reduction on the accuracy of the result is analysed. For clearance ratio $\psi = 0.003$, elliptical ration $\delta = 0.556$, $\omega = 1000$ r/min, the periodic solutions of the system were solved by eight eigenmodes model and full eigenmodes model, respectively. If an initial iterative position is given as $\boldsymbol{X}_s(t_0)$, periodic solutions $\boldsymbol{X}_s$ are solved by the PNF method. The kth periodic errors is ε_k by the PNF method. The mode of the leading Floquet multiplier is $|f_{\max}| = 0.938861 = -0.0637515 + 0.936694j$. Periodic solutions at B1 station are $x_{B1} = 2.001 \times 10^{-4}$, $y_{B1} = 2.128 \times 10^{-5}$, $\varphi_{B1} = 7.443 \times 10^{-4}$, $\psi_{B1} = -9.405 \times 10^{-7}$ by reduced model (eight eigenmodes model). The mode of the leading Floquet multiplier is $|f_{\max}| = 0.938924 = -0.064017 + 0.936739j$. Periodic solution at B1 station are $x_{B1} = 1.998 \times 10^{-4}$, $y_{B1} = 2.116 \times 10^{-5}$, $\varphi_{B1} = 7.460 \times 10^{-4}$, $\psi_{B1} = -9.423 \times 10^{-7}$ by the original model (i.e., full freedom-of-freedom model). Table 1 infer that the nonlinear oil film forces and their

Table 1. Periodic errors calculated by the different model.

Time: k	1	2	3	4
Error ε_k (reduced model):	9.99875×10^{-1}	1.70111×10^{-2}	8.86754×10^{-4}	1.37143×10^{-8}
Error ε_k (original model):	9.99832×10^{-1}	1.88753×10^{-2}	9.57457×10^{-4}	1.65259×10^{-8}

Jacobians have compatible accuracy and enough accuracy has been achieved when eight eigenmodes are used. For $\psi = 0.003$, $\delta = 0.556$, groove-to-width ratio $B/B = 0.4$, $0 < \omega < 1526$ r/min (i.e., $\omega = 1526$ r/min, $|f_{\max}| = 1.0000$ by using the established method), the periodic solution is stable.

6. CONCLUSION

For the localized nonlinear feature of bearing-flexible rotor system, a modified modal synthesis technique with free-interface is presented to reduce linear degrees-of-freedom of model of the flexible rotor system. The accuracy of nonlinear analysis of the system is ensured and the computing costs are cut down. Variational constraint approach is introduced to continuously revise the variational form of Reynolds equation. A periodic solution method consisting of the predictor-corrector mechanism and the PNF method are presented to calculate the periodic motions and bifurcation points of the system. The numerical examples show that the schemes of this study not only save computing efforts but also have good precision.

ACKNOWLEDGEMENT

Project is supported by *National Natural Science Foundation of China* (No. 50275116) and *National 863 of China* (No. 2002AA414060, 2002AA503020).

REFERENCES

1. T.S. Zheng and N. Hasebe (1999). An efficient analysis of high order dynamic system with local nonlinearity. *ASME Journal Vibration and Acoustics*, 121, pp. 408–416.
2. J.Z. Zhang, Q.Y. Xu and T.S. Zheng (1998), A method for determining the periodic solution and its stability of a dynamic system with local nonlinearities (in Chinese). *Acta Mechanica Sinica*, 30, 5, pp. 572–579.
3. R.R. Craig Jr (1985), A review of time-domain and frequency-domain component modes synthesis methods. *Combined Experimental/Analytical Modeling of Dynamic Structural Systems Using Substructure Synthesis*, D.R. Martinez and A.K. Miller (ed.), pp. 1–31.
4. D. Kinderlenhrer and G. Stanpacchia (1980), *An Introduction to Variational Inequalities and Their Applications*. Academic Press, NewYork.
5. Y.J. Lu, L. Yu and H. Liu (2004), Stability and bifurcation of nonlinear bearing-rotor system (in Chinese). *Chinese Journal of Mechanical Engineering*, 40, 10, pp. 62–67.

ON THE IMPERFECTIONS OF CYLINDRICAL SHELLS ON LOCAL SUPPORTS

Wesley Vanlaere, Rudy Van Impe, Guy Lagae, K.B. Katnam
and Matthieu De Beule
Laboratory for Research on Structural Models, Ghent University,
Technologiepark-Zwijnaarde 904, B9052 Zwijnaarde, Belgium

Abstract This paper deals with stringer stiffened cylinders on local supports. For shell structures, the imperfections of the constructions tend to have a major influence on the structural behaviour. In this contribution, the results of a number of numerical simulations of discretely supported cylinders with different imperfection shapes are presented.

Keywords: cylindrical shell, buckling, imperfection, numerical simulation.

1. INTRODUCTION

In the field of structural engineering, thin-walled shell structures are frequently applied. As some examples of this kind of structures, pressure vessels, tanks, silos and pipelines can be mentioned. The behaviour of this kind of structures is however complex and the present understanding is mainly due to the extensive research that was conducted in the past by renowned researchers as Timoshenko [1], Flügge [2], Koiter [3],... In the first decades of last century, this research was mainly on a theoretical base and led to an expression for the buckling stress of a perfect elastic cylinder subjected to uniform axial compression (*classical elastic critical stress*). The correspondence of this expression with the experiments however was dreadful: real cylinders buckled at a distinctly lower load and furthermore the scatter of the results was remarkable. In the decades that followed, this discrepancy was investigated and nowadays the presence of unavoidable geometrical imperfections in the shell was is widely accepted as one of the main reasons for this discrepancy.

G. R. Liu et al. (eds.), Computational Methods, 1767–1771.

Figure 1. An example of a stiffened cylinder.

In the past century, the focus was mainly on uniformly loaded cylinders, because the structural behaviour of this shell shape had to be thoroughly understood before shell structures with increased complexity could be investigated. At the dawn of the new century, the situation has changed and cylindrical shells subjected to other load cases are also intensely studied. The traditional build-up of a elevated steel silo consists of a cylindrical body and a conical bottom. The silo is supported by a limited number of columns which are attached to the cylindrical wall at the joint between the two shell structures. For the cylindrical shell, this supporting method leads to a non-uniform loading and local instability of the cylinder can be expected if the shell thickness is to low. Evidently this instability has to be avoided and one way of preventing this failure phenomenon is increasing the wall thickness of the entire cylinder. This is investigated by Guggenberger et al. [4]. The possible deformation pattern of locally supported cylinders is however very localized and therefore more economical ways of preventing the instability may be possible. The present research of the authors deals with the structural behaviour of locally supported cylinders where the possible buckling zone is stiffened by means of longitudinal stiffeners in combination with two ring stiffeners. Since the buckles can appear in the cylinder in the zone just above a support, two longitudinal stiffeners with limited height are attached to the cylindrical wall above each support. Above and below these stringers, ring stiffeners are welded to the cylinder. An example of such a stiffened cylinder can be seen in Figure 1.

For this kind of cylinders, adequate design rules have yet to be developed. Therefore, it is the goal of the present research to develop these design rules for stringer stiffened cylinders on local supports. In this contribution, the focus is on the effect of the geometrical imperfections on these structures.

2. THE NUMERICAL MODEL

The research is divided into two main parts: numerical simulations of the cylindrical shells and experiments on scale models as a validation of the numerical model. An example of a scale model used in the experiments is shown in Figure 1. In the test set-up, the cylinder is placed upon four supports. The main dimensions of the cylinder and the stiffeners are summarized in Table 1. For more information regarding the test set-up, refer to Van Impe et al. [5]. A numerical model with the dimensions of Table 1 was developed and numerical simulations with the finite element package ABAQUS were performed. The supports were modelled as rigid and the top rim of the cylinder was subjected to a uniform axial load. This numerical model was validated with the results of a number of tests on scale models [6].

With this model, a number of simulations were performed. In this contribution, the results of geometrically and materially non-linear analyses of imperfect shell structures (GMNIA) are discussed. The results of five analyses are summarized in Table 2. The first column of Table 2 indicates the imperfection shape that was taken. For each simulation, an eigenmode (EM) of the perfect structure

Table 1. Dimensions of the cylinder and the stiffeners.

Characteristic	Value [mm]	Characteristic	Value [mm]
Cylinder radius	350	Stringer spacing	37
Cylinder wall thickness	0.7	Support width	80
Cylinder length	700	Thickness upper ring	1.5
Stringer width	20	Width upper ring	20
Stringer height	280	Thickness lower ring	0.7
Stringer thickness	1.5	Width lower ring	45

Table 2. Results of the numerical simulations.

Imperfection shape	Failure load [kN]	Failure pattern
EM.1	78.3	Stiffener
EM.2	78.1	Stiffener
EM.3	73.8	Upper ring
EM.4	73.9	Upper ring
EM.5	77.5	Upper ring

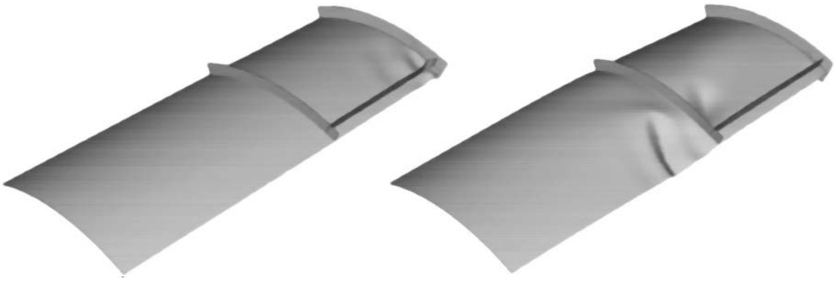


Figure 2. Failure pattern *stiffener* (left) and *upper ring* (right).

was chosen as imperfection shape and the imperfection amplitude was 3 mm. So, for each of the first five eigenmodes of the perfect structure, a simulation with this imperfection shape was performed. In the second column, the failure load of this structure is given, while the third column indicates the type of failure pattern. Basically, two possible failure patterns exist. The first failure pattern is characterized by buckles in the stringers and the cylindrical shell near the supports. Here, this failure pattern is called *stiffener*. For the second failure pattern, the buckles appear in the cylindrical shell above the upper ring. Secondary buckles may also appear just below this upper ring. This failure pattern is called *upper ring*. Both failure patterns are shown in Figure 2. In order to be able to draw conclusions from these result, the applied imperfection shapes have to be explained. For eigenmodes 1, 2 and 5, the buckles appear near the support. For eigenmodes 3 and 4, the buckles appear above the upper ring. The results of Table 2 clearly indicate that this imperfection shape has a strong influence on the failure pattern as well as on the failure load. Furthermore, the results show that the first eigenmode is not necessarily the most detrimental imperfection shape.

3. CONCLUSIONS

In this contribution, the effect of the imperfections on the structural behaviour of stiffened cylindrical shells on local supports is investigated. These imperfections have a major influence on the failure load and failure pattern. Furthermore, the results show that the first eigenmode of the perfect construction is not necessarily the most detrimental imperfection shape.

ACKNOWLEDGEMENTS

The authors would like to thank the Fund for Scientific Research—Flanders (Belgium) (F.W.O.—Vlaanderen) for its financial support.

REFERENCES

1. S.P. Timoshenko (1910), Eininge stabilitätsprobleme der elasticitätstheorie. *Zeitschrift für Mathematik und Physik*, 58, pp. 378–385.
2. W. Flügge (1932), Die stabilität der kreiszylinderschale. *Ingenieur Archiv*, 3, pp. 463–506.
3. W.T. Koiter (1945), *Over de stabiliteit van het elastisch evenwicht*, Ph.D Thesis, TU Delft, The Netherlands.
4. W. Guggenberger, R. Greiner and J.M. Rotter (2000), The behaviour of locally-supported cylindrical shells: unstiffened shells. *Journal of Constructional Steel Research*, 56, pp. 175–197.
5. R. Van Impe, J. Belis, P. Buffel and G. Lagae (2001), Experimental buckling analysis of cylinders with discrete stringer stiffeners above the local supports, In: *Ninth Nordic Steel Construction Conference*, Helsinki, Finland, pp. 221–228.
6. W. Vanlaere, G. Lagae, R. Van Impe, P. Buffel, J. Belis and M. De Beule (2003), Boundary conditions for stringer stiffened cylindrical shells on local supports, In: *International Conference on Design, Inspection, Maintenance and Operations of Cylindrical Steel Tanks and Pipelines*, Prague, Czech Republic, pp. 55–61.

STUDY ON DOUBLY PERIODIC RIGID LINE INCLUSIONS UNDER ANTIPLANE SHEAR

Y.L. Xu and L. Huang
Department of Engineering Mechanics, Yanshan University, Qinhuangdao 066004, P. R. China

Abstract The problem of doubly periodic rigid line inclusions in a rectangular array under a uniform antiplane shear stress at infinity is dealt with. By using the complex potential and the conformal mapping technique, the problem is reduced to a Hilbert one and the exact solution is obtained. The stress singularity factor at the tip of each rigid line is given in closed form and numerical results show the contrary dependence on the 'column' and 'stack' spacings of the rigid lines. Also, the effective longitudinal shear modulus of the inhomogeneous materials containing a rectangular array of rigid line inclusions is given as a function of the rigid line spacings.

Keywords: double period, rigid line, inhomogeneity materials, antiplane shear.

1. INTRODUCTION

Most solid materials contain some defects in the form of cracks, voids or inclusions which can cause serious stress concentrations and often lead to failure. From the viewpoint of inhomogeneities in solids, a slit crack and a rigid line (sometimes called a hard crack or an inverse crack) are the two extreme cases of a flat inhomogeneity, namely, $E \to 0$ for a slit crack and $E \to \infty$ for a rigid line, where E is Young's modulus. These are the most dangerous extreme cases. Crack problems have received much attention and have been widely studied. Rigid line inclusion problems have also aroused much interest. The references [1–11] provide examples of the contributions to the studies on rigid line inclusions. The interested readers consult also Chapter 11 of Ting's [12] book.

In many problems of multiple rigid lines, it is sufficient to assume that the stress field in the vicinity of each rigid line is approximately the same as the stress field around a single rigid line in a large (i.e., infinite) body. It will be

G. R. Liu et al. (eds.), Computational Methods, 1773–1782.

seen that rigid lines which are separated by less than several rigid line lengths interact to alter significantly the state of stress near the rigid line tips. Many natural and manufactural solid materials may possess a periodic microstructure. The study on doubly periodic rigid line inclusions is useful in understanding such an interesting interaction phenomena of the stress field among multiple rigid line inclusions. Obviously, exact solutions to such problems are most desirable. To our best knowledge, the exact solution to the doubly periodic rigid line problem has not been reported. It is noted, however, that several studies on doubly periodic cracks have been reported [13–17].

In this paper, the problem of doubly periodic rigid line inclusions in a rectangular array under a uniform antiplane shear stress at infinity is dealt with. By using the complex potential and a conformal mapping from a rectangular region to a half-plane, the problem is reduced to a Hilbert one. The exact solution is obtained, and the interactions among rigid line inclusions are studied.

2. STATEMENT OF THE PROBLEM

The problem to be considered is shown in Figure 1. An infinite elastic medium with shear modulus G contains doubly periodic rigid line inclusions parallel to y-axis in a rectangular array. The rigid lines along the same lines are said to form a 'column' and a 'stack' is formed in the direction normal to the

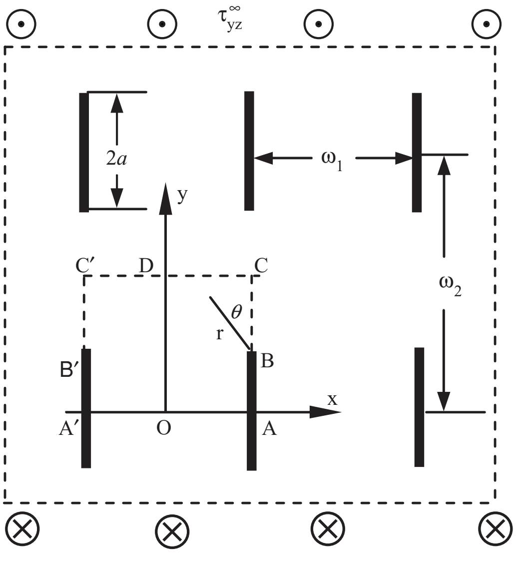


Figure 1. Doubly periodic rigid line array.

rigid lines. The length of each rigid line is 2a, and the periods are ω_1 and ω_2 in the directions of column and stack, respectively. $\boldsymbol{\tau}_{yz}^{\infty}$ is the uniform antiplane shear stress at infinity. $\boldsymbol{\tau}_{xz}^{\infty}$ is set to be zero since the stress field induced by it is not disturbed by the rigid lines.

To formulate the problem, we use the complex potential f(z) (an analytic function of the complex variable $z = x + iy$), in terms of which, the antiplane displacement w, antiplane shear stress components τ_{xz} and τ_{yz} can be expressed as [18–20]

$$w = [f(z) + \overline{f(z)}]/2 \tag{1}$$

$$\tau_{xz} - i\tau_{yz} = Gf'(z) = GF(z) \tag{2}$$

Where the overbar denotes the complex conjugate and the prime represents the derivative with respect to the argument.

According to the periodicity of the problem, consider a rectangle ABCDC′B′A′A (Figure 1). Since w is antisymmetric about A′A and C′C, $\partial w/\partial x = 0$ on A′A and C′C. Since the displacement w is symmetric about B′C′ and BC, $\partial w/\partial x = 0$ on B′C′ and BC. The rigid line is non-deformable, so $\partial w/\partial y = 0$ on A′B′and AB. From Equations (1) and (2), it is seen that $\partial w/\partial x = \mathrm{Re}F(z)$, $\partial w/\partial y = -\mathrm{Im}F(z)$, where Re and Im denote the real and imaginary parts of a complex function, respectively. Thus, the boundary condition of the rectangle ABCDC′B′A′A can be expressed as

$$\mathrm{Im}\,F(z) = 0 \quad \text{on A}'\text{B}' \text{ and AB} \tag{3}$$

$$\mathrm{Re}\,F(z) = 0 \quad \text{on the other boundaries} \tag{4}$$

To determine solely the solution of the problem, the resultant stress on the boundary CD must be prescribed. Noting the symmetry and periodicity of the problem, from the stress state at infinity and the equilibrium condition, it is seen that

$$\int_{CD} \tau_{yz}\, dx = \tau_{yz}^{\infty}\omega_1/2 \tag{5}$$

where the length of CD is $\omega_1/2$.

3. ANALYSIS AND SOLUTION

Making use of the Jacobian elliptical function

$$\zeta = \mathrm{sn}\left(\frac{z}{A}, k\right) \tag{6}$$

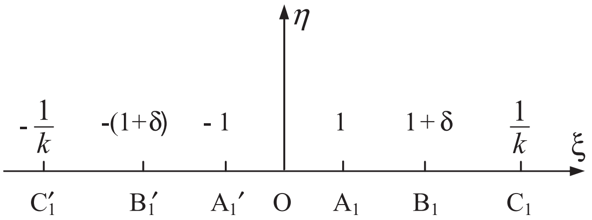


Figure 2. Conformal transformation (ζ-plane).

or

$$z = A\int_0^{\zeta} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}} \tag{7}$$

where A is a constant and k is the modulus of the function sn, the rectangular region ABCDC′B′A′A in the z-plane (Figure 1) is conformally transformed onto the upper half plane of the complex ζ−plane (Figure 2). The transformation characteristic points are indicated in Figure 2 (the point D in the z-plane is corresponding to the point infinity in the ζ-plane).

Now, we determine the coordinates of the transformation characteristic points. In the z-plane OA $= \omega_1/2$, AC $= \omega_2/2$. From Equation (7), it is seen that

$$\frac{\omega_1}{2} = A\int_0^{1} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}} \tag{8}$$

$$\frac{\omega_2}{2} i = A\int_1^{1/k} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}} \tag{9}$$

from Equations (8) and (9), constant A is determined as following

$$A = \frac{\omega_1}{2K(k)} = \frac{\omega_2}{2K(k')} \tag{10}$$

where k′ is called as the complementary modulus and

$$k'^2 = 1 - k^2 \tag{11}$$

K(k) and K(k′) are the complete elliptical integral of the first kind.

$$K(k) = \int_0^{1} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}} \tag{12}$$

From Equations (10) and (11), the constant A, modulus k and complementary modulus k′ can be determined.

In the z-plane, AB = a. From Equation (7), it is seen that

$$\mathrm{ia} = \mathrm{A}\int_{1}^{1+\delta}\frac{\mathrm{dt}}{\sqrt{(1-\mathrm{t}^2)(1-\mathrm{k}^2\mathrm{t}^2)}} \tag{13}$$

Letting

$$\mathrm{t} = \frac{1}{\sqrt{1-\mathrm{k}'^2\tau^2}} \tag{14}$$

Equation (13) becomes

$$\mathrm{a} = \mathrm{A}\int_{0}^{\frac{\sqrt{(1+\delta)^2-1}}{(1+\delta)\mathrm{k}'}}\frac{\mathrm{d}\tau}{\sqrt{(1-\tau^2)(1-\mathrm{k}'^2\tau^2)}} \tag{15}$$

From Equations (10) and (15), it is seen that

$$1+\delta = 1/\mathrm{dn}\left[\frac{2\mathrm{a}}{\omega_2}\mathrm{K}(\mathrm{k}'),\mathrm{k}'\right] \tag{16}$$

Letting

$$\mathrm{F}_1(\zeta) = \mathrm{iF(z)} = \mathrm{iF}\left[\mathrm{A}\int_{0}^{\zeta}\frac{\mathrm{dt}}{\sqrt{(1-\mathrm{t}^2)(1-\mathrm{k}^2\mathrm{t}^2)}}\right] \tag{17}$$

Equation (2) becomes

$$\tau_{\mathrm{yz}} + \mathrm{i}\tau_{\mathrm{xz}} = \mathrm{GF}_1(\zeta) \tag{18}$$

The substitution of Equation (17) into Equations (3) and (4) yields the boundary conditions of $\mathrm{F}_1(\zeta)$ on the real axis

$$\mathrm{ReF}_1(\zeta) = 0 \quad \text{on } [-(1+\delta),-1] \quad \text{and} \quad [1,(1+\delta)] \tag{19}$$

$$\mathrm{ImF}_1(\zeta) = 0 \quad \text{on the other parts} \tag{20}$$

The stress has singularities at the rigid tips and is equal to zero at the midpoints of rigid lines, so $\mathrm{F}_1(\zeta)$ has singularities of order $-1/2$ at $\zeta = \pm(1+\delta)$ and zeros of order 1/2 at $\zeta = \pm 1$. The stress is finite at the point D in the z-plane, so $\mathrm{F}_1(\zeta)$ is finite at infinity in the ζ-plane. Thus, the solution of the boundary problem can be expressed as

$$\mathrm{F}_1(\zeta) = \mathrm{C}_0\sqrt{\frac{\zeta^2-1}{\zeta^2-(1+\delta)^2}} \tag{21}$$

Where C_0 is a constant, which is determined by Equation (5). The substitution of Equation (21) into Equation (5) yields

$$C_0AG\int_{\frac{1}{k}}^{\infty}\frac{d\zeta}{\sqrt{[\zeta^2-(1+\delta)^2](k^2\zeta^2-1)}}=\frac{\omega_1}{2}\tau_{yz} \tag{22}$$

Letting

$$\zeta=\frac{1}{k\tau} \qquad k_1=k(1+\delta) \tag{23}$$

The substitution of Equations (23) and (10) into Equation (22) yields

$$C_0=\frac{\tau_{yz}^{\infty}}{G}\frac{K(k)}{K(k_1)} \tag{24}$$

From Equations (18), (21) and (24), the stress in complex form is obtained

$$(\tau_{yz}+i\tau_{xz})=\tau_{yz}^{\infty}\frac{K(k)}{K(k_1)}\sqrt{\frac{\zeta^2-1}{\zeta^2-(1+\delta)^2}} \tag{25}$$

4. STRESS SINGULARITY FACTOR

The stress singularity at a rigid line tip is of great concern in fracture analysis. To this end, introduce polar coordinates r, θ with the origin at a rigid line tip as shown in Figure 1, and define the stress singularity factor S_{III}

$$S_{III}=\lim_{\substack{\theta=0\\ r\to 0}}\sqrt{2r}\,\tau_{yz}=\lim_{\substack{\theta=0\\ r\to 0}}\sqrt{2r}\,\tau_{yz}^{\infty}\frac{K(k)}{K(k_1)}\sqrt{\frac{\zeta^2-1}{\zeta^2-(1+\delta)^2}} \tag{26}$$

Equation (26) has a limit in closed form:

$$S_{III}=\frac{\tau_{yz}^{\infty}}{K(k_1)}\sqrt{\frac{\omega_1\,K(k)\mathrm{sn}\left[\frac{2a}{\omega_2}K(k'),k'\right]\mathrm{dn}\left[\frac{2a}{\omega_2}K(k'),k'\right]}{2\mathrm{cn}\left[\frac{2a}{\omega_2}K(k'),k'\right]}} \tag{27}$$

It is of interest to examine the influence of the column and stack spacings of the rigid lines on the stress singularity. Introduce the dimensionless stress singularity factor $\beta=S_{III}/(\tau_{yz}^{\infty}\sqrt{a})$, the dimensionless rigid line spacings $b_1=\omega_1/(2a)$ in the stack and $b_2=(\omega_2-2a)/(2a)$ in the column. Using solution (27), the numerical results of the dimensionless stress singularity factor β for several values of the dimensionless spacings b_1 and b_2 are listed in Table 1.

Table 1. β for several values of b_1 and b_2.

	β				
b_1	$b_2 = \infty$	$b_2 = 4$	$b_2 = 1$	$b_2 = 0.5$	$b_2 = 0.2$
∞	1.0000	1.0170	1.1284	1.2861	1.6885
5	0.9840	0.9849	1.0369	1.1357	1.3994
4	0.9750	0.9758	1.0163	1.1035	1.3424
3	0.9579	0.9580	0.9836	1.0535	1.2557
2	0.9138	0.9138	0.9231	0.9662	1.1135
1	0.7641	0.7641	0.7675	0.7726	0.8325
0.5	0.5631	0.5631	0.5631	0.5634	0.5752
0.2	0.3568	0.3568	0.3568	0.3568	0.3570

It is observed from Table 1 that decreasing the dimensionless spacing b_2 of the rigid lines in the column while holding the dimensionless spacing b_1 of the rigid lines in the stack fixed, tends to increase the dimensionless singularity factor β, but decreasing b_1 while holding b_2 fixed, tends to decrease β. The numerical results show the contrary dependence of β on the rigid line spacings in the stack and in the column.

It is also observed from Table 1 that β is greater unity when $b_1 = \infty$ and b_2 is finite, but that its value drops off with the decrease of b_1. Thus, for each value of b_2, there must be a value of b_1 for which β is equal to unity, the value of β for a single rigid line. Values of b_1 and b_2 for which β remains constants have been determined, and the resulting curves are plotted in Figure 3.

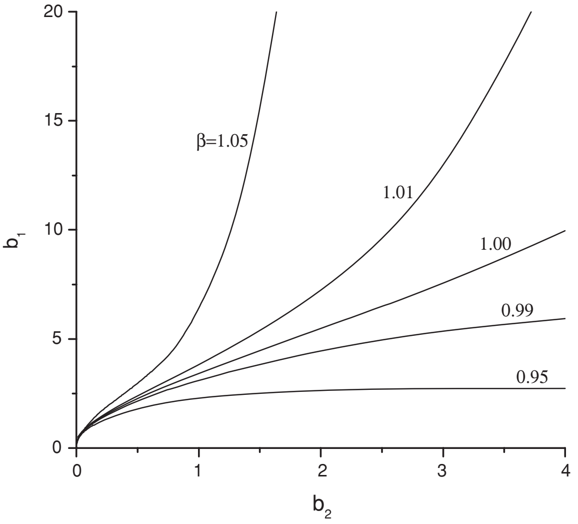


Figure 3. Relationship between b_1 and b_2 for constant values of β.

If we assume that fracture occurs when the stress singularity factor S_{III} reaches an experimentally determined critical value, it may be concluded that a body containing a rectangular array of rigid lines for which the point (b_1, b_2) lies on or below the curve $\beta = 1$ in Figure 3 will be of the same strength or stronger (i.e., will fracture at a higher level of applied stress) than the same body containing a single rigid line of the same length. Otherwise, the solid will be weaker than the same material with a single rigid line. It is seen from Table 1 and Figure 3 that increasing the density of rigid lines in the stack tends to strengthen the solid under antiplane shear, whereas increasing the density of rigid lines in the column tends to weaken it.

5. EFFECTIVE LONGITUDINAL SHEAR MODULUS

The present solution can also be used to determine the effective longitudinal shear modulus of the inhomogeneous materials containing a rectangular array of rigid line inclusions. Due to the presence of doubly periodic rigid lines, the material appears macroscopically continuous and orthotropic. The effective longitudinal shear moduli $G_{e,xz}$ and $G_{e,yz}$ of such an equivalent orthotropic material (refer to Figure 1) can be estimated by the following averaged stress and strain relations

$$\overline{\tau}_{xz} = G_{e,xz}\,\overline{\gamma}_{xz}, \qquad \overline{\tau}_{yz} = G_{e,yz}\,\overline{\gamma}_{yz} \tag{28}$$

where the overbar denotes averaging.

If only uniform τ_{xz}^{∞} (not be included in Figure 1) is imposed at infinity, the material will respond as if the rigid lines were not present, that is, $G_{e,xz}$ remains unchanged.

$$G_{e,xz} = G \tag{29}$$

Taking $1/b_1$ and $1/b_2$ as the dimensionless density of the rigid lines in the stack and in the column, respectively, the variations of the dimensionless effective longitudinal shear modulus $G_{e,yz}/G$ with $1/b_1$ for several values of $1/b_2$ are plotted in Figure 4.

It is observed from Figure 4 that the dimensionless effective longitudinal shear modulus $G_{e,yz}/G$ increases progressively with increasing the dimensionless densities $1/b_1$ or $1/b_2$. However, the influence of $1/b_2$ is much stronger than the influence of $1/b_1$. $G_{e,yz}/G$ tends to infinity as $1/b_2$ goes up to infinity for a fixed finite value of $1/b_1$, whereas it tends to a finite limit $\omega_2 \big/ (\omega_2 - 2a)$ as $1/b_1$ goes up to infinity for a fixed finite value of $1/b_2$.

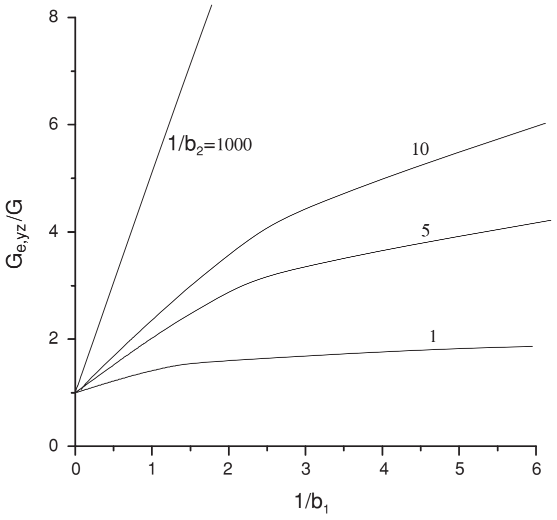


Figure 4. Variations of $G_{e,yz}/G$ with $1/b_1$ for different $1/b_2$.

6. CONCLUSIONS

By using the complex potential and the conformal mapping technique, the exact solution for doubly periodic rigid line inclusions under antiplane shear is obtained, and the stress singularity factor at each rigid line tip is given in closed form. Also, the effective longitudinal shear modulus is estimated.

The present solution shows that the rigid line densities in the column and in the stack are two important parameters for the material containing a doubly periodic array of rigid lines. Under antiplane shear, increasing of the density of rigid lines in the stack tends to strengthen the material, whereas increasing the density of rigid lines in the column tends to weaken it. A material containing doubly periodic rigid line inclusions in a rectangular array appears macroscopically orthotropic. The effective longitudinal shear modulus increases with increasing the rigid lines densities in the column or in the stack, and the sensitivity of the modulus to the density parameters in two directions is very different.

REFERENCES

1. Z.Y. Wang, H.T. Zhang and Y.T. Chou (1986), Stress singularity at the tip of a rigid line inhomogeneity under antiplane shear loading. *ASME Journal of Applied Mechanics*, 53, pp. 459–461.

2. R. Ballarini (1987), An integral equation approach for rigid line inhomogeneity problems. *International Journal of Fracture*, 33, pp. R23–R26.
3. Y.K. Cheung and Y.Z. Chen (1989), Multiple rigid line problems in an infinite plate. *Engineering Fracture Mechanics*, 34, pp. 379–391.
4. Q. Li and T.C.T. Ting (1989), Line inclusions in anisotropic elastic solids. *ASME Journal of Applied Mechanics*, 56, pp. 556–563.
5. T.H. Hao and Y.C. Wu (1989), Elastic plane problem of collinear periodical rigid lines. *Engineering Fracture Mechanics*, 33, pp. 979–981.
6. C.P. Jiang (1991), The plane problem of collinear rigid lines under arbitrary loads. *Engineering Fracture Mechanics*, 39, 2, pp. 299–308.
7. C.P. Jiang and C.T. Liu (1992), Stress distribution around a rigid line in dissimilar media. *Engineering Fracture Mechanics*, 42, 1, pp. 27–32.
8. C.P. Jiang and Y.K. Cheung (1995), Antiplane problems of collinear rigid line inclusions in dissimilar media. *Engineering Fracture Mechanics*, 52, 5, pp. 907–916.
9. A. Asundi and W. Deng (1995), Rigid inclusions on the interface between dissimilar anisotropic media. *Journal of Mechanics and Physics of Solids*, 43, 7, pp. 1045–1058.
10. X. Markenscoff and L. Ni (1996), The debounded interface anticrack. *ASME Journal of Applied Mechanics*, 63, pp. 621–627.
11. L.Z. Wu and S.Y. Du (2000), A rigid line in confocal elliptic piezoelectric inhomogeneity embedded in an infinite piezoelectric medium. *International Journal of Solids and Structures*, 37, 10, pp. 1453–1469.
12. T.C.T. Ting (1996), *Anisotropic Elasticity: Theory and Applications*. Oxford University Press, New York.
13. W.R. Delameter, G. Herrmann and D.M. Barnett (1975), Weakening of an elastic solid by a rectangular array of cracks. *ASME Journal of Applied Mechanics*, pp. 74–80.
14. N.I. Ioakimidis and P.S. Theocaris (1978), Doubly periodic array of cracks in an infinite isotropic medium. *Journal of Elasticity*, 8, 2, pp. 157–169.
15. S.S. Chang (1983), The general solutions of the doubly periods cracks. *Engineering Fracture Mechanics*, 18, 4, pp. 887–893.
16. T.H. Hao, J.D. Cheng and J.S. Yu (1985), The antiplane problem of double period nonuniform distribution crack field. *Applied Mathematics and Mechanics* (English Edition), 6, 2, pp. 191–196.
17. X. Li (2001), The effect of a homogeneous cylindrical inlay on cracks in the doubly-periodic complete plane strain problem. *International Journal of Fracture*, 109, pp. 403–411.
18. N.I. Muskhelishvili (1953), *Some Basic Problems of the Mathematical Theory of Elasticity*. Noorhoff Ltd., Groningen, Holland.
19. G.I. Barenblatt and G.P. Cherepanov (1961), On brittle cracks under longitudinal shear. *Applied Mathematics and Mechanics*, 25, pp. 1654–1666.
20. G.C. Sih (1965), Stress distribution near internal crack tip for longitudinal shear problems. *ASME Journal of Applied Mechanics*, 32, pp. 51–58.

PRELIMINARY ANALYSIS OF NORMAL STRENGTH CONCRETE WALLS WITH OPENINGS USING LAYERED FINITE ELEMENT METHOD

D.J. Lee, H. Guan, S. Fragomeni, J.H. Doh
School of Engineering, Gold Coast campus, Griffith University, PMB 50 Gold Coast Mail Centre, Queensland 9726, Australia

Abstract A nonlinear Layered Finite Element Method (LFEM) is applied to investigate the structural behaviour of reinforced concrete walls with openings. Four half-scale concrete walls with openings, tested recently in the laboratory, are analysed. The test variables are the wall slenderness ratio and the size and location of openings. The walls are axially loaded at an eccentricity of one-sixth of the wall thickness. The ultimate loads, the load–deflection responses and the crack patterns predicted by the LFEM are compared with the experimental results. The comparison shows that the LFEM is effective and accurate.

Keywords: layered finite element method, reinforced concrete walls, openings, vertical and eccentric loading, slenderness ratio, ultimate loads.

1. INTRODUCTION

This investigation focuses on the ultimate load behaviour of normal strength concrete walls with openings in one-way action, with normal concrete strengths and varying slenderness ratios. Previous research on walls with openings was carried out by only a few researchers [1, 2]. Saheb and Desayi [1] tested 12 wall panels and Doh and Fragomeni [2] tested only eight wall panels. Hence, the amount of experimental data is very limited. To achieve more reliable empirical formulae, more testing and numerical analysis are therefore required.

A nonlinear Layered Finite Element Method (LFEM) was developed by Guan and Loo [3, 4] for the analysis of punching shear behaviour of flat plate structures. The LFEM has been used successfully to analyse planar continuum

G. R. Liu et al. (eds.), Computational Methods, 1783–1791.

structures without openings, such as flat slabs and walls [5]. The method has been further extended to analyse slab–column connections with openings [6]. In this investigation, the LFEM is used to analyse normal strength concrete walls with openings.

Test results by Doh and Fragomeni [2] on normal strength half-scale reinforced concrete walls with openings will be used to investigate the effectiveness of the LFEM. The tested panels were supported top and bottom with varying opening size and location. The LFEM ultimate loads, load–deflection responses and crack patterns of selected panels are presented and compared to the experimental data.

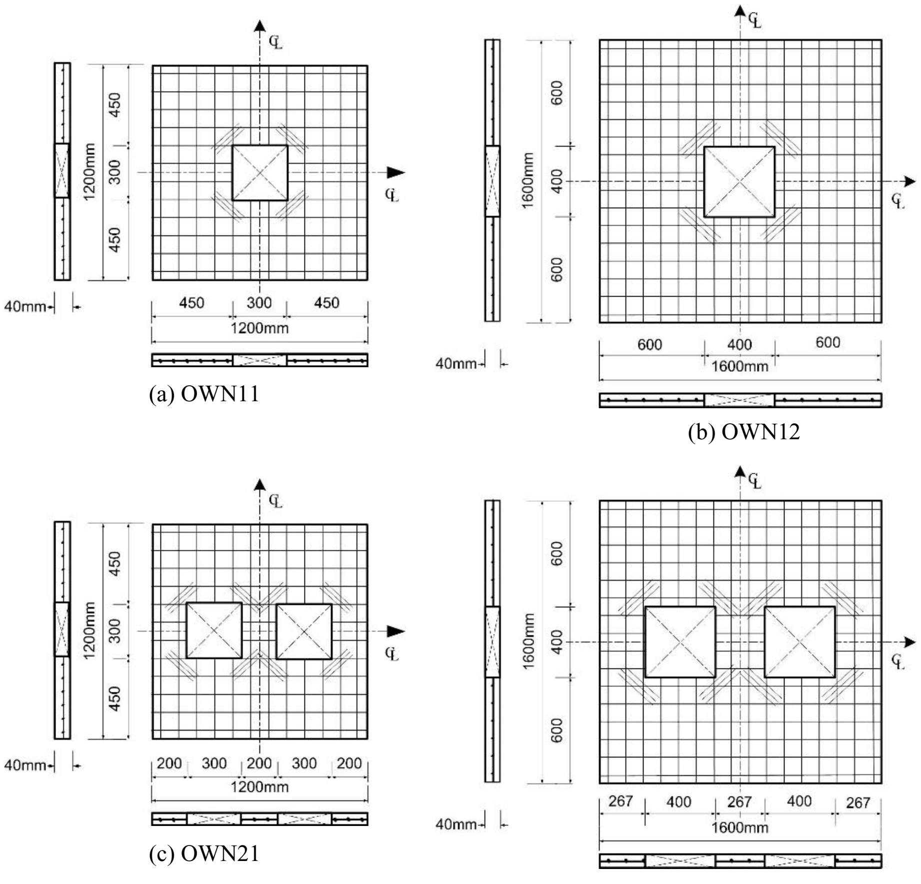
2. SELECTED EXPERIMENTAL PANELS

Geometrical properties and designations of the selected wall specimens to be analysed are given in Table 1. In each designation, the first letter **O** means one-way action; **W** means wall; **N** refers to normal strength concrete; the first number refers to the number of openings; the second, the type of specimens (e.g. OWN11). Three cylinders were tested for each wall panel to obtain the compressive strength. The average mean value of the cylinders was taken as the concrete compressive mean strength of the wall (f_{cm}) at the day of testing. The thickness and slenderness ratio of wall specimens as well as concrete compressive strength are also presented in Table 1.

Figure 1 shows the geometrical shapes, size and location of the opening(s) as well as the arrangement of the steel mesh for each wall panel. The opening size has been varied depending on the wall specimen. OWN11 and OWN12 each have one opening located at the wall centre. OWN21 and OWN22 each have two openings positioned at the centre and distributed evenly in the horizontal direction. The steel reinforcement (F41 mesh), which consists of 4 mm diameter bars at 100 mm vertical and horizontal spacing, is located centrally in a single layer, and satisfies the minimum requirement of reinforcement ratio in both

Table 1. Geometric and material properties of specimens.

Designation of specimen	OWN11	OWN21	OWN12	OWN22
Size of wall, height × length (mm × mm)	1200 × 1200	1200 × 1200	1600 × 1600	1600 × 1600
Size of opening (mm × mm)	300 × 300	300 × 300	400 × 400	400 × 400
Number of openings	1	2	1	2
Thickness, t_w (mm)	40	40	40	40
Slenderness ratio, H/t_w	30	30	40	40
Concrete mean strength, f_{cm} (MPa)	53	50	47	51



(a) OWN11

(b) OWN12

(c) OWN21

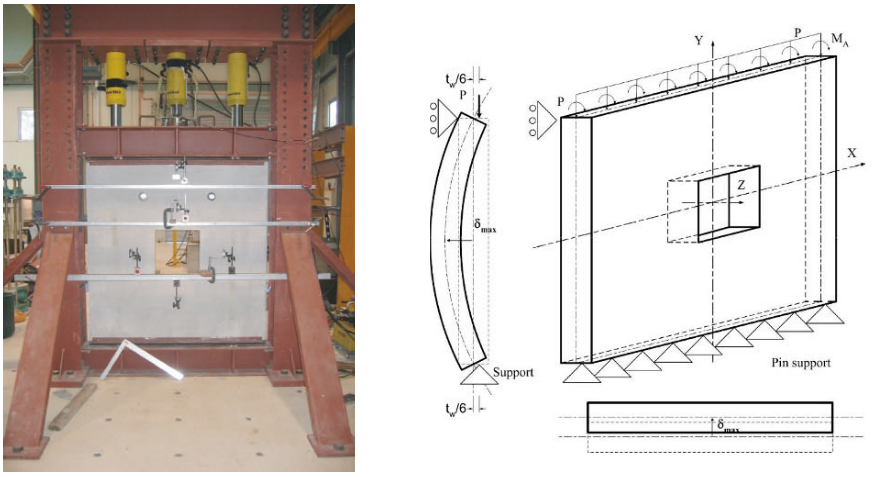
(d) OWN22

Figure 1. Dimension and arrangement of steel mesh in wall specimens.

directions as specified in AS 3600-01 [7]. The steel reinforcement mesh and three diagonal reinforcements at each corner of openings were arranged to prevent early shrinkage cracking of walls.

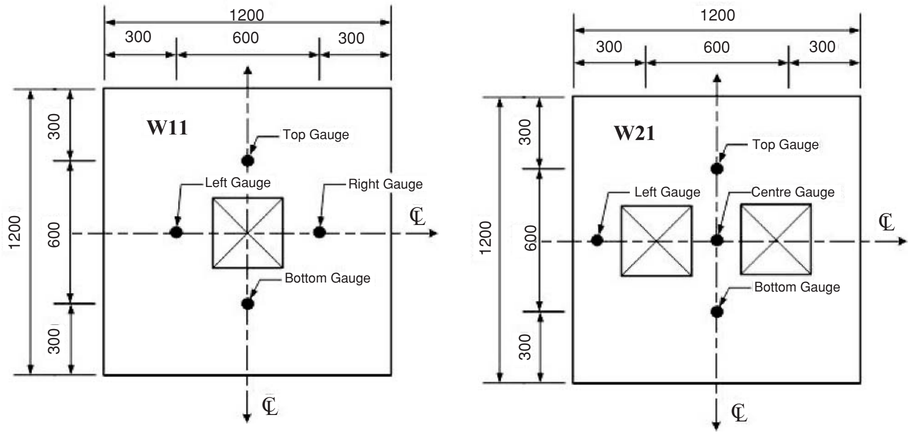
The test rig was previously designed by Doh [8] as shown in Figure 2a and the test set-up and test regime is described in more detail by Doh and Fragomeni [2]. Figure 2b shows a schematic of typical deflection expected from the one-way loading. To ensure one-way action the walls were loaded at an eccentricity of t/6. The loading was undertaken at an incremental rate of 1.5 tonnes, where deflection readings were also taken using dial gauges at the points indicated in Figures 2c and d.

The selected concrete walls are modelled in LFEM using the degenerated shell elements, each of which is subdivided into a number of layers fully bonded together. Eight concrete layers are adopted. Each layer has a uniform thickness



(a) Test rig set-up

(b) One-way action



(c) Wall panel with one opening

(d) Wall panel with two openings

Figure 2. Test set-up and location of dial gauges.

and its own material properties, which are assumed to be constant through the layer thickness. In this study, both front and back layers representing the compression and tension surfaces are made thinner as compared to the inner layers. The horizontal and vertical steel reinforcement meshes are represented by two smeared layers of equivalent thickness as shown in Figure 3. Each layer has gauss points on its mid-surface. The stresses of each layer are computed at the gauss points and are assumed to be constant over the layer thickness. This assumption results in a stepwise approximation of the stress distribution over the thickness. The material responses at the gauss points can be elastic, plastic or fracture according to the loading history.

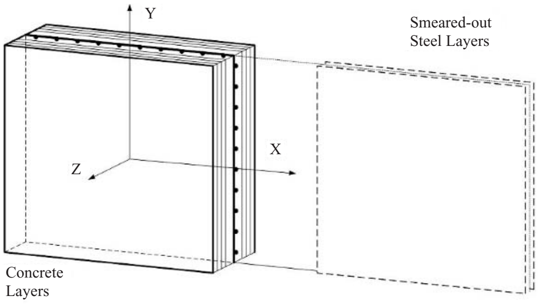


Figure 3. Construction of typical layered system.

In the LFEM, the incremental and iterative procedure, based on the Newton–Raphson method, is used to compute the nonlinear solution. Both the geometric and material nonlinearities are taken into account in the analysis. The LFEM is also able to offer post-processing functions including the graphical representation of the deformed shape and crack patterns of the wall panels.

Due to symmetry, only half of each wall needs to be analysed using the LFEM. A convergence study was initially carried out for wall specimen OWN12 to determine the most appropriate finite element mesh. In total, five mesh schemes were attempted. They are referred to as MS1 to MS5, and are shown in Figure 4.

MS1 is a very coarse mesh having uniform element size. This mesh is further refined to achieve an accurate analysis from MS2 to MS5. The size of elements around the opening is made relatively smaller to deal with stress concentration around the corner of the openings. The corresponding nodal and element numbers of the five mesh schemes are given in Table 2. The ultimate load N_u and the maximum deflection δ_{max} for each mesh scheme are also presented.

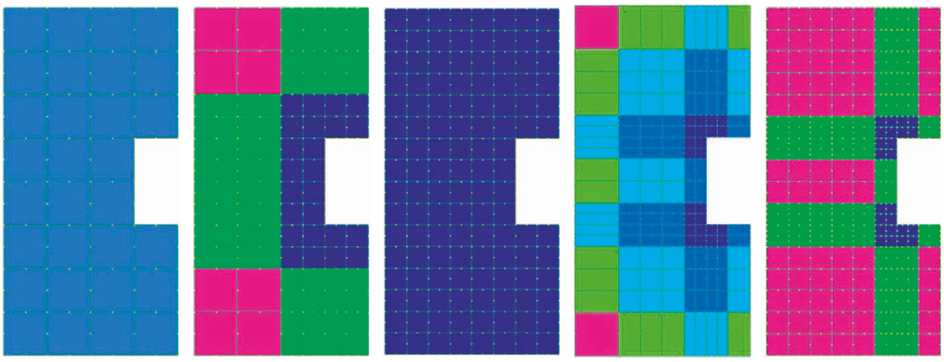

(a) MS1 (b) MS2 (c) MS3 (d) MS4 (e) MS5

Figure 4. Mesh schemes for OWN12.

Table 2. Convergence results of OWN12

Mesh scheme	MS1	MS2	MS3	MS4	MS5
Number of elements	30	64	120	144	182
Number of nodes	117	233	413	493	613
Ultimate load, $N_{u,LFEM}$ (kN)	281.1	278.1	278.1	278.1	278.1

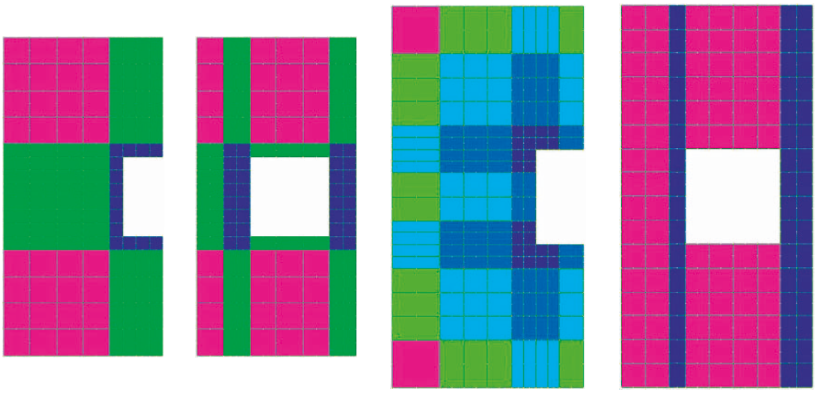

(a) OWN11 (b) OWN21 (c) OWN12 (d) OWN22

Figure 5. Mesh scheme of specimens.

The ultimate loads N_u, derived by the LFEM, have apparently converged to 278.1 kN. MS4 is selected as the appropriate mesh for OWN12 taking into account the finer mesh around the corner of the opening. This can better simulate the stress distribution around the range. In addition, MS4 is more economical than MS5 in terms of lesser elements and nodal numbers. Based on this convergence study, the mesh scheme for the remaining wall specimens can be readily determined. The mesh schemes for all the specimens are shown in Figure 5.

3. ANALYSIS AND COMPARISON

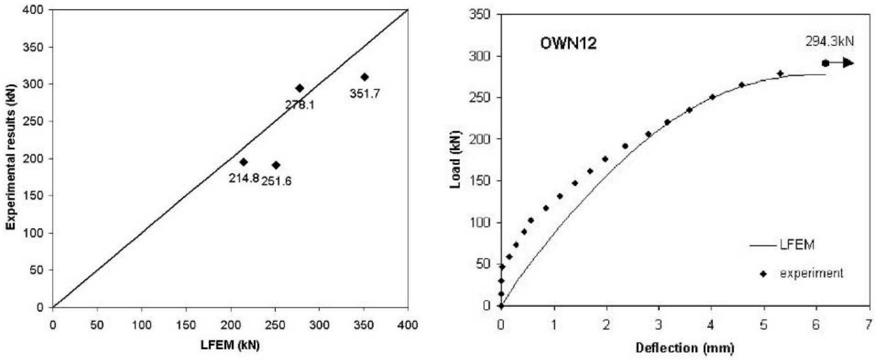
The experimental and LFEM ultimate loads and maximum deflections are presented in Table 3. The ratio between the experimental and the LFEM predictions are also included.

The comparison shows that the LFEM gives reasonable and satisfactory prediction, where the average of the load and deflection ratios are close to unity. The comparison between the experimental and the LFEM ultimate loads is presented in Figure 6a and the load–deflection response of OWN12 is shown in Figure 6b. These results are further evidence of good prediction by the LFEM.

The deflection predicted by the LFEM as presented in Figure 7 for wall panels, OWN12 and OWN22, demonstrates typical one-way action. It is evident

Table 3. Experimental and LFEM results

Designation of specimen	OWN11	OWN21	OWN12	OWN22	Average
$N_{u,exp}$ (kN)	309.0	191.3	**294.3**	195.7	
$N_{u,LFEM}$ (kN)	**351.7**	**251.6**	**278.1**	**214.8**	
Ratio, $N_{u,exp}/N_{u,LFEM}$	0.879	0.760	1.058	0.911	0.902
$\delta_{max,exp}$ (mm)	5.7	4.8	5.3	5.0	
$\delta_{max,LFEM}$ (mm)	5.5	4.9	6.2	5.8	
Ratio, $\delta_{max,exp}/\delta_{max,LFEM}$	1.027	0.976	0.857	0.861	0.930

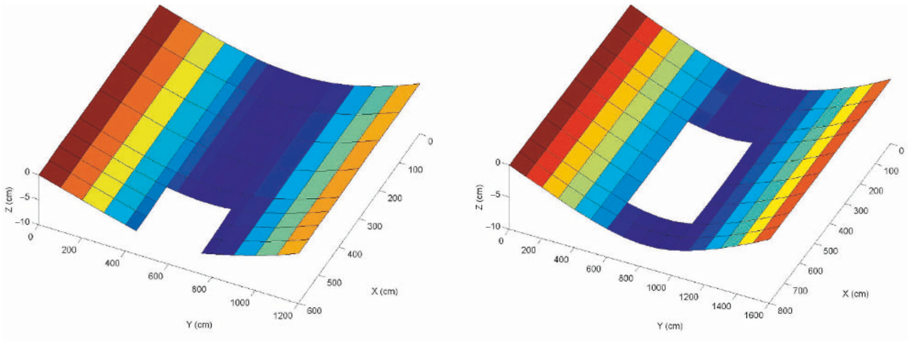


(a) Ultimate loads of each specimens (b) Load–deflection response

Figure 6. Comparison between experimental and LFEM results.

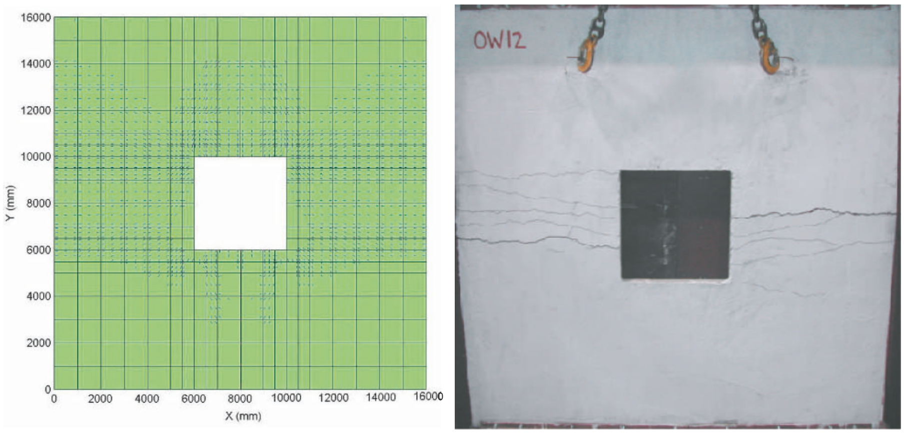
that the eccentrically loaded walls, simply supported at the top and bottom, deflect in a single curvature. This is identical to the experimental observation.

The crack patterns predicted by the LFEM closely resemble the test results, as shown in Figure 8 for OWN12 and OWN22. The major horizontal cracks

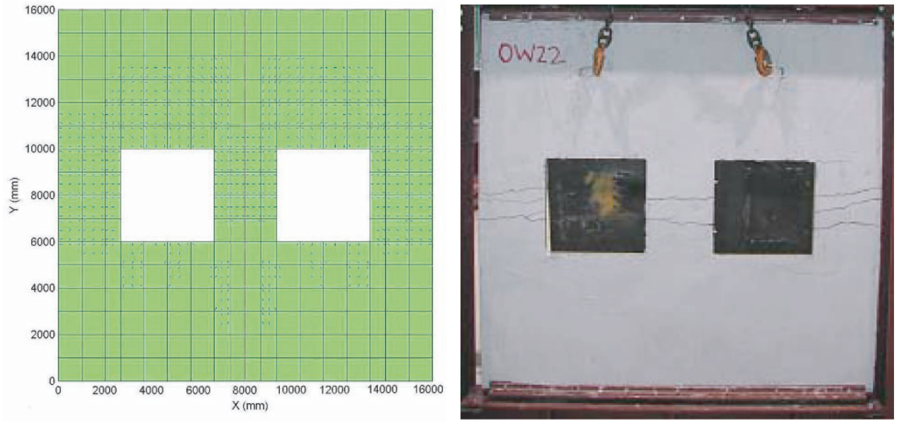


(a) OWN12 (b) OWN22

Figure 7. Typical deformed shapes predicted by the LFEM.



(a) LFEM prediction-OWN12 (b) Experimental result-OWN12



(c) LFEM prediction-OWN22 (d) Experimental result-OWN22

Figure 8. Typical crack patterns of OWN12 and OWN22.

through the openings are accurately predicted by the LFEM. But there seem to be more cracks above and below the openings in the LFEM predictions, whereas actual cracks were more distinct. This is generally expected by any finite element analysis [9].

4. CONCLUSIONS

Based on the analysis of four one-way wall panels with openings, the following conclusions are made,

1. The convergence study has enabled a selection of the most appropriate meshes for the analysis of each wall panel using the LFEM.
2. The mean ratio of the ultimate load between the experimental and the LFEM results is 0.902 and that of the maximum deflection is 0.930, which are considered to be satisfactory.
3. The crack patterns predicted by the LFEM closely resemble the test results.
4. The LFEM is effective and reliable in analysing reinforced concrete walls with either one or two openings.

Based on the satisfactory outcomes, further analysis work will be conducted to cover walls with openings in not only one-way action but also two-way action (i.e. walls supported on all sides). Further, walls with openings made of high strength concrete will be investigated using the LFEM.

REFERENCES

1. S.M. Saheb and P. Desayi (1990), Ultimate strength of R.C. Wall Panels with openings. *Journal of Structural Engineering*, 116, 6, pp. 1565–1578.
2. J.H. Doh and S. Fragomeni (2004), Evaluation and experimental work for concrete walls with openings in one and two way action. In: *The third Civil Engineering Conference in the Asian Region*, Korea.
3. H. Guan and Y.C. Loo (1997), Flexural and shear failure analysis of reinforced concrete slabs and flat plates, *Advances in Structural Engineering*, 1, 1, pp. 71–85.
4. H. Guan and Y.C. Loo (1997b), Layered finite element method in cracking and failure analysis of beams and beam–column–slab connections, *Structural Engineering and Mechanics*, 5, 5, pp. 645–662.
5. H. Guan and Y.C. Loo (2002), Deflection, cracking and failure analysis of planar continuum reinforced concrete structures. In: *The Second International Conference on Advances in Structural Engineering and Mechanics (ASEM'02)*, C.K. Choi and W.C. Schnobrich (eds.), Pusan, Korea, August 21–23, CD-ROM Proceedings, pp. 1–8.
6. H. Guan, N.A. Kaoustos and Y.C. Loo (2004), Comparative study on punching shear behaviour of slab–edge column connections with openings. In: *Proceedings of the 18th Australasian Conference on the Mechanics of Structures and Materials (ACMSM18)*, Perth, Australia, December 2004.
7. Standards Australia (2001), *Australian Standard Concrete structures*. AS3600-2001, Section 11, pp. 120–123.
8. J.H. Doh (2002), Experimental and theoretical studies of normal and high strength concrete wall panels. PhD thesis, Griffith University.
9. Y.C. Loo and H. Guan (1997), Cracking and punching shear failure analysis of RC flat plates. *Journal of structural engineering*, 123, 10, pp. 1321–1330.

FLEXURAL BEHAVIOUR OF CONCRETE BEAMS REINFORCED WITH INTERNAL TENSILE STEEL AND EXTERNAL CFRP

W. Karunasena and W.C. Anderson
Discipline of Civil & Environmental Engineering, School of Engineering, James Cook University, Townsville, Queensland 4810, Australia

Abstract A simple analytical model that can easily be implemented in MATLAB (Matrix Laboratory) is presented for accurately predicting the moment-curvature and load-deflection relations of concrete beams reinforced with internal tensile steel bars and rehabilitated with externally bonded carbon fibre reinforced polymer (CFRP) strips. The model has been verified by comparing model results with those obtained experimentally. The model results and test data show that, with adequate bond strength, application of a CFRP strip to a reinforced concrete beam significantly increases the capacity. Currently, the model does not consider debonding failure of CFRP from the beam.

Keywords: concrete beam, fibre reinforced polymer, rehabilitation, computer model.

1. INTRODUCTION

Reinforced concrete structures undergo continuous deterioration throughout their service life, and many are required to support loads in excess of their original design loads. Such structures need to be replaced or rehabilitated, with rehabilitation being the cheapest alternative. Rehabilitation of concrete beams with CFRP is becoming an increasingly popular method. Desirable characteristics of CFRP include: high strength-to-weight ratio, corrosion resistance, high impact resistance and low maintenance. Work by many researchers [1, 2] indicate that, with adequate bond strength and anchorage, application of a CFRP strip to the tension face of a beam significantly increases capacity. Recently, Duthinh and Starnes [3] presented an iterative numerical approach to predict the flexural behaviour of CFRP rehabilitated concrete beams. Their approach

G. R. Liu et al. (eds.), Computational Methods, 1793–1797.

involved iteration over two variables. In this chapter, an approach involving only one variable is presented.

2. SIMPLIFIED COMPUTATIONAL MODEL

Figure 1 shows the four point bending arrangement used for testing of the CFRP reinforced concrete beams discussed in this chapter. The length variables used in the formulation of the simplified computational model to predict flexural behaviour of the beam are defined in Figure 1.

The stress–strain behaviour for steel is assumed to be linear elastic until steel yields and then perfectly plastic with a constant yield stress of f_{sy} in the post-yield region. The constitutive behaviour of concrete is represented by the Hognestad stress parabola. The model assumes beam failure by compressive concrete crushing at ultimate concrete strain, ε_{cu}, of 0.003. The problem is formulated to solve iteratively over the top fibre concrete strain, ε_{cm}. For the iteration, we choose strain increment $\Delta\varepsilon$ as ε_{cm}/N, where N is the total number of steps. A value of 20 for N has been used in the present work. The current value of the extreme compressive fibre concrete strain can be defined by

$$\varepsilon_{cm} = (n-1)\Delta\varepsilon \tag{1}$$

where n is the current step number.

Using linear elastic stress–strain relations, the steel force, F_s, for the current step can be expressed in terms of the neutral axis depth, d_n as

$$F_s = A_s E_s \varepsilon_{cm}\left(\frac{d-d_n}{d_n}\right) \tag{2}$$

where A_s and E_s, respectively, denote the area and Young's modulus of steel, and d represents the depth to the tensile steel. Similarly, the FRP strip force,

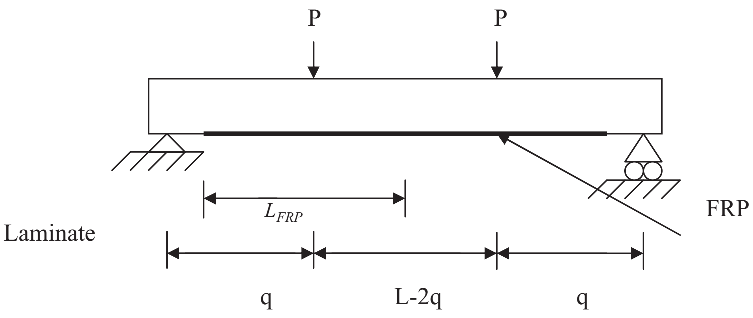


Figure 1. Four-point bending arrangement.

F_{frp}, can be expressed as

$$F_{frp} = A_{frp}\,E_{frp}\,\varepsilon_{cm}\left(\frac{h - d_n}{d_n}\right) \tag{3}$$

where A_{frp} and E_{frp} are area and Young's modulus, respectively, of FRP.

The concrete force, F_c, can be expressed by the integral of Hognestad stress parabola as

$$F_c = bf'_c d_n\left(\frac{\varepsilon_{cm}}{\varepsilon_0} - \frac{\varepsilon_{cm}^2}{3\varepsilon_0^2}\right) \tag{4}$$

where ε_o is the strain at which the concrete attains its compressive strength, f'_c. The neutral axis depth, d_n, can now be determined by solving the following equilibrium equation:

$$F_c = F_s + F_{frp}. \tag{5}$$

Knowing d_n, one can determine the FRP and steel strains, ε_{frp} and ε_s, respectively. If FRP strain exceeds ultimate tensile strain, ε_{frpu}, F_{frp} is set to zero. Similarly, if steel strain exceeds yield strain, ε_{sy}, F_s is set to $A_s f_{sy}$, where f_{sy} is the yield stress of steel. Now equation (5) is solved again for the new value of d_n.

The moment of resistance, M, and the mid-span curvature, ϕ_o can be determined from

$$M = F_s(d - c) + F_{frp}(h - c) \tag{6a}$$

$$\phi_o = \varepsilon_{cm}/d_n. \tag{6b}$$

where c is the depth to the centroid of the Hognestad stress parabola. Next the step number, n, is increased by 1, and new values for M and ϕ_o are calculated. The procedure is repeated until n reaches the value N. The result of this will be a series of moment-curvature values. These moment-curvature values are used to determine the beam curvature distribution along the length. In the final phase of the model, the central deflection, δ of the beam is determined by invoking the virtual work method.

3. NUMERICAL RESULTS AND DISCUSSION

The simple computational model presented in the previous section was implemented in MATLAB software package. A comparison of the moment-curvature and load-deflection results from the simple model with those obtained experimentally from two 3 m long (L = 3 m) beams (beam I and beam II) at James Cook University are presented in Figure 2 and 3.

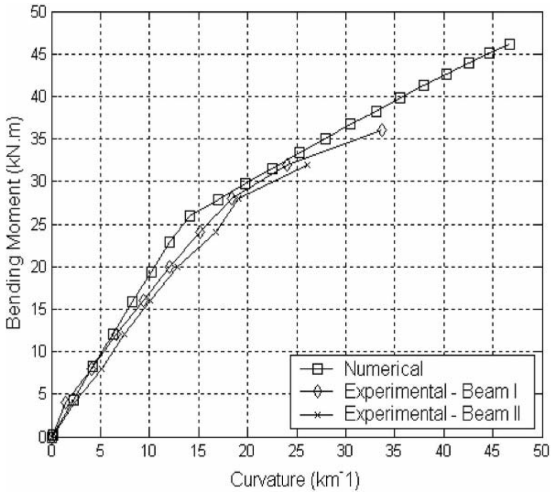


Figure 2. Moment—curvature relation for beam I and beam II.

The experimental beams had cross-section dimensions of 150 mm width × 250 mm height. The internal reinforcing steel consisted of 2Y12 reinforcing bars (with $A_s = 226\,\text{mm}^2$ and $f_{sy} = 400$ MPa) 25 mm from the extreme tensile fibres. No shear reinforcement was provided. Both beams were cured for 28 days after construction. Before applying CFRP, the beams were loaded in four point bending (with q = L/3) until steel starts yielding and then unloaded. Approximately 2 months later, the beams were rehabilitated by bonding a 80 mm wide,

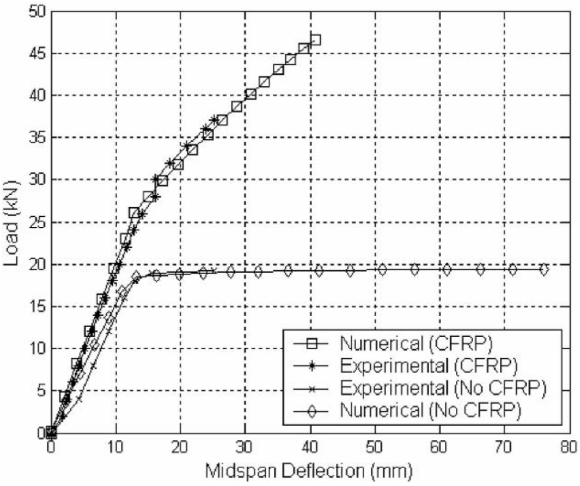


Figure 3. Load—deflection relation for beam II.

1.2 mm thick, unidirectional CFRP laminate (Sika Carbodur 80) to the tension face of each beam. The rehabilitated beams were tested in four point bending 7 days later and experimental data required for moment-curvature and load-deflection curves were recorded. For numerical computations, E_{frp} was taken as the manufacturer's recommended value of 150 GPa and E_s as 200 GPa. The ultimate tensile strength of CFRP was taken as 2400 MPa. A $f_c^{'}$ value of 31 MPa based on concrete cylinder testing was used.

Figures 2 and 3 illustrate strong congruence of moment-curvature and load-deflection relations between experimental and model prediction results. This indicates that the simple algorithm can be used to predict the flexural behaviour of CFRP rehabilitated concrete beams. Currently, the model assumes a perfect CFRP/concrete bond, prior to beam failure by CFRP tensile rupture or compressive concrete crushing. In contrast to this, the experimental results show that ultimate beam failure has occurred by debonding near the CFRP/concrete interface, before either of the aforementioned failure modes. Therefore, the model has not predicted ultimate moment capacity accurately. Authors are currently further investigating this failure mode.

4. CONCLUSION

A simple computational model that can predict the moment-curvature and load-deflection relations for CFRP rehabilitated beams with internal tensile steel reinforcement has been presented. The model results have been verified by comparing with experimental results. The model takes into account CFRP tension failure and compressive concrete crushing failure modes. However, currently, the debonding near the CFRP/concrete interface is not taken into account and, therefore, the ultimate capacity predictions from model are not accurate.

REFERENCES

1. M. Arduini, A.D. Tommaso and A. Nanni (1997), Brittle failure in FRP plate and sheet bonded beams. *Structural Journal*, ACI, 94, 4, pp. 363–370.
2. El-Mihilmy, M.T. and Tedesco, J.W. (2000), Deflection of reinforced concrete beams strengthened with fibre-reinforced polymer (FRP) plates. *Structural Journal*, ACI, 97, 5, pp. 679–688.
3. D. Duthinh and M. Starnes (2004), Strength and ductility prediction of concrete beams reinforced with carbon fiber-reinforced polymers plates and steel. *Journal of Composites for Construction*, ASCE, 8, 1, pp. 59–69.

3D MODELLING OF BRITTLE FRACTURE IN HETEROGENEOUS ROCKS

Z.Z. Liang[1,2], C.A. Tang[1] and Y.B. Zhang[1,2]
[1] *School of Civil and Hydraulic Engineering, Dalian University of Technology, Dalian, 116024, Chaina*
[2] *Research Center for Numerical Tests on Material Failure, Dalian University, Dalian 116622, China*

Abstract Owing to the vast developments in computer science and technologies, recent years have seen a renewal of interest in the computational modelling of material failure in meso-macro scale. The multi-scale capability of the method is recognized as a promising tool in attacking formidable problems in fracture mechanics for heterogeneous media, such as rock, concrete or ceramic, and, indeed, has been successfully applied to the various engineering problems for industrial materials. These models have established themselves as a powerful and realistic alternative to the non-local continuum models for softening damage and fracturing. However, the quantitative macro response (such as stress–strain) from most of the current meso-mechanical models is not close enough to the behaviour of real materials. One reason for this shortcoming is that most of these models are two-dimensional. In this paper, a three-dimensional material failure process analysis model, MFPA^{3D}, is proposed. The failure behaviour of brittle materials can be simulated by this three-dimensional model. The model can realistically simulate the inelastic triaxial behaviour, strength limits, and post-peak response for both tension and compression. The capabilities of the MFPA^{3D} model to generate a wide range of damage morphologies are examined in this paper.

Keywords: failure process, numerical simulation, elastic damage.

1. INTRODUCTION

There is greatly increasing interest and activity in the modelling failure mechanics in heterogeneous materials, such as rocks, concretes, ceramics etc., as the vast developments in computer science and technologies in the past

G. R. Liu et al. (eds.), Computational Methods, 1799–1803.

decades. These materials consist of crystals, grains, cementitious materials, voids, pores, cracks and the like. These heterogeneities give rise to local differential strain and stress within stressed brittle materials. Heterogeneities play an important role during fracture process to create new cracks or make existing cracks propagate when subjected to various loading cases. A lot of numerical methods were applied to the study of the microscopic processes involved in brittle fracture of rock-like materials. However, due to the limited computing performance and complicated data processing, very few numerical codes can simulate brittle failure process satisfactorily. The numbers of the total elements in most of the numerical simulations were not more than 20,000 [1, 2]. It is well know that a great number of elements are needed to simulate the failure process in heterogeneous materials in order to take various weakness or initial cracks into account. The failure process of rocks is three dimensional. Even though many models can simulate the failure process, including crack initiation, propagation and coalescence, they are developed just based on two dimension [3]. Most of the cracks in geologic body are three dimensional. Cracks propagate and coalesce in three-dimensional space. Many rock mechanical problems except plain strain problems and plain stress problems cannot be simplified to two-dimension problems. Apparently, the ability to model the brittle failure of heterogeneous materials lies in the high speed stress solver in computer calculating and data process to deal with great number of elements and suitable technique to demonstrate three-dimensional failure process properly. In this respect, we present a newly developed numerical code RFPA^{3D} (Material Failure Process Analysis-3D) to study the brittle failure process of heterogeneous materials in this paper.

2. BRIEF DESCRIPTION OF MFPA^{3D}

2.1 Heterogeneity in MFPA^{3D}

The MFPA^{3D} code has been developed by considering the deformation of an elastic material containing an initial random distribution of micro-features to simulate the progressive failure in a more visual way, including simulation of the failure process, failure induced further rupture and failure induced stress redistribution. Mechanical properties in each element, such as elastic modulus, peak strength, passion ratio and weight, are the same. For heterogeneity, the material properties for elements are randomly distributed throughout the specimen by following a Weibull distribution. We define m as the homogeneity index of the rock. A larger m implies that more elements with the mechanical properties approximated to the mean value and a more homogeneous rock specimen (Figure 1).

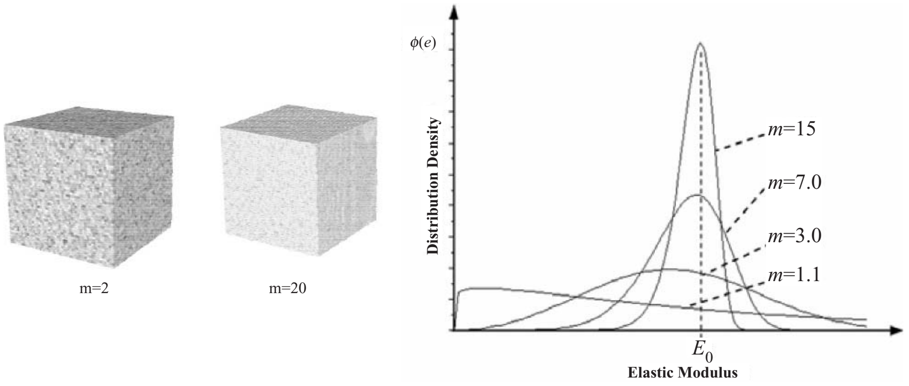


Figure 1. MFPA^{3D} models with different homogeneity indices.

2.2 Elastic damage evolution in MFPA^{3D}

Each element in the numerical model has an elastic-brittle constitutive law during the loading process. When the stress increases to a value leading to the failure of the element, the elastic modulus of the element may degrade gradually as damage progresses. In the presented paper, it was assumed that each element has a Mohr–Coulomb failure criterion envelope with a tensile cut-off on mesoscopic scale, and the failure of elements may be either in shear or in tension mode. If the stress condition of the element meets the shear failure criterion, damage variable Dcan be described by the following function:

$$D = \begin{cases} 0 & \varepsilon_1 < \varepsilon_{c0} \\ 1 - \dfrac{\sigma_{rc}}{E_0 \varepsilon_1} & \varepsilon_1 \geq \varepsilon_{c0} \end{cases} \tag{1}$$

Where σ_{rt} is the residual compressive strength and ε is the effective strain. E_0 is the initial elastic modulus. If the stress condition of the element meets the tensile failure criterion, the evolution of damage variable D can be summarized as following:

$$D = \begin{cases} 0 & (\bar{\varepsilon} < \varepsilon_{t0}) \\ 1 - \dfrac{\sigma_{rt}}{\bar{\varepsilon} E_0} & (\varepsilon_{t0} \leq \bar{\varepsilon} \leq \varepsilon_{ut}) \\ 1 & (\bar{\varepsilon} > \varepsilon_{ut}) \end{cases} \tag{2}$$

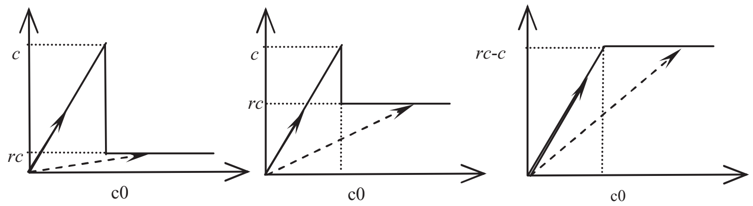
where σ_{rt} is the residual tensile strength of the element and ε is the effective strain. ε_{t0} and ε_{ut} are strain at the failure point and ultimate effective strain.

2.3 Residual strength

A parameter termed residual coefficient β is introduced to define the post-strength in the presented elastobrittle model, and it can be described as the following function: $\sigma_{rc}=\beta\sigma_c$. The peak strength of each element keeps constant before the stress of the element satisfies the strength criterion. Residual coefficient describes the variation of the degradation behaviour of the specimens. The stress–strain curves are controlled by residual coefficient and display different strain-softening features. Numerical specimens with higher residual strength coefficient show more ductile features after the peak strength, while low residual strength will lead to brittle failure. The slope of the dashed lines are elastic modulus after failure.

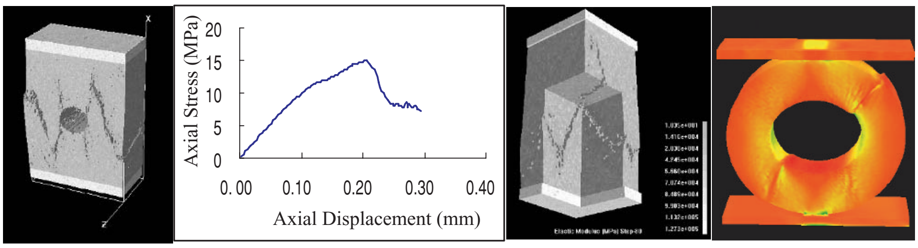
3. NUMERICAL SIMULATION RESULTS

In this section, three numerical tests simulated by MFPA^{3D} were undertaken to analyse the complete failure process of brittle materials. All numerical specimens were divided into 128,000 elements. A high speed FEM solver was adopted in MFPA^{3D} code. Not only the crack initiation, propagation and coalescence but also the complete progressive failure process and complete stress–strain curves can be obtained. Figure 2a shows the failure mode of a rock sample containing a circular opening subjected to uniaxial compression. Stress concentration near the circular opening was released gradually when cracks around the opening initiated and propagated along the axial stress orientation. Figure 2b is the axial stress-displacement curve. It experienced four stages: linear deformation, deformation accelerating caused by crack propagation near the opening, collapse caused by coalescence of the cracks and residual deformation. Figure 2c shows the brittle failure of a numerical specimen subjected to axial compressive loading. We can find only two main shear cracks which leaded to the brittle fracturing from the surface of the specimen. From



$\beta=0.1$ $\beta=0.5$ $\beta=1.0$

Figure 2. Stress–strain curves with different residual coefficients.



(*a*) (*b*) (*c*) (*d*)

Figure 3. Numerical results simulated by MFPA^{3D}.

three-dimensional numerical pictures, we can find that two shear cracks on the surface were caused by a tensile crack inside the specimen. Figure 2d shows the maximum principal stress field of a ring subjected to compression imposed by a relative motion of the upper and lower loading plates. Compressive stress was found in the parts near the plates. Two compressive stress zones formed which prevented the propagation of cracks from propagating along the compressive stress orientation.

The presented model can simulate non-linear deformability of a quasi-brittle behaviour with an ideal brittle constitutive law with a strength and elastic modulus reduction of weaken element by introducing micro-heterogeneities into the model. Although the constitutive law for the individual element in the numerical model is nearly brittle, a substantial non-linearity exists before the ring reached its peak strength. Even though each element obeys a simple constitutive law at mesoscopic level, the quantitative macro response demonstrate strong non-linearity during the complete progressive failure process.

ACKNOWLEDGEMENTS

This work is funded by NSFC (No. 50134040, 50134040, 50374020) and HKSARG (HKU7029/02E).

REFERENCES

1. A. Arlsan, R. Ince and B.L. Karihaloo (2002), Improved lattice model for concrete fracture. Journal of Engineering Mechanics, 128, 1, pp. 57–65.
2. J.G.M. Van Mier, M.R.A. van Vliet and T.K. Wang (2002), Fracture mechanisms in particle composites: Statistical aspects in lattice type analysis. Mechanics of Materials, 34, 11, pp. 705–724.
3. C.A. Tang (1997), Numerical simulation on progressive failure loading to collapse and associated seismicity. International Journal of Rock Mechanics and Minerals Science, 34, pp. 249–262.

STRENGTH DETERIORATION OF NONPRISMATIC REINFORCED CONCRETE BEAMS

M. Teguh[1,2], F. Saleh[3], C.F. Duffield[1] and G.L. Hutchinson[1]
[1]*Department of Civil and Environmental Engineering, the University of Melbourne, Parkville Campus VIC 3010, Australia*
[2]*Department of Civil Engineering, the Islamic University of Indonesia, Yogyakarta 55223, Indonesia*
[3]*Department of Civil Engineering, the Muhammadiyah University of Yogyakarta, Indonesia*

Abstract Cracking of reinforced concrete structures reduces stiffness and a subsequent deterioration in the capacity of structural elements. This study has investigated the use of a central difference approximation to obtain curvature mode shapes for structural elements based on the measured natural frequency and modal displacements for each load step. The findings of numerical analysis, conducted using finite element software, has been compared with experimental results. This paper also presents a proposed method of damage detection for structural elements based on an estimate of stiffness degradation using nonprismatic beam element called the Modified Flexure Damage Index (MFDI). This approach estimates damage in terms of the change in the fundamental frequency of the structural elements. It has been demonstrated that the MFDI and static strength approach provide a realistic estimate of crack damage on nonprismatic reinforced concrete beams.

Keywords: nonprismatic, strength deterioration, curvature mode, reinforced concrete.

1. INTRODUCTION

Nonprismatic beams; e.g., to resist shear deformations caused by induced loads are commonly used in long highway bridge structures. They are less common in building structures [1]. Analytical procedures to predict the capacity of

G. R. Liu et al. (eds.), Computational Methods, 1805–1815.

nonprismatic members are complex due to both the modulus of elasticity E_x and moment of inertia I_x which vary along length of the member. In general, design criteria for the prismatic members, for examples beam and column elements, in a ductile earthquake resistant structures, are well documented and codified, e.g., the American Concrete Institute (ACI), Australian/New Zealand (AS/NZ) and Japanese Codes. Similar, Standards are not available for nonprismatic members and thus such members require special consideration in structural analysis and design. Fertis [1] proposed the concept and method of the equivalent systems for the analysis of prismatic and nonprismatic members. This method provided an accurate and efficient method for such problems, and it simplified a great deal the computational work that was required in inelastic solution subjected to static loads. In addition, this method dealt with the derivation and used of equivalent systems of constant stiffness rigidity EI, where both E_x and I_x of the original member varied in any arbitrary manner, and where the material of the member along its length was permitted to be stressed well beyond its elastic limit. However, this research has not adequately addressed the strength deterioration of nonprismatic reinforced concrete beams.

Stress concentrations in particular regions facilitate propagation of cracks or localized damage and these cause strength deterioration of the members. It has been observed that once cracks occur in a beam the stiffness is reduced and the damping in the beam increases. The natural frequency decreases and the vibration mode changes are affected by stiffness reductions along the beam. Most of the previous investigations have placed emphasis on using the decrease in frequency or the increase in damping to detect cracks. Very little work regarding the strength deterioration in nonprismatic reinforced concrete beams has been done by using the changes in the mode shapes to predict the crack. The changes of the mode shapes to detect damaged structures was considered by Yuen [2] through the use of a finite element method to analyse a cantilever reinforced concrete beam to obtain the natural frequencies and the mode shapes of a damaged beam. The frequency changes can be used to determine the presence of crack or damage in a structure. However, the determination of crack location is not directly correlated to the frequency changes. This is because cracks at two different locations associated with certain crack lengths may cause the same amount of frequency change. Other parameters need to be determined which will directly identify the crack location in a structure [3].

In this paper, curvature mode shapes are introduced to detect the strength deterioration in nonprismatic reinforced concrete beams. The different in the curvature mode shapes between the intact and damaged beams is utilized to detect crack locations. The changes in the curvature mode shapes are shown to be localized in the region of damage compared to the changes in the displacement mode shapes. Simply supported nonprismatic beams have been investigated in the experimental tests and were used in the numerical analysis using finite element method (FEM).

2. THEORETICAL REVIEW

2.1 Structural system of damaged beam

A damaged beam can be modelled by reducing its flexural rigidity (*EI*) only if plastic hinges have not been found. A damaged beam also has changed dynamic characteristics, i.e., the natural frequency reduces and thus is the potential for changes in the mode shapes. Curvature (κ) mode shapes are related to the flexural stiffness of cross-sections [3] and is defined by

$$\kappa = \frac{M}{EI} \tag{1}$$

where M is the bending moment at a section, E is the modulus of elasticity and I is the second moment of the cross-sectional area.

When a crack occurs it reduces the flexural rigidity of the beam (particularly at damaged region) and an increase in the magnitude of curvature of the beam results. The curvature changes are local in nature and hence can be used to detect and locate a crack in the beam. The curvature change increases with reduction in the value of the stiffness rigidity, and therefore, the amount of damage can be obtained from the magnitude of changes in the curvature. Dynamic characteristics of an intact beam are not affected by an external force where $\{F\} = \{0\}$. However, decrease in the flexural rigidity causes a change in the stiffness matrix. This assumes any change to the mass matrix can be neglected. It was assumed that damage in a beam will affect only to the stiffness matrix and not to the mass matrix in the eigenvalue problem formulation and thus is consistent the assumption of Pandey [3]. The eigenvalue representation of an intact beam can be written as Equation (2)

$$([K] - \omega_i^2[M])\{a_i\} = \{0\} \tag{2}$$

and for the cracked beam as Equation (3)

$$([K']) - \omega_i'^2[M])\{a_i'\} = \{0\} \tag{3}$$

where K is the stiffness matrix of an intact beam, K' is the stiffness matrix of a damaged beam, M is the mass of the intact beam, M' is the mass of the damaged beam, ω_i is the ith natural frequency (eigenvalue) of the intact beam,

ω_i' is the ith natural frequency (eigenvalue) of the damaged beam, a_i is the ith displacement eigenvector of the intact beam and a_i' is the ith displacement eigenvector of the damaged beam. The damage level was modelled by adjusting the modulus of elasticity (E) in the finite element analysis.

The displacement mode shape results from the finite element analysis were used to estimate the curvature mode shapes of nonprismatic beams by using a central difference approximation as described by Equation (4)

$$\kappa = \frac{y_{i+1} - 2y_i + y_{i-1}}{l^2} \tag{4}$$

where y_i is the ith displacement mode and l is an effective length of the element. At each specified damage location, the percentage change in the natural frequency, the absolute difference of the displacement mode shapes and absolute difference of the curvature mode shapes were computed. The experiment conducted in the University of Gadjah Mada, the displacement modes were measured at each time increment and then by using the central difference method its curvature mode shape were obtained [4].

2.2 Estimation of residual strength

A post-cracking/residual flexural rigidity of the reinforced concrete beam can be estimated by computing the absolute difference of mode curvature based on the intact and damaged beams. Pandey [3] detected damage locations based on increases in the mode curvature. The curvature mode of the beam is derived from the displacement by central difference approximation (Equation 4) and the curvature at each point is also defined by using the relative displacements described by Equation (4). The natural frequency and curvature mode at each step of loading were measured, and based on these measurements the flexural rigidity can be written in terms of the function of natural frequency ω_n as detailed in Equation (5)

$$EI = \frac{m\omega_n^2 l^3}{c} \tag{5}$$

where m is the uniform mass density and c is a viscous damping constant. Substituting may be Equation (4) into Equation (1) obtaining the bending moment at any cross section M. The second integral of the general differential equation of a beam element enables computation of the deflection and maximum stress at any cross section along the beam. Application of this equation and assuming the first crack occurs at the first yield, the damage index (*DI*) can be formulated

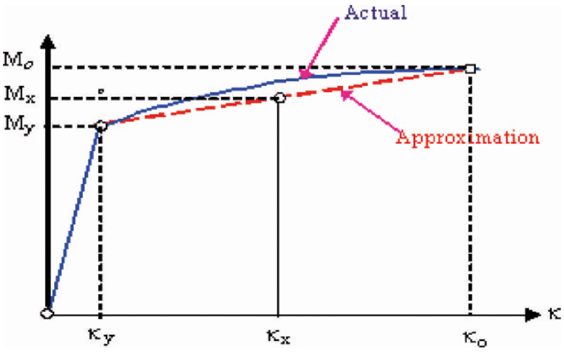


Figure 1. Moment curvature relationship for predicting the damage index.

as described by Equation (6)

$$DI = \frac{(1/\omega_x^2) - (1/\omega_r^2)}{(1/\omega_0^2 - 1/\omega_r^2)} \tag{6}$$

which ω_x = natural frequency at the estimated point, ω_r = natural frequency at the initial crack and ω_0 = natural frequency at the maximum damage. Previous researchers [5] have modified the damage index, i.e., called Modified Flexure Damage Index (*MFDI*), which relies on the use of the bilinear moment curvature relationship, refer Figure 1. This approach has been adopted in the formulation of the damage index presented in Equation (7). It has been assured that the first crack occurs at the first yield and the damage index can be redefined using Equation (7).

$$MFDI = \frac{(\kappa_x/M_x) - (\kappa_y/M_y)}{(\kappa_0/M_0) - (\kappa_y/M_y)} \tag{7}$$

where the indexing of x, y and 0 represents the specific point being estimated, yield condition and maximum damage condition.

3. EXPERIMENTAL PROGRAM

An experimental program was conducted to detect damage to nonprismatic beams resulting from a series of load regimes. The investigation formed particularly on changes of curvature mode shapes of concrete beam elements when localized damage occurred. Eight 3 m long, nonprismatic beam specimens were fabricated simultaneously in the University of Gadjah Mada, Structures Laboratory [4]. The specimens were divided into four groups of beams with each group had two typical reinforced beams. The 28-day concrete compressive

Table 1. Beam specimen details.

Typical	Specimen	Longitudinal reinforcement	Transversal reinforcement	Treatment	Material properties
Beam 1	A	4Ø6	Ø6–200	—	
	B	4Ø6	Ø6–200	—	Concrete:
Beam 2	C	4Ø10	Ø6–75 and Ø6–150	—	$f_c' = 33.6$ MPa
	D	4Ø10	Ø6–75 and Ø6–150	—	$E_c = 2.8 \times 10^4$ MPa
Beam 3	E	4Ø10	Ø6–75 and Ø6–150	—	Steel:
	F	4Ø10	Ø6–75 and Ø6–150	—	$f_y = 290$ MPa
Beam 4	G	4Ø10	Ø6–75 and Ø6–150	App. initial crack	$E_y = 2 \times 10^5$ MPa
	H	4Ø10	Ø6–75 and Ø6–150	App. initial crack	

strength of the ready-mix concrete averaged 33.6 MPa for testing and analysis of the dynamic response. The average yield stress of the 4 mm and 6 mm bar diameter steel reinforcement was 290 MPa. Reinforcement details and geometry of the beam specimens are presented in Table 1 and Figure 2.

The natural frequency and damping ratio of the beam were determined by measuring the free vibration on beam specimens 1A and 1B due to impact and static loads. Static load was applied from zero in increment of 980 N until collapse occurred. Preliminary analysis of the beam indicated that maximum displacement would occur at nodal point 15. Hence, a concentrated load was applied at this point in the experiment. At each load step, these measurements of the frequency and amplitude of vibration were measured at each nodal and an average value recorded. The value of natural frequency could be directly obtained by processing the recorded data using FFTSCOPE program developed by Sulistiyo and Haryanto [6]. During the investigation, the displacement

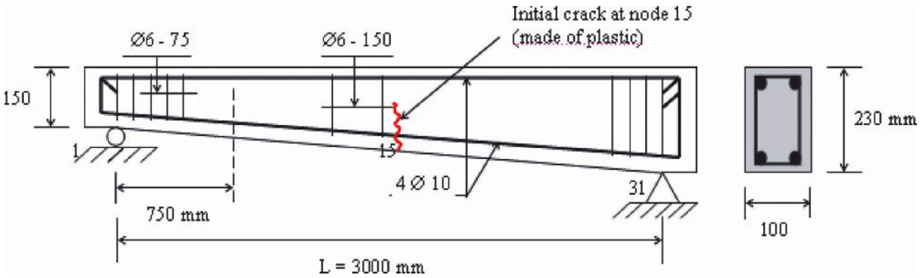


Figure 2. Typical specimens of nonprismatic beams.

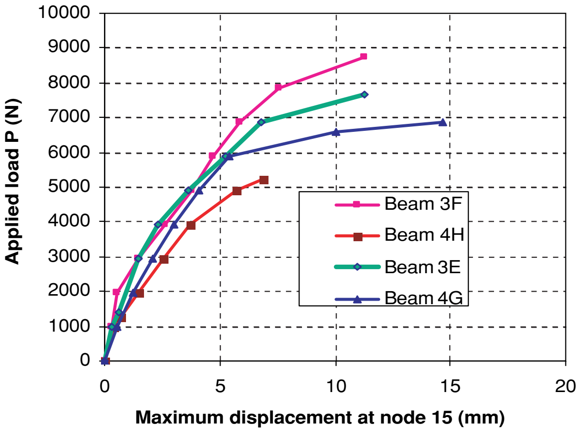


Figure 3. Measured load-displacement response.

at nodal points 11, 13, 15, 17 and 19 were subsequently observed at each increased load. Displacement measurements at five points were also recorded by dial gauges. Figure 3 shows applied load-maximum displacement response for four types of specimen and Figure 4 shows the frequency-applied load relationship.

4. NUMERICAL ANALYSIS

The numerical analysis of the nonprismatic beams focused on the displacement at node 15 (i.e., the maximum displacement) and the computed curvature

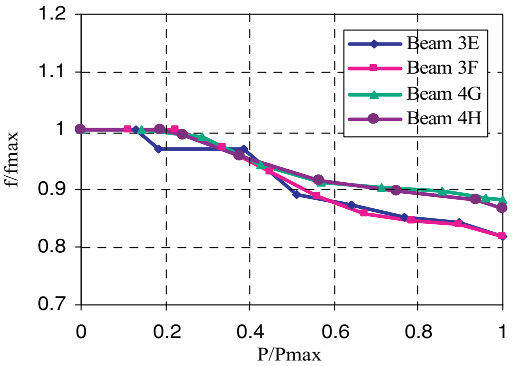
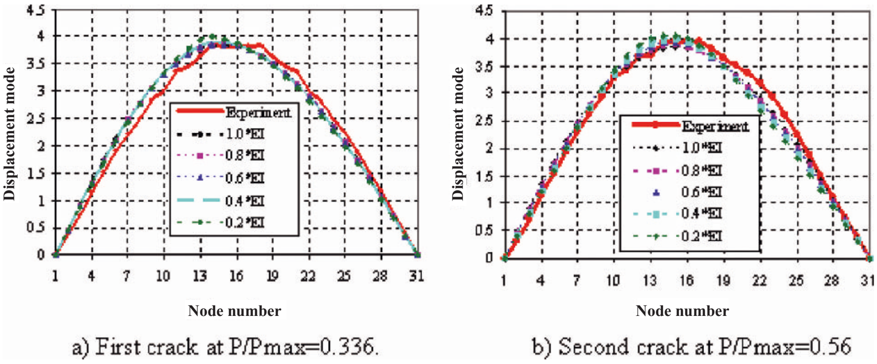


Figure 4. Measured frequency-load.



a) First crack at P/Pmax=0.336.

b) Second crack at P/Pmax=0.56

Figure 5. Typical displacement mode shapes of beam 3F at the first and second cracks.

mode and absolute difference of the curvature mode were compared with the experimental results. The two dimensional finite element program (SAP2000) was used to compute the natural frequency for each specimen. In this analysis, beams were modelled as solid and beam elements [7]. The equivalent method was adopted to calculate the equivalent moment of inertia at each element. Each 3 m, nonprismatic beam specimen was descretized into 30 elements with length of 100 mm and 31 nodes. The flexural rigidity of the intact element was assumed of 100% (*EImax*), and the *EI* of damage element was reduced 20% gradually. The reduction of *EI* in the numerical computation is relied on the element range of the experimental damage patterns, so that the determination of damaged element was different at each applied load. In other words, reduction on an element *EI* was depended on the crack patterns and the applied load increase. Based on this numerical computation, the displacement and the natural frequency were identified at each reduction of the *EI* and the displacement mode and curvature mode calculated and compared with the experimental results. The absolute difference of the curvature mode was also identified to determine the damage index and residual strength. Figure 5 shows a typical of plot of the displacement modes at the first and second cracks for beam 3F.

5. RESULTS AND DISCUSSION

The first two beams were used to measure natural frequency. Excitations for overall specimens were achieved by the measured of natural frequencies were in a range of 32.2–34.6 Hz. The frequency of mechanical vibrator was always lower than the natural frequency. At small applied loads, the frequency decreases tend to be linear curves, while at higher static loads the curve decreases gradually and nonlinearly (Figure 4). This result was consistent with the fact that the natural frequency of the beam reduced along with the level of damage propagation.

During the investigation, the cracked damage beams were observed at the first to third cracks representing a sequence of the damage. The damage locations and crack patterns were then measured by a micro crack instrument and plotted into the descretized elements assuming as a basis of stiffness reduction in the numerical analysis. At each crack load, the numerical computation was conducted by reducing of the flexural rigidity gradually.

Based on the crack loads, the residual strength of the beam specimen can be estimated. Theoretically the damage index can be estimated depending on the residual strength or vice versa. The damage index is a function of the natural frequency between the damaged and intact beam. Defining the curvature mode and the absolute difference of the curvature mode was dependent on the first displacement mode shape resulting from the finite element analysis. The residual strength of the damaged beam was achieved in three ways; firstly by using the damage index formulation found from the recorded dial gauge and static load, secondly by using the natural frequency substituted into Equation (6) or (7) and thirdly by plotting the absolute difference of the experimental curvature mode and the numerical curvature mode. The complete results are summarized in Table 2.

6. CONCLUSIONS

The numerical results for the simply supported nonprismatic beam model demonstrate the capability of curvature mode shapes in detecting and locating a state of damage. It has been shown that changes in the curvature mode shapes are localized in the damage regions. In contrast, changes in the displacement mode shapes are not localized and therefore they do not provide any indication of the damage location. To obtain the curvature mode shapes experimentally is very tricky. However they can be calculated accurately by nonlinear finite element analyses. The uncertainty of damage models makes the prediction less reliable particularly in the nonlinear material behaviour. It is shown that the damage index predicted by the MFDI and static strength approach has a good compromise and more realistic measure of the cracked damage in nonprismatic reinforced concrete beams. This method requires some modifications for future study in nonlinear finite element analysis.

Table 2. Classification of damage beams.

			Damage index				
Beam type	P/Pmax	Crack level	Static Strength	MFDI	Difference abs. of curvature mode	Damage location (element number)	Damage level
3E	0.000–0.513	1st	0.000–0.404	0.000–0.461	0.000–0.263	15	1
	0.513–0.641	2nd	0.404–0.682	0.461–0.575	0.263–0.459	7–20	2–3
	0.641–1.000	3rd	0.682–1.000	0.575–1.000	0.459–1.000	7–20	3
3F	0.000–0.448	1st	0.000–0.554	0.000–0.224	0.000–0.207	15	1
	0.448–0.560	2nd	0.554–0.732	0.224–0.4S0	0.207–0.471	9–18	2
	0.560–1.000	3rd	0.732–1.000	0.4S0–1.000	0.471–1.000	6–20	2–3
4G	0.000–0.426	1st	0.000–0.358	0.000–0.339	0.000–0.261	15	1
	0.426–0.571	2nd	0.358–0.642	0.339–0.709	0.261–0.480	7–20	2–3
	0.571–1.000	3rd	0.642–1.000	0.709–1.000	0.481–1.000	5–22	3
4H	0.000–0.375	1st	0.000–0.450	0.000–0.248	0.000–0.251	14–15	1
	0.375–0.563	2nd	0.450–0.710	0.248–0.566	0.251–0.460	13–18	2–3
	0.563–1.000	3rd	0.710–1.000	0.566–1.000	0.460–1.000	7–20	3

Note of damage level: 1 = low (0–0.25). 2 = moderate (0.26–0.50) and 3 = high (0.51–1.00).

REFERENCES

1. D.G. Fertis (1994), Inelastic analysis of prismatic and nonprismatic aluminum members. *Computers and Structures*, 52, 2, pp. 287–295.
2. M.M.F. Yuen (1985), A numerical study of the eigenparameters of a damaged cantilever. *Journal of Sound and Vibration*, 103, pp. 301–310.
3. A.K. Pandey, M. Biswas and M.M. Samman (1991), Damage detection from changes in curvature mode shapes. *Journal of Sound and Vibration*, 145, 2, pp. 321–332.
4. F. Saleh (2000), Detection of Cracks in Nonprismatic Beams by Curvature Mode Shapes. Master Thesis. University of Gadjah Mada, Yogyakarta, Indonesia (in Indonesian).
5. B. Suhendro. Analysis of Nonlinear-inelastic Dynamic and Concept of Damage Factor for Seismic Design of Structural Reinforced Concrete Frame. *National Seminar on Seismic Resistant of Structural High-rise Frames*, Islamic University of Indonesia in cooperation with Gresik Cement Company, Yogyakarta, Indonesia (in Indonesian).
6. P. Sulistiyo and H. Haryanto (1997), Data Processing for Vibration Analogue by Fast Fourier Transformation (FFT) Program. *Civil Engineering Department Seminar*, University of Gadjah Mada, Yogyakarta, Indonesia (in Indonesian).
7. E.L. Wilson (1996), *SAP2000 Integrated Finite Element Analysis and Design Structures*. CSI Computers & Structures Inc., Berkeley, CA.

NUMERICAL STUDY ON CONFINING PRESSURE EFFECT IN THE PROCESS OF ROCK FAILURE

D.P. Qiao[1], Y.N. Sun[1], S. H. Wang[2] and C.A. Tang[2]
[1]*Faculty of Land Resources Engineering, Kunming University of Science and Technology, Kunming 650093, P. R. China*
[2]*School of Resource and Civil Engineering, Northeastern University, Shenyang 110004, P. R. China*

Abstract The deformation mechanism of rock under different constant confining pressure was briefly analysed based on continuum damage mechanics and the effects of confining pressure on deformation, strength and macroscopic fracture patterns of model rock specimens are also studied using the Rock Failure Process Analysis (RFPA^{2D}) code. The theoretical analysis and numerically obtained results duplicate the deformation, strength (such as Young's modulus, compressive strength, etc.) and macroscopic fracture patterns observed in laboratory. The theoretical studies and numerical simulations are extremely instructive and indicative for investigating some catastrophic hazard phenomena such as rock bursts and instability induced by excavation.

Keywords: rock fracture, confining pressure effect, RFPA code, fracture patterns.

1. INTRODUCTION

Constitutive relations of rocks under the effects of confining pressure are very important to energy exploitation and prevention of natural disasters. Nowadays the study on the effect of the confining pressure on macroscopic mechanics is confined to the laboratory tests. It should be noted that differences in rock samples and test methods could result in big differences in test results. For this reason, a numerical study was conducted on the deformation strength and macroscopic damage mode under different confining pressures through the RFPA^{2D} system, in order to obtain some references for the forecasting of disasters in rock engineering.

G. R. Liu et al. (eds.), Computational Methods, 1817–1822.

Table 1. Parameters of numerical model.

Parameters	Value
Even degree m	1.5
Average elastic module E_0/GPa	30
Average compression strength σ_0/MPa	200
Friction angle $\phi/^\circ$	30
Compression/tension C/T	10
Poisson's ratio μ	0.25
Coefficient of the residual strength ξ	0.1
Maximum extension strain coefficient	1.5
Maximum compression strain coefficient	200
Confining pressure σ_3/MPa	0, 2, 4, 8

2. NUMERICAL MODELS

The numerical tool used in this paper is the Rock Failure Process Analysis (RFPA^{2D}) which is developed to account for the randomness and heterogeneity of the REV, which generates the Weibull statistic distribution. The REV is supposed to be crisp elastic. Because the tensile strength of rock is far smaller than its compressive strength, we suppose that the REV has two types of damages: tensile failure and shear failure. When the minimum principal stress is larger than the tensile strength, tensile failure appears; when the shear stress meets the tensile cut-off of the Mohr-Coulomb principle, shear failure appears.

The plane strain model is adopted with a model size of 100 × 50 mm. The model is divided into 200 ×100 REVs. The loading process is by displacement control and the load step is Δ s = 0.002 mm. The specific mechanic parameters are given in Table 1.

3. THE EFFECTS OF THE CONFINING PRESSURE ON ROCK DEFORMATION

Due to the fact that the stress status of actual rock is nonhomogenous, it is unlikely to display all the compound modes of the stress status when it is damaged. However, we still can study the effects of the stress status on the damage by studying the ultimate strength of the rock under different confining pressures. In order to eliminate the effects of specimen variation, we did the compressive loading test under different confining pressure on the same specimen. From Figure 1, we can see that increasing the confining pressure increases the Young's modulus and the ultimate compression strength. Figure 2 is the experimental stress–strain curves for granite under different constant confining

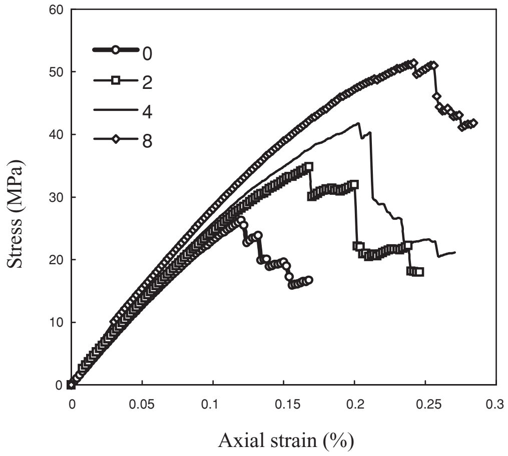


Figure 1. Stress–strain curves for rock specimens under different constant confining pressures.

pressures. From Figures 1 and 2, we can see that most of the simulation results are in agreement with the laboratory results.

4. EFFECTS OF THE CONFINING PRESSURE ON STRENGTH

According to Figure 1 and single axle failure strength of rock samples, the curves of the confining pressure and the ultimate compression strength

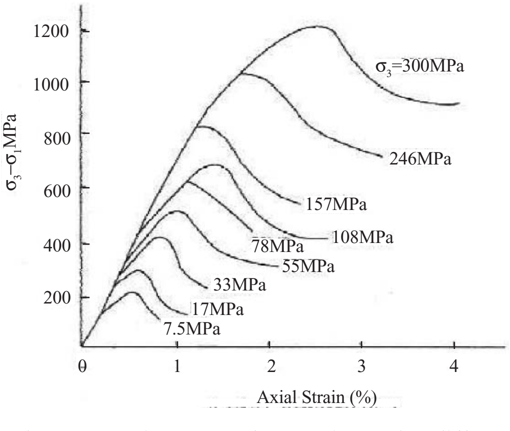


Figure 2. Experimental stress–strain curves for granite under different constant confining pressures.

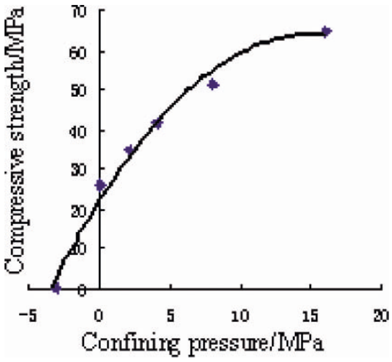


Figure 3. Curve between compressive strength of rock specimens and confining pressure.

can be drawn (Figure 3). Although the damage principles of the REV adopt the linear-elastic Coulomb principle, the numerical test results indicate that the relation between confining pressure (or minimum principle stress) and ultimate compression strength of rock samples is nonlinear.

Based on the ultimate compression strength of rock samples under different confining pressures, the Mohr circles and envelope line of every kind of stress conditions is plotted in Figure 4. Figure 4 shows that the simulation results are very similar to that of the physical test results. Although the damage principle which simulates the REV is the linear Coulomb principle, the macroscopic representation is nonlinear. In the RFPA^{2D} simulation, the damage principle of each element is very simple—amended Coulomb principle with tensile cut-off. The most important reasons are that in this model, the radical characters of the rock, namely, the heterogeneity and the interaction between the REV are taken into account.

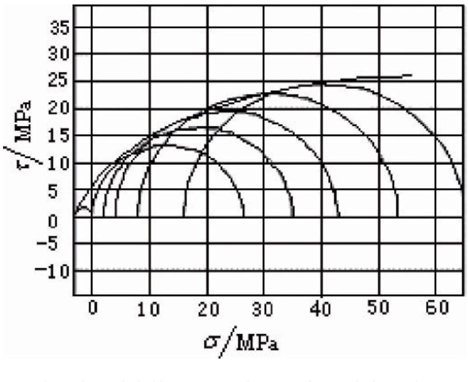


Figure 4. Simulated failure envelope of model rock specimens.

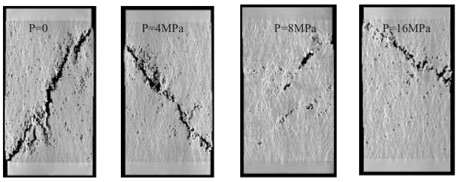


Figure 5. Macroscopic failure patterns of model specimens under different constant confining pressures.

5. THE EFFECT OF THE CONFINING PRESSURE ON THE FRACTURE MODE

A large number of experimental data show that the fracture mechanics of the intact rock has close relationship with confining pressure. In practical situations, the rock is under 2 or 3 dimensional stress, and different confining pressure conditions obviously affect the fracture types.

Figure 5 shows the simulation results of the macroscopic fracture mode of the rock under different confining pressures. We can find that with the enhancing of the stress level of the confining pressure, the angle of the macroscopic fracture plane changes gradually. When the confining pressure is 0, the angle between the macroscopic fracture plane and the maximum principle stress increases gradually (approximately 35°–40°), at the same time, the damage of the rock transits and develops from tensile failure to shear failure. But under high confining pressure, the angle between the ultimate fracture plane and the compressive axis is about 45° and the macroscopic characters of the rock take show some ductility.

6. CONCLUSIONS

In this paper, we built the damage mechanics model under different confining pressure based on the microscopic continuum damage mechanics. The model gives a good reflection of actual mechanical parameters such as the increase in elastic modulus and ultimate compression strength with the increase in confining pressure.

Although in RFPA^{2D}, the elastic crisp constitutive relation and simple failure principle of the REV are adopted, the simulated results show that with the increasing of the confining pressure, the simulation still gives a nonlinear deformation behaviour which transit from crisp damage to ductile damage. All these simulation results are identical to the phenomena observed in laboratory, which indicates that the evolutionary principles of REV adopted in RFPA^{2D} are suitable to simulate complicated and nonlinear evolutionary damage of the rock material. This study and its results are helpful to give theoretical and practical directions.

REFERENCES

1. T.N. Gowd and F. Rummel (1980), Effect of confining pressure on the fracture behavior of a porous rock. *International Journal of Rock Mechanics, Mining Sciences and Geomechanical Abstracts*, 17, pp. 225–229.
2. M.F. Ashby and C.G. Sammis (1990), The damage mechanics of brittle solids in compression. *Pure and Applied Geophysics*, 133, pp. 489–521.
3. C.A. Tang (1997), Numerical simulation of progressive rock failure and associated seismicity. *International Journal of Rock Mechanics and Mining Science*, 34, 2, pp. 249–261.
4. C.A. Tang, L.G. Tham, P.K.K. Lee, et al. (2000), Numerical studies of the influence of microstructure on rock failure in uniaxial compression—Part II: Constraint, slenderness and size effect. *International Journal of Rock Mechanics and Mining Science*, 37, 4, pp. 571–583.
5. M.F. Ashby and C.G. Sammis (1990), The damage mechanics of brittle solids in compression. *Pure and Applied Geophysics*, 133, pp. 489–521.
6. M.Q. You and A.Z. Hua (1998), Fracture of rock specimen and decrement of bearing capacity in uniaxial compression. *Chinese Journal of Rock Mechanics and Engineering*, 17, 1, pp. 24–29.
7. X.R. Ge, J.X. Ren, Y.B. Pu, et al. (1999), A real-in-time cttriaxial testing study of meso-damage evolution law of coal. *Journal of Applied Mechanics*, 18, 5, pp. 497–502.

APPLICATION OF GEOSTATISTICAL WEIGHTS IN SOLVING PROBLEMS GOVERNED BY 2-D POISSON'S EQUATION USING FINITE POINT METHOD

N. Fathianpour, H. Hashemolhosseini and M. Tabatabaie
Department of Mining Engineering, Isfahan University of Technology, Isfahan 84156, Iran

Abstract The modified Finite Point Method (FPM) is considered as one of the latest developments among the meshless numerical methods. In order to achieve accurate results using modified FPM, it is essential to impose suitable mixed Boundary Conditions (BC) along with simultaneously satisfying both the governing Partial Differential Equation (PDE) at internal nodes and the associated BC at the nodes located on inhomogeneous interfaces in addition to the nodes at external boundaries of a normalized spatial domain. Through solving Homogeneous and Inhomogeneous Geoelectric Models (HGM and IGM) of geophysical science using weighted FPM we show that employing geostatistically optimized weights obtained via solving Kriging equations leads to results with greater accuracy compared with those obtained using conventional Gaussian weight.

Keywords: FPM, BC, PDE, HGM, IGM, geostatistical weights, Kriging equations.

1. INTRODUCTION

The theoretical basis of meshless methods and in particular the so called Finite point Method (FPM) have been explained by many researchers [1–5]. The improved weighted FPM as developed by Onate et al. [4] using Finite Calculus (FC) is considered as a truly meshless numerical method that has the advantages of both solution stability and precisely handling Drichlet and Neuman boundary conditions. In the general form of weighted finite point formulation the approximation at a point is done in least squares sense likewise the diffuse element (DE) and Element Free Galerkin (EFG) method. Also the final discretized system of equations are obtained by sampling the governing Partial Differential

G. R. Liu et al. (eds.), Computational Methods, 1823–1828.

Equations (PDE) at all nodal points as is normally done in generalized Finite Difference (FD) method. In this paper, we investigate a different aspect of FPM that is the selection of optimum weight coefficients in solving Boundary Value Problems (BVP) governed by Poisson's equations. First we will present some important properties of geostatistical estimators followed by deriving Kriging weight coefficients to be incorporated in the FPM formulation.

2. PROPERTIES OF GEOSTATISTICAL INTERPOLATING ESTIMATORS

In all numerical methods based on least square minimization of approximation errors one of the main issues is the interpolation procedure. In order to have an optimum estimator, three important conditions are to be satisfied. These conditions are (1) the estimator should have an exact interpolating property which means that the estimation of each nodal value should equal the exact value of the node value, (2) the approximation should have the minimum estimation variance and (3) the estimation should be unbiased. Investigating the spatial behaviour of the analytical solutions of Poisson's equation with the impulse source shows that they are inherently functions of distances from their impulse source and hence applying the geostatistical estimators as the main interpolating weights might led to a more accurate solutions using weighted FPM. In geostatistical literature the mean of squared sum of differences between the pair nodal values h meters apart is called the point variogram of the unknown function and is denoted by $2\gamma(h)$.

$$2\gamma\,(h) = E\left[(z(x_i + h) - z(x))^2\right] = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [z(x_i + h) - z(x_i)]^2$$

Each variogram is characterized by three basic parameters called nugget effect (random portion of variation), range (radius of influence) and sill value (the spatial portion of the variation). The common theoretical variogram models having sill values are known as linear, spherical, exponential, Gaussian and sinusoid models. Here we have used the least squares approach to find the best fit between the experimental variogram and one of the theoretical models. Figures 1 and 2 show the experimental variograms computed for HGM and IGM models, respectively. The best obtained model for both media is the Gaussian model with the following equations.

$$\gamma(h) = C\left(1 - e^{\frac{-h^2}{a^2}}\right) + C_0 \quad \gamma(h)_{HGM} = 4.5\left(1 - e^{\frac{-h^2}{0.0529}}\right)$$

$$\gamma(h)_{IGM} = 6.2\left(1 - e^{\frac{-h^2}{0.0625}}\right)$$

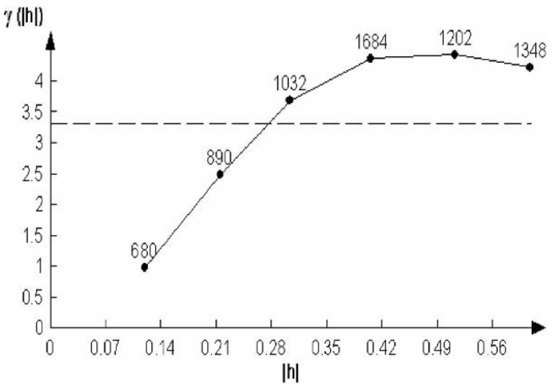


Figure 1. Experimental and fitted variograms for HGM model.

in which C_0 is the nugget effect, *a* is the range and C is the sill value of the variogram.

3. APPLYING KRIGING WEIGHTS IN FINITE POINT METHOD

The Kriging estimator is based on a weighted moving average interpolation using neighbourhood nodal values. Here the recommended Kriging estimator is defined as a simple moving average interpolator as follows.

$$Z_v^* = \sum_{i=1}^{n} \lambda_i Z_{v_i}$$

in which λ_i is the weight for the i-th nodal value within the influence radius of unknown nodal value being estimated, Z_{v_i} is the value of i-th nodal point

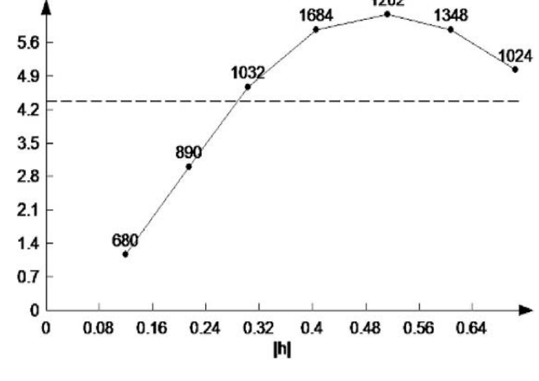


Figure 2. Experimental and fitted variograms for IGM model.

and Z_v^* is the estimated value at the nodal point under estimation. Under the second stationary condition the Kriging equations to be solved for weighting coefficients is: A*X = B

$$A = \begin{bmatrix} \overline{\gamma}_{11} & \overline{\gamma}_{12} & \mathrm{K} & \overline{\gamma}_{1\mathrm{n}} & 1 \\ \overline{\gamma}_{21} & \overline{\gamma}_{22} & \mathrm{K} & \overline{\gamma}_{2\mathrm{n}} & 1 \\ \mathrm{M} & \mathrm{M} & \mathrm{K} & \mathrm{M} & \mathrm{M} \\ \overline{\gamma}_{\mathrm{n}1} & \overline{\gamma}_{\mathrm{n}2} & \mathrm{K} & \overline{\gamma}_{\mathrm{nn}} & 1 \\ 1 & 1 & \mathrm{K} & 1 & 0 \end{bmatrix} X = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \mathrm{M} \\ \lambda_n \\ \mu \end{bmatrix} B = \begin{bmatrix} \overline{\gamma}_{1v} \\ \overline{\gamma}_{2v} \\ \mathrm{M} \\ \overline{\gamma}_{nv} \\ 1 \end{bmatrix}$$

in which $\bar{\gamma}_{ij}$ is the average variogram between nodal points i and j, γ_i is the Kriging weight for the i-th nodal value and $\bar{\gamma}_{iv}$ is the average variogram value between i-th nodal value and the current unknown nodal point. Accordingly it is necessary to search for the nodal values in the range of the modelled variogram distance.

3.1 Example 1: Electric potential distribution in HGM model

In this paper, we adopt the FPM formulation originally developed by Onate et al. [3]. The Poisson's equation and the accompanied boundary conditions on a normalized domain are assumed. In the case of one current electrode located in a HGM with normalized spatial extent and a constant electrical resistivity (10 Ohm-Meter) it can be shown that the potential is a function of radial distance from current electrode and is given by $u = (\rho I/4\pi)1/r$. Figure 3 shows the

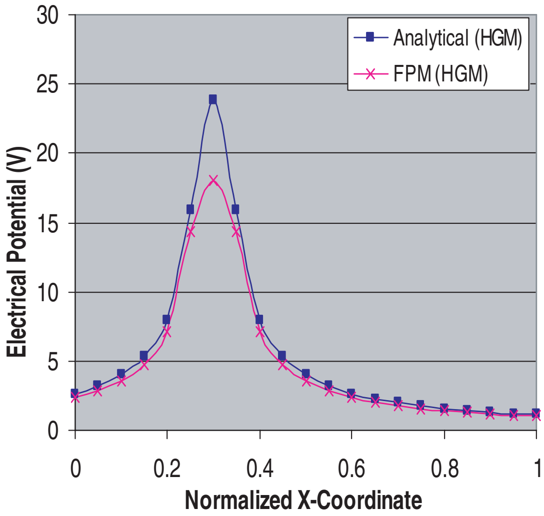


Figure 3. Analytical (red) and FPM solutions of electric potentials for the HGM model.

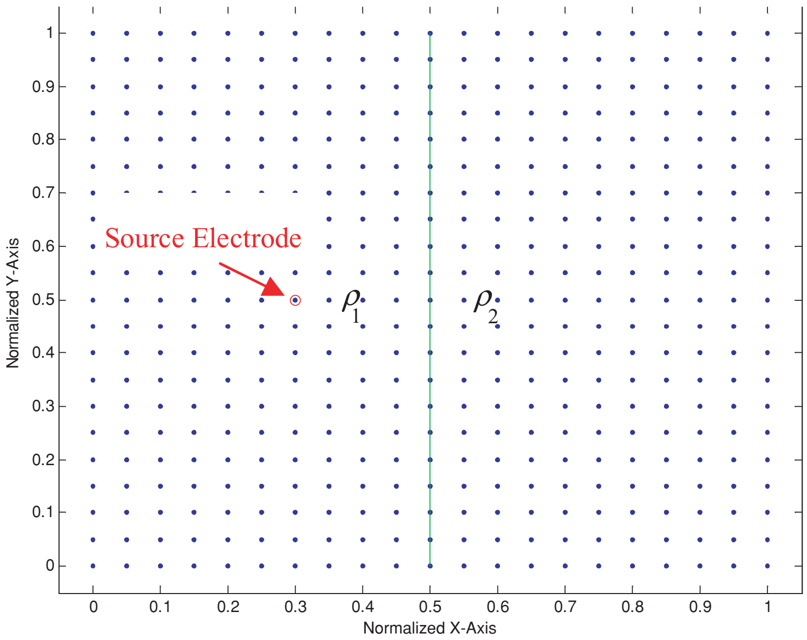


Figure 4. The geometry used for the IGM model with vertical contact located at $x = 0.6$.

results of applying Kriging weights in finite point approach for HGM on a normalized spatial extent using a regular grid of 20 by 20. The source current electrode is located at (0.3,0.5).

3.2 Example 2: Electric potential distribution in IGM model

Figure 4 shows the IGM model consisting of two medium with differing resistivities (10 and 100 Ohm-Meters) separated with a vertical contact. Figure 5 depicts the results obtained for the IGM model using Kriging coefficients on the same domain as Example 1.

4. CONCLUSIONS

In the previous works presented by Onate et al. [3, 4] a standard normalized Gaussian weight function have been used giving neither proof for its effects nor specific criteria for selecting searching radius for solving any particular PDE. However in this study, we have shown that the best weighting coefficients for the solution of Poisson’s type BVP’s obeys the Gaussian variogram model

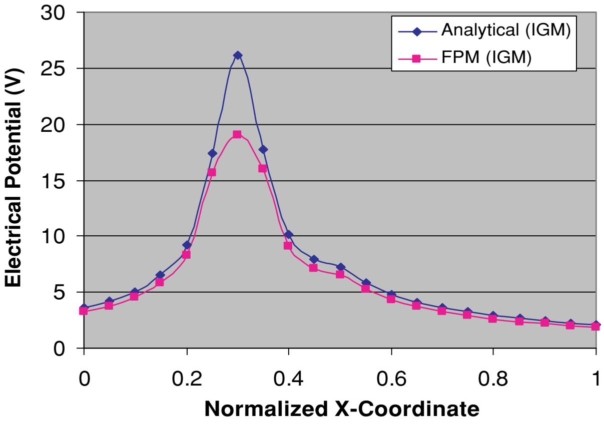


Figure 5. Analytical (red) and FPM solutions of electric potentials for the IGM model.

resulting in more than 5% reduction on average approximation errors compared to conventional Gaussian weights. Obviously using such geostatistical weighs requires the solution of Kriging equations for every nodal point distributed all over the domain of problem in a moving average sense.

REFERENCES

1. T. Belytschko, Y.Y. Lu and L. Gu (1994), Element free Galerkin methods. *International Journal for Numerical Methods in Engineering,* 37, pp. 229–256.
2. G.R. Liu (2003), *Mesh Free Methods: Moving Beyond the Finite Element Method.* CRC Press, Boca Raton, USA, pp. 83–161.
3. E. Onate, S. Idelson, O.C. Zienkiewicz and T. Fisher (October 1995), A finite point method for analysis of fluid flow problems. In: Proceedings of the 9th International Conference Finite Element Methods in Fluids, Veniz, Italy, pp. 15–21.
4. E. Onate, S. Idelson, O.C. Zienkiewicz, R.L. Taylor and C. Sacco (1996), A stabilized finite point method for analysis of fluid mechanics problems. *Computer Methods application Mechanical Engineering,* 139, pp. 315–346.
5. E. Onate, F. Perazzo and J. Miquel (2001), A finite point method for elasticity problems. *Computers and structures,* 79, pp. 2151–2163.

NUMERICAL APPROACH TO FRACTURES SATURATION BEHAVIOUR IN HETEROGENEOUS MATERIAL SUBJECTED TO THERMAL LOADING

L.C. Li[1], C.A. Tang[1], T.H. Yang[1], and L.G. Tham[2]
[1]*Center for Rock Instability and Seismicity Research, Northeastern University, Shenyang, China, 110004*
[2]*Department of Civil Engineering, The Hong Kong University, Hong Kong, China*

Abstract Based on the heterogeneous characteristics of rock and concrete at mesoscopic level, a 2-D mesoscopic thermo-mechanical-damage (*TMD* model) coupling model, implemented in RFPA^{2D}, is proposed by introducing elastic damage mechanics and thermal-elastic theory. The TMD model is used to study the thermal induced cracking process, including crack formation, extension and coalescence in heterogeneous materials subjected to mechanical and thermal loading. The numerical results compare well with the corresponding mathematic resolution and experimental results, which prove that the proposed model as well as the numerical system (RFPA^{2D}) is reasonable and effective in investigating the failure of heterogeneous materials subjected to thermomechanical action.

Keywords: numerical simulation, heterogeneity, damages, thermal conductivity.

1. INTRODUCTION

The physicochemical environment of geologic systems is host to various coupled thermal, hydraulic, mechanical and chemical processes. Scientific interest in these coupled physicochemical processes in the earth's crust, in general, and economic and environmental concerns related to waste geologic disposal, in particular, have resulted in many research efforts aimed at understanding the coupled thermal, hydrologic chemical and mechanical (*THMC*) behaviour of geologic systems subject to complex natural or man-made perturbations. Due to the complexity of the problem, one of the most promising approaches is

G. R. Liu et al. (eds.), Computational Methods, 1829–1834.

to develop numerical methods for analyses. Lately two international forums, namely DECOVALEX and INTERCLAY, provided great opportunities for development of *THM* research [1].

To reach such a goal, a numerical method was developed. However, only physical phenomena, namely, thermal and mechanical will be addressed. Implemented with a Rock Failure Process Analysis code (RFPA^{2D}) [2], the *TMD* (thermomechanical-damage) model can be used to investigate the behaviour of heat transfers, damage evolution and the overall response at the mesoscopic level.

2. BRIEF OUTLINE OF THE ROCK FAILURE PROCESS ANALYSIS NUMERICAL CODE, RFPA^{2D}

The Rock Failure Process Analysis (RFPA^{2D}) code based on *FEM*, is a numerical tool capable of handling the fracture initiation, interaction and coalesce process in heterogeneous materials. A model of coupling between thermal, mechanical and damage (*TMD*) has been proposed and implemented within RFPA^{2D}.

For heterogeneity, the material properties (failure strength σ_c and elastic modulus E_c) for mesoscopic elements are randomly distributed throughout the rock sample by following a Weibull distribution [2]:

$$f(u) = \frac{m}{u_0}\left(\frac{u}{u_0}\right)^{m-1} \exp\left(-\frac{u}{u_0}\right)^{m} \tag{1}$$

where u is the parameter of the element; the scale parameter u_0 is the average of element parameter and the parameter m defines the shape of the distribution function. The parameter m defines the degree of material homogeneity and is called the homogeneity index. Materials with higher m values represent more homogeneous materials, whereas those with lower m values, more heterogeneous materials. Figure 1 shows the variations of f with respect to m.

3. MODEL VALIDATION

For the sake of validating the proposed model, thermal induced stresses in a homogeneous material, homogeneity index $m = 1000$, was numerically studied and Compared with an analytical solution derived from classical theory of thermoelasticity. A two-dimensional numerical rectangle board of 50-mm wide and 100-mm long is designed as shown in Figure 2. The Young's modulus (E) of the numerical sample is 50,000 MPa. The numerical sample was divided

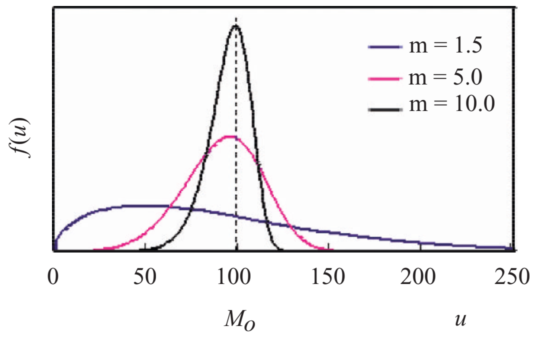


Figure 1. Weibull's distribution for mechanical properties with different homogeneity index.

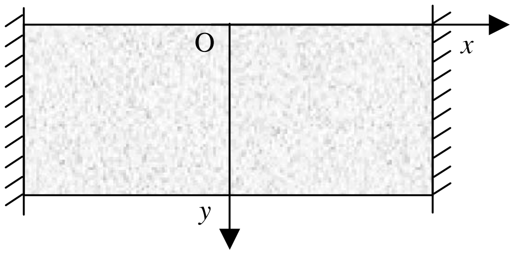


Figure 2. Configuration for model validation.

into 50 × 100 meshes. Temperature of 100°C and 10°C were applied to bottom and upper boundary respectively.

The temperature distribution in the sample may be expressed as: $t = a_0 y$. Then theoretically, the stress in the sample are $\sigma_x E\alpha a_0 y$, $\sigma_y = 0$, $\tau_{xy} = 0$ (where *a* is the thermal expansion coefficient). Figure 3 shows the stress distribution along the coordinate *y* axis from origin point *O*. The simulation results are in good agreement with the the analytical solution derived from classical theory of thermoelasticity.

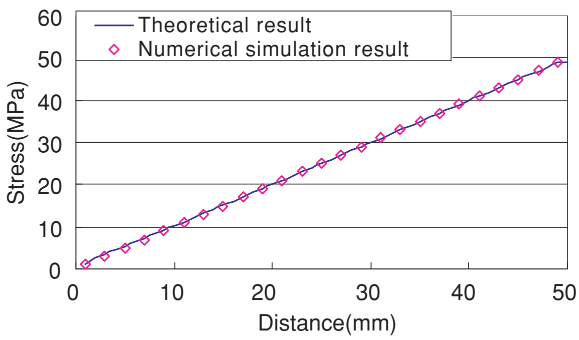


Figure 3. Stress distribution along the *y* axis.

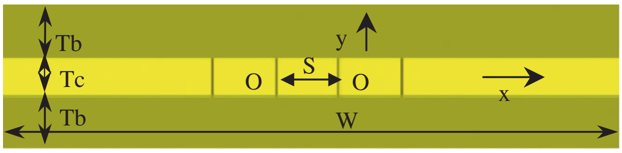


Figure 4. The three-layer model for numerical simulation of fractures saturation.

4. NUMERICAL SIMULATION OF FRACTURES SATURATION PHENOMENA

Having the model validated, further analyses have been carried out to study fracture saturation, the famously geological phenomena [3]. We use a three-layer model, as used in many previous studies. The mesh for the model contains $400 \times 60 = 24,000$ elements with geometry of 8000×1200 mm. The homogeneity index m is chose to be 3. Thickness of the central layer (T_c) is 300 mm. The overall thickness of the model (T) is 1200 mm, which can be defined as $T = T_c + 2T_b$. As shown in Figure 4, the right and left boundary are fixed in x direction, both the bottom boundary and the top boundary are free. The three layers are subjected to a uniform temperature decline from 0°C to −88°C at a step of −0.4°C. The Young's modulus (E_b), compressive strength (σ_b) and Poisson's ratio (υ_b) are same for adjacent layers, $E_b = 10$ GPa, $\sigma_b = 200$ MPa, $\upsilon_b = 0.4$, and those for the central layer are 50 GPa, 100 MPa and 0.25, respectively. In addition, the thermal properties for adjacent layers are thermal conductivity coefficient 1.2 W/m·°C and thermal expansion coefficient 1.0×10^{-5} 1/°C, and those for the central layer are 1.1 W/m·°C and 0.9×10^{-5} 1/°C. The space of adjacent fractures is denoted as S. As illustrated in previous study [3], the critical value (the spacing to fractured layer thickness ratio S/T_c) determines the stress state between the adjacent fractures: when the ratio is below this critical value the stress along line O–O is compressive, while above the critical value the stress is tensile.

The process of fracture saturation is clearly shown in Figure 5. The different grey scales represent different values of normalized maximum shear stress. At temperature −2°C, there is no fracture. When the temperature descends to −5°C, new fractures come into being. With the descending of temperature, more fractures form. At last, the fractures are spaced so closely that no more new fractures can infill, even with descending temperature (−74°C to −88°C) and the three-layer model breakdown completely. In saturation state, we obtain that the spacing to fracture thickness ratio S/T_c is 1.47, which is bigger than the critical value (1.0) mentioned above. The reason is that the tensile stress between adjacent fractures is bigger than zero but it is not big enough to offer the tensile stress for new fracture formation.

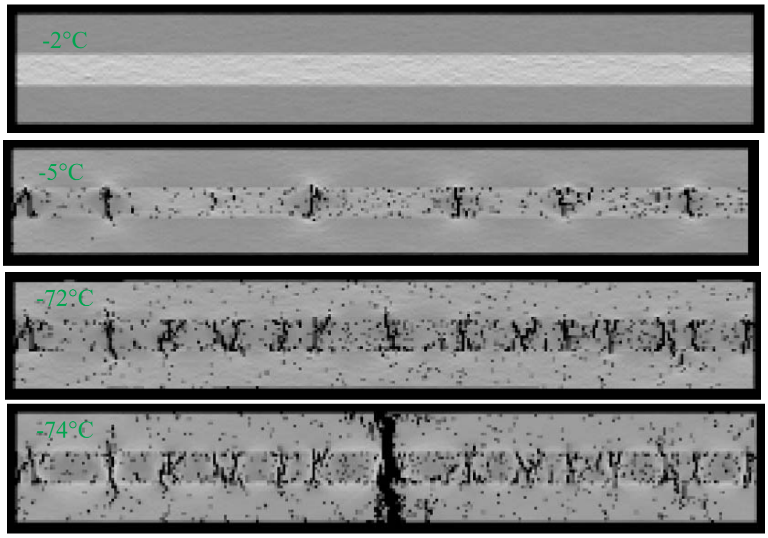


Figure 5. The process of fracture saturation simulated with RFPA^{2D}.

One important method of observing damage during the rock deformation experiments is monitoring the acoustic emissions (AE). As shown in Figure 6, there are certain AE events before new fractures form, but the magnitude is small. This indicates that small joints initiate in the model. As one new fracture forms, and large amounts of AE events occur correspondingly. While in saturation state, the magnitude of AE events is also small and the cumulative AE counts main constant. This indicates that no more new fractures form in the saturation state. The total number of the large-scale AE occurrences is 17, which means that 17 fractures saturate the whole layer. This number is consistent with the results as shown in Figure 5.

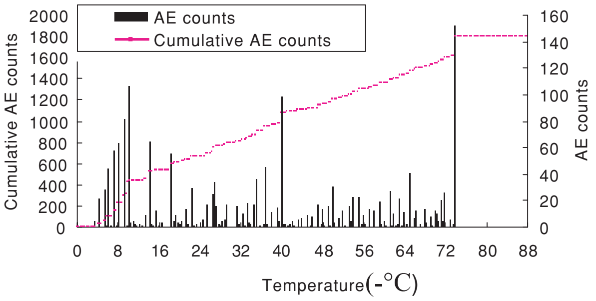


Figure 6. AE counts with temperature obtained from RFPA^{2D}.

5. CONCLUSIONS

It indicates that the RFPA^{2D} is a valid tool in understanding the physical essence of the evolutionary nature of fracture phenomena in heterogeneous materials subjected to thermal and external loading. The analysis provides a deep insight in the influence of spatial variation in mesomechanical properties on the fracture or damage process as well as thermal conductivity through a statistical modelling scheme. The phenomenological approach shows in details the nucleation, propagation and coalescence of micro-fractures. It gives a clear explanation to fracture saturation phenomena, the famously geological phenomena.

ACKNOWLEDGEMENTS

This work presented in this paper has been supported by the National Natural Science Foundation (Grant No. 50490274, 50134040 and 50204003) of China.

REFERENCES

1. C.F. Tsang (ed.) (1987), *Coupled Processes Associated with Nuclear Waste Repositories*. Academic Press, Orlando, FL.
2. C.A. Tang and P. K. Kaiser (1998), Numerical simulation of cumulative damage and seismic energy release in unstable failure of brittle rock—part 1 Fundamentals. *International Journal of Rock Mechanics and Mining Sciences*, 35, pp. 113–121.
3. T. Bai, D.D. Pollard and H. Gao (2000), Explanation for fracture spacing in layered materials. *Nature*, 403, pp. 753–756.

MICROMECHANICAL MODEL FOR SIMULATING THE HYDRAULIC FRACTURES OF ROCK

T.H. Yang[1], L.C. Li[1], L.G. Tham[2] and C.A. Tang[1]
[1]*Center for Rock Instability and Seismicity Research, Northeastern University, Shenyang, China, 110004*
[2]*Department of Civil Engineering, The Hong Kong University, Hong Kong, China*

Abstract Rock permeability is not a constant but a function of stresses and stress-induced damage. In this paper, a seepage-damage coupling model for heterogeneous rocks by taking into account the growth of the existing fractures and formation of new fractures is proposed. Implemented with the Flow-coupled Rock Failure Process Analysis code, this coupling model has been used to investigate the behavior of fluid seepage, damage evolution and their coupling action in samples subjected to both hydraulic and biaxial compressive loading. The modeling results suggest that the nature of fluid flow in rocks varies from material to material and depends strongly upon the damage of the rocks.

Keywords: mescoscopic, heterogeneous, hydraulic fracturing, simulation.

1. INTRODUCTION

Except for being used in determining in situ stresses in rock masses, hydraulic fracturing is also a method used by oil companies to stimulate reservoir production. There are lots of different interpretation theories for hydraulic fracturing. But there are at least two drawbacks in most of the hydraulic fracture theories. Firstly, the materials studied in most of the hydraulic fracture models are impermeable. Secondly, in most of the previous flow-coupled models, the influence of heterogeneity existed in rock on the fracture pattern or hydraulic fracture path cannot be taken into account.

Due to the difficulty to gain a complete solution of a hydraulic fracturing problem, numerical simulation methods are widely used. In this paper, we

G. R. Liu et al. (eds.), Computational Methods, 1835–1840.

employed the Flow-coupled Rock Failure Process Analysis code (F-RFPA^{2D}) to investigate the hydraulic fracturing phenomenon. The results indicate that both the rock heterogeneity and the permeability affect the fracture initiation and propagation significantly, and that the simplistic premise that rock is homogeneous and impermeable may apply to limited, but not general cases in hydraulic fracturing

2. NUMERICAL SIMULATION

In this section, we simulate the rock failure and the associated fluid flow numerically using F-RFPA^{2D} [1]. A benchmark case of fracturing of an elastic rock stratum by applying pressure in a borehole has been analyzed by using the present model (Figure 1). The main mechanical parameters are listed below: Poisson's ratio is 0.25, Permeability is 00.01 m/s, Young's modulus is 6 GPa, Compression strength is 60 MPa and Tension strength is 6 MPa, respectively. The horizontal, σ_h, and vertical, σ_v, in situ stresses are 1.0 and 2.0 MPa, respectively. After drilling a hole, a pressure of 1 MPa is applied in the hole. The heterogeneity is introduced into the model by assigning E, σ_c randomly to each element within the model by assuming a Weibull distribution, as defined in equation:

$$\varphi(s,m) = \frac{m}{s_o}\left(\frac{s}{s_o}\right)^{m-1} \exp\left[-\left(\frac{s}{s_o}\right)^{m}\right] \tag{1}$$

where s is the element strength σ_c or elastic modulus E, and s_o is the mean strength σ_0 or elastic modulus of elements E_0. The parameter m is defined

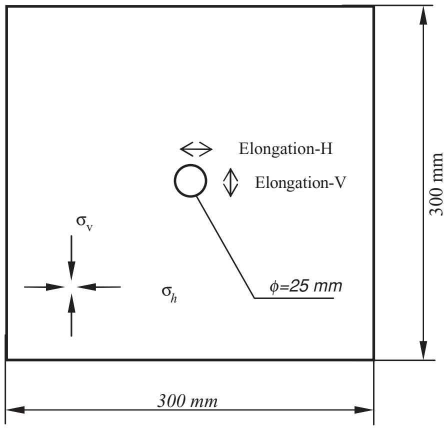


Figure 1. Sample geometry for simulating hydraulic fracturing.

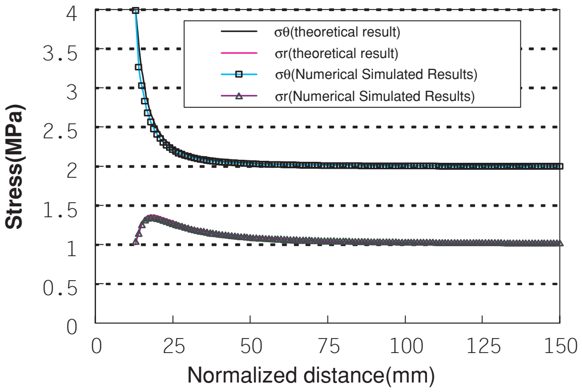


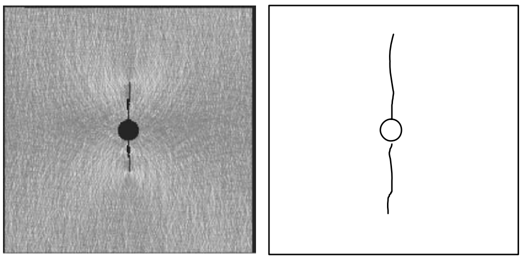
Figure 2. Stress distribution along borehole in the homogenous materials.

as homogeneity index. The homogeneous index is chosen to be 1000, that is the rock stratum is almost homogeneous. The variations of the tangential and radial stress along the horizontal diameter are plotted in Figure 2. The computed results are in good agreement with the published one and the analytical solution [2]. Having the model validated, further analyses have been carried out to study the effects the flow in fracture, the material homogeneity, the in situ stress ratio and the pore pressure field on the breakdown pressure and the development of the fractures. The findings are reported as follows

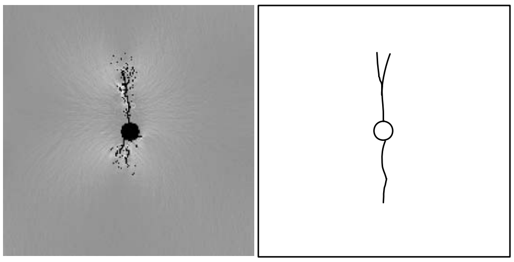
2.1 Hydraulic fracture paths with different in situ stress ratio

Further simulations have been carried out to study the effect of the σ_h/σ_v stress ratio on the fracture pattern. The stress ratio varies from 1.5 (uniform pressure) to 1.0. The fracture patterns for different stress ratio are shown in Figure 3. In the analysis, the homogeneity index is taken to be 3. At high stress ratios, that is stress ratio equal to 1.5, a pair of fractures extends in the maximum tensile stress direction from both ends of the hole (Figure 3a). The fractures are also rather straight.

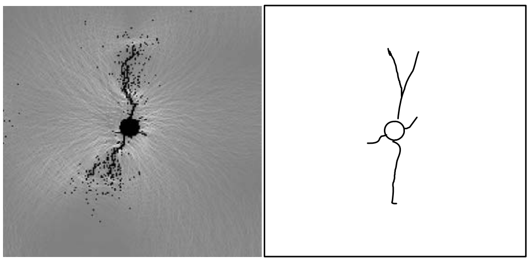
As the stress ratio decreases, the main fractures, while oriented in the maximum horizontal stress direction, are no longer straight and show a tendency to branch out along the grain boundaries. Figure 3b shows the fracture pattern for stress ratio equal to 1.25. Though cracks open at both ends of the vertical diameters, they branch out after growing for a short distance. One can also note that isolated fractures also open within the rock mass. Such fractures should represent the existence of weak elements. Figures 3c and d show the hydraulic



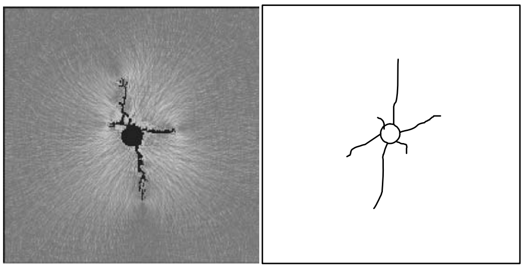
Shears stress evolution during failure Experimentally obtained result



Shear stress evolution during failure Experimentally obtained result



Shears stress evolution during failure Experimentally obtained result



Shears stress evolution during failure Experimentally obtained result

Figure 3. Experiment and numerical simulated results.

fracture path of rock models with the stress ratio equal to 1.13 and 1.0, respectively. The figures indicate multiple major traces without any preferred orientations are formed. There are significant branching and isolated fracturing. Comparison of the present results with the stress patterns published by Doe and Boyce [3] shows that the present model can predict the initiation and

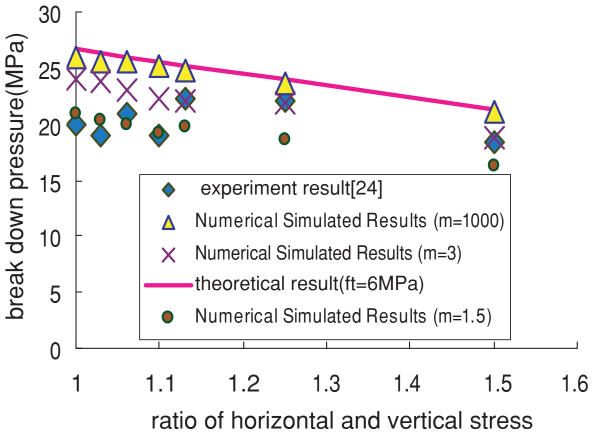


Figure 4. Breakdown pressure with different pressure ratio.

development of fractures fairly accurately. These results indicate that the crack pattern depends on the homogeneity when the stress ratio is close to one.

Due to the heterogeneity of rocks, the hydraulic fracture paths are irregular. The numerically obtained random nature of hydraulic fracture path during hydraulic fracturing and its dependence on mescoscopic homogeneity is clearly illustrated in Figure 3. As we can see in the figure, the hydraulic fracture propagation is controlled by the pre-existing field of defects.

The hydraulic fracture deterministically selects a path of least resistance through the material with statistical features, and the random location of the individual in homogeneities results in an irregular hydraulic fracture trajectory.

In Figure 4, the breakdown pressures, which is the pressure that fractures are initiated, predicted by the present model are compared to those obtained by the experiment and analytical approaches [4]. When parameter $m = 1.5$ and $m = 3$, the numerical simulation results are accord with those obtained from experiment. The results show that the present method has a more accurate prediction than the analytical ones.

3. CONCLUSIONS

The numerical simulations using F-RFPA^{2D} clearly indicates the results indicate that both the rock heterogeneity and the permeability affect fracture initiation and propagation significantly, The main findings are: Firstly, the heterogeneity of rock has a significant influence on initiation and breakdown pressure. Both fracture initiation and breakdown pressure values are much higher for the homogeneous sample than that for the heterogeneous sample. Secondly, the

F-RFPA^{2D} clearly simulates the hydraulic fracturing in heterogeneous rocks in a more realistic way than other numerical models. The capability of F-RFPA^{2D} in identifying hydraulic fracturing mechanisms, rather than prejudicing towards certain mechanisms, is obvious. In addition, the authors would like to emphasize that the work presented here is mainly intended to demonstrate a new model for coupling stress, flow and damage. Coupling with damage is the main point. We just provide a method with many parameters that you can input based on your own experience or knowledge you have in your research.

ACKNOWLEDGEMENT

The study presented in this paper is supported by the China National Natural Science No. 50490274 and No. 50204003). It is also partially supported by the Hong Kong Research Grants Council (No. HKU7029/02E).

REFERENCES

1. C.A. Tang, L.G. Tham, P.K.K. Lee and T.H. Yang (2002), Coupled analysis of flow, stress and damage (FSD) in rock failure. *International Journal of Rock Mechanics and Mining Sciences*, 39, pp. 477–489.
2. P.A. Charlez (1991), *Rock Mechanics* (*II: Petroleum applications*). Technical Publisher, Paris.
3. T.W. Doe and G. Boyce (1989), Orientation of hydraulic fractures in salt under hydrostatic and non-hydrostatic stress. *International Journal of Rock Mechanics and Sciences and Geomechanics*, 26, 6, pp. 605–611.
4. M.S. Bruno (1991), Pore pressure influence on tensile fracture propagation in sedimenta. *International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts*, 28, 4, pp. 261–273.

A NEW NUMERICAL APPROACH FOR STUDYING SELF-ORGANIZED CRITICALITY BEHAVIOUR IN ROCK FAILURE PROCESS

Y.H. Li[1], H.Q. Zhang[1], C.A. Tang[1], Q.L., Yu[1], and Z.Y. Zhao[2]
[1]*Centre for Rock Instability and Seismicity Research, Northeastern University, Shenyang City, Liaoning Province, PR China 116622*
[2]*School of Civil and Environmental Engineering, Nanyang Technological University, Singapore*

Abstract Self-Organization in rock failure progress is simulated by introducing a new numerical approach, rock failure progress analysis (RPPA^{2D}). Through the analysis of acoustic emission (AE) event time series and the frequency distribution of damage group size by correlation function and rescaled range (R/S) analysis method, it is found that the frequency distribution of damage group size complies with a power law (fractal geometry configuration) and the time series of AE event exhibits the similar scale-invariant properties and temporal long-range correlation. These fractal geometry configuration properties and long-range correlations are two fingerprints of self-organized criticality, which denonstrates the occurrence of SOC.

Keywords: self-organized criticality, power law distribution, long-range correlation, numerical simulation.

1. INTRODUCTION

Since the concept of self-organized criticality (SOC) was introduced by Bak, Tang and Wiesenfeld [1], an increasing number of physical situations showing this behaviour have been studied. The sand-pile model became one of prototype abstract models exhibiting self-organized criticality. The critical state is reached asymptotically in the limit of infinitely slow driving [2]. Due to inertia effects, the experiments on real sand-piles did not confirm SOC behaviour [3]. Its variants consisting of a cellular automaton and fully

G. R. Liu et al. (eds.), Computational Methods, 1841–1846.

deterministic versions were also studied, showing periodic or self-similar but non-random behaviour [4,5]. Since the critical point is somewhat similar to the invariable point of renormalization group calculations, self-organized criticality is often addressed using the tool of renormalization group. But it is hard to see how they relate to each other [6]. A closely related puzzle is whether criticality and self-organized criticality are compatible. In analogy with usual critical systems, the behaviour of SOC systems is often characterized by fractal geometry configuration and spatio-temporal long-range correlations. In this study, we try to explain the intriguing self-organized criticality behaviour in heterogeneous materials by introducing a new numerical method, rock failure process analysis (RFPA^{2D}), which has been developed at Centre for Rock Instability and Seismicity Research, Northeastern University, People's Republic of China.

2. BRIEF INTRODUCTION OF RFPA^{2D} AND SIMULATED MECHANISM OF SOC

RFPA^{2D} [7] is relied on FEM to perform stress analysis. The model is firstly discretized into a large number of small elements. On account of the heterogeneity of rock-like materials, the element local mechanical parameters are assumed to follow Weibull's distribution. During the simulation, the model is loaded in a displacement control mode. At each loading increment, the stress and strain in the elements are calculated, after which the stress field is examined and those elements which satisfy with the failure criterion, a revised-Coulomb criterion [8], get damaged. In RFPA^{2D}, breaking an element means reducing the element's the stiffness and strength. Then the model with new parameters switches to a new equilibrium. By this means, the stress of failed element will transfer to its neighbouring element. Then element failure is judged again. Consequently RFPA^{2D} may possibly perform many calculation circulations (stress and failure analysis) in the same loading step. The next load increment is added only when there are no more elements fail. In the same load increment, the subsequent element damage results from the stress transference of previously failed elements. In the whole loading process, there must exist one critical point at which a minimal increment is applied on the specimen, one element failure will trigger such a violent avalanche that it can cut through the whole rock specimen (long-range correlations). As no outside disturbance, the system self-organizes into a critical state. So RFPA^{2D} exhibits the avalanche behaviour in rock failure progress quite vividly and it explains why the avalanche occurs very clearly. The more the times of calculation circulations is, the more violent the avalanche behaviour is.

Table 1. Material mechanical property.

Parameter	Value
Homogeneity index (m)	4
Poisson's ratio (μ)	0.25
Elastic modulus (E_u)	60,000 (MPa)
Compressive strength (σ_c)	200 (MPa)
Friction angle (ϕ)	30°
Tension cut-off	10%

3. PROBLEM DESCRIPTION

The geometry of rock specimen is 120 mm × 80 mm, which is discretized into 120 × 80 elements. The specimen is under uniaxial compressive condition and an increasing displacement with 0.001 mm/step is applied in the vertical direction. So the stress condition is closer to plane stress. The element local mechanical parameters are listed in Table 1.

4. SOC BEBAVIOUR OF AE EVENTS IN ROCK FAILURE PROCESS

Besides stress analysis, RFPA^{2D} can simulate seismicities associated with progressive fracture process by recording failed elements and estimate the size of damage group by computing the counts of neighbouring failed elements. The shear stress distribution (above) and the location of AE events occurring during the whole failure process (below) are showed in Figure 1. The relationship of stress (σ), AE counts and loading step are showed in Figure 2. As seen the symbol marked in Figure 2, the point 'A, B, C and D' are yield strength point, maximum compressive strength point, long-time strength point and rock break point, respectively. The whole failure process is divided into five stages: (1) O–A, linear elastic deformation; (2) A–B, non-linear elastic deformation; (3) B–C, strain weakening; (4) C–D, sudden breaking: (5) D-, residual deformation. Figure. 3 shows us the frequency distribution of the size of damage group at the 98th loading step (see the point 'D' in Figure 2). According to the definition of RFPA^{2D}, two neighbouring failed elements belong to the same damage group. We can see from Figure 3 that the biggest damage group comprises 88 elements, which indicates long-range correlation occurs. It is clear that the frequency distribution of damage group turns into a power law. These data satisfy with an equation $y = 51.453 \times^{-1.0775}$. This power law distribution reflects the spatial fractal gecometry configuration; it can be regarded as spatial fingerprints for SOC. Figure 4 shows us the analytical

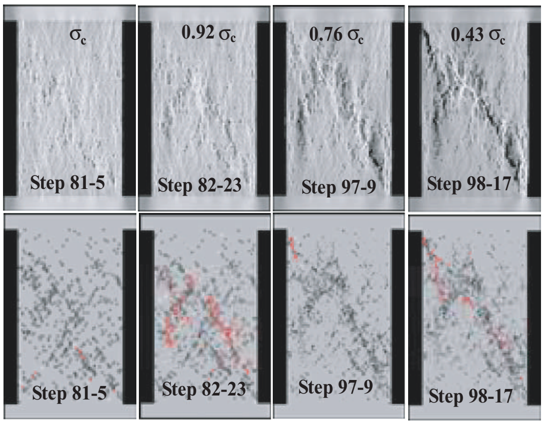


Figure 1. Stress field (above) and AE events (below) in fracture process. Step 98-17: the 17th calculation circulation in step 98.

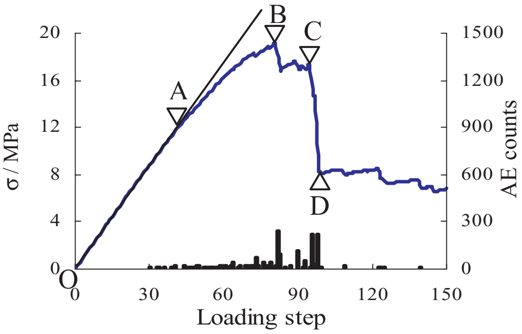


Figure 2. The relationship between stress (σ), AE counts and loading step.

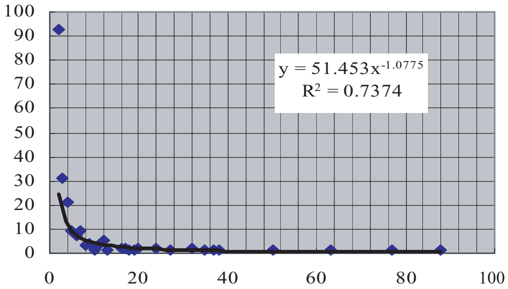


Figure 3. The frequency (N) distribution of damage group size (S) at the 98th loading step.

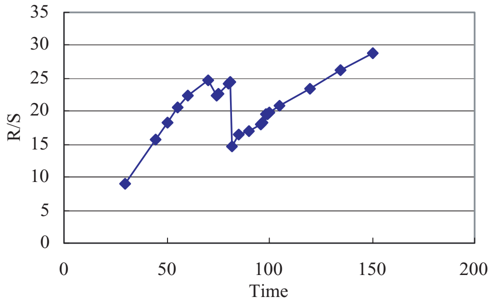


Figure 4. The persistence of AE event time series through rescaled range (R/S) analysis.

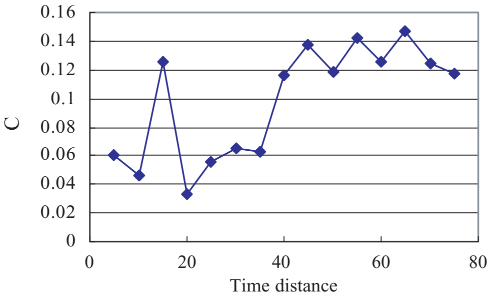


Figure 5. Absolute value (C) of autocorrelation coefficient of AE event time series obtained by correlation function.

results gained by rescaled range (R/S) analysis method, which can evaluate the persistence/antipersistence of AE event time series. A detailed description for the R/S method is presented in the document by Hurst [9]. It can be seen that the last part of the curve in Figure 4 is almost fitted with a line, especially for the part from step 98 to step 150 which exhibits the similar scale-invariant properties of AE event time series. Figure 5 shows the analytical results obtained by the correlation function, which can evaluate the autocorrelation of the AE event time series. A detailed introduction for the correlation function, please refer to the literature [10]. It can be seen from Figure 5 that the absolute value of autocorrelation coefficient enhances to a relatively stable good standard with the inerease of time delay, which reflects the temporal long-range correlation of AE event time series. The similar scale-invariant properties and long-range correlation of AE event time series are the temporal fingerprints for SOC.

5. CONCLUSIONS

Self-organization in rock failure progress is simulated by introducing a new numerial approach. RPPA^{2D}. Through the analysis of AE event time series and the frequency distribution of damage group size by correlation function and R/S analysis method, it is found that the frequency distribution of damage group size of complies with a power law (fractal geometry configuration) and the AE event time series exhibits the similar scale-invariant properties and temporal long-range correlation. These long-range correlations and fractal geometry configuration properties are two figerprints of self-organized criticality, which demonstrates the occurrence of SOC.

ACKNOWLEDGEMENTS

This work is funded by NSFC (No. 50374020, 50134040, 50174013).

REFERENCES

1. P. Bak, C. Tang and K. Wiesenfield (1987), Self-organized criticality: an explanation of 1/f noise. *Physical Review Letters*, 59, pp. 381–384.
2. C. Guido, D.D.T. Francesco and P. Albert (1996), Self-organization and annealed disorder in a fracturing process. *Physical Review Letters*, 77, pp. 2503–2506.
3. J. Rajchenbach (2004), Some remarks on the rheology of dense granular flows: a commentary on 'on dense granular flows' by GDR MiDi. *The European Physical Journal E*, 14, pp. 367–371.
4. S.S. Manna and H.J. Herrmann (2000), Intermittent granular flow and clogging with internal avalanches. *The European Physical Journal E*, pp. 341–344.
5. Lübeck, Rajewsky and Wolf (2000),
6. Y. Huang, H. Saleur, C.G. Sammis and D. Sornette (2002), Precursors, aftershocks, criticality and self-organized criticality. *Pure and Applied Geophysics*, 155, pp. 509–515.
7. C.A. Tang (1997), Numerical simulation of progressive rock failure and associated seismicity. *International Journal of Rock Mechanics and Mining Science*. 34, pp. 249–262.
8. B.H.G. Brady and E.T. Brown (1993), *Rock Mechanics*. Chapman & Hall Press, London, UK.
9. E.H. Hurst (1956), Method of using long-term storage in reservoirs. *Proceedings of the Institute Civil Engineers*, 5, pp. 519–xxx.
10. K. Ludek (2002), Correlation functions of random media. *Pure and Applied Geophysics*, 159, pp. 1811–1831.

NUMERICAL APPROACH TO MINING INDUCED INSTANTANEOUS OUTBURSTS

T. Xu[1], C.A. Tang[2], L.C. Li[1] and Z.Z. Liang[2]
[1]*Research Center for Numerical Tests on Material Failure, Dalian University, Dalian 116622, China*
[2]*Department of Engineering Mechanic, Dalian 116024, China*

Abstract Instantaneous outburst in underground coal mines has been the cause of major disasters in the world mining industry and has plagued underground mining operations of gassy coal seams in many countries for over a century. In this paper, a numerical quantitative model for the coupled gas flow in coal and rock failure is proposed to simulate and visualize mining induced instantaneous outbursts and associated stress fields. Numerical simulations on instantaneous outburst reveal that using the proposed numerical model to investigate the instantaneous outbursts and attempt to gain insight into the mechanisms of outbursts is appropriate and practical.

Keywords: instantaneous outburst, gas flow, numerical simulation, RFPA^{2D}-flow.

1. INTRODUCTION

Instantaneous outburst is a sudden and usually unexpected expulsion of coal and gas away from a freshly exposed coal face during underground mining or drilling in coal seams. It has been the cause of major disasters in the world mining industry. In the last 150 years since the first reported outburst occurred in France in 1843, as many as 30,000 outbursts have occurred in the world. This has forced mining industry to develop techniques to understand this complex phenomenon, and procedures to minimize the effect of outbursts. As of today, despite extensive research into outbursts occurred in coal mines, surprisingly little progress has been made in the past decades in understanding the complex phenomena. Especially, a quantitative numerical model that describes

G. R. Liu et al. (eds.), Computational Methods, 1847–1851.

progressive failure as well as instantaneous outbursts has not appeared. It is the aim of this paper to present such a model and to show how the model replicates instantaneous outbursts associated with the observations.

2. OUTLINE OF RFPA^{2D}-FLOW MODEL

In this paper, based on the Rock Failure Process Analysis code [1], an improved numerical model, rock failure process analysis for fluid flow, was developed to simulate mining induced instantaneous coal and gas outbursts and try to gain an insight into the physical and mechanical mechanisms and essence of instantaneous outbursts. Here, we mainly focus on the coupled gas flow in coal and rock, and the details of the model on stress and failure description can refer to the related publications [1, 2].

The fundamental assumption behind the outburst model presented here is that the gas flow through coal and rock at a meso-scopic level follows linear filtration law [3]. Hereby, the equation of the isothermal infiltration gas flow in gassy coal and rock based on seepage theory can be obtained [2]

$$\alpha_p \nabla^2 P = \frac{\partial P}{\partial t} \tag{1}$$

where, $\alpha p = 4\lambda A^{-1} P^{\frac{3}{4}}$, λ is the coefficient of gas infiltration in $m^2/(MPa^2\ S)$, A is the empirical coefficient of gas content in $m^3/(m^3\ MPa^{1/2})$ and P is the square of gas pressure in MPa^2. For the deformation of coal and rock, the governing equation considering the gas pressure in coal/rock can be represented as [2]

$$(K + G)u_{j,ji} + Gu_{i,jj} + f_i + (\alpha p)_{,i} = 0 \tag{2}$$

where G is shear modulus, K is Lame's constant, u is the displacement of element, f_i is the body force, α is the coefficient of gas pore pressure and p is gas pressure in MPa. In addition, for damage induced gas permeability change, it has been found that the microcracking induced damage causes a remarkable increase in permeability of rock [4]. The gas permeability coefficient in stressed rock can be described [1]

$$\lambda = \xi \lambda_0 e^{-\beta(\sigma - \alpha p)} \tag{3}$$

where λ_0 is the initial gas permeability for unloaded coal and rock, β is the coupling factor of stress to pore pressure and ξ is the coefficient of sudden jump of gas permeability for stressed elements. Thus, the above governing equations were incorporated into the model to simulate the mining induced instantaneous outbursts.

Table 1. Numerical model parameters.

Index	Coal	Rock
m	1.5	5
E_0/GPa	5	50
σ_0/MPa	100	300
Possion's ratio, μ	0.3	0.25
λ, m^2/(MPa^2 S)	$1.2 * 10^{-5}$	$1.2 * 10^{-8}$
A, m^3/(m^3 MPa$^{1/2}$)	2	0.1
α	0.5	0.01
β	0.2	0.1

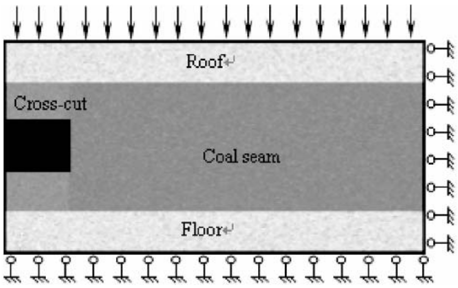


Figure 1. Numerical model.

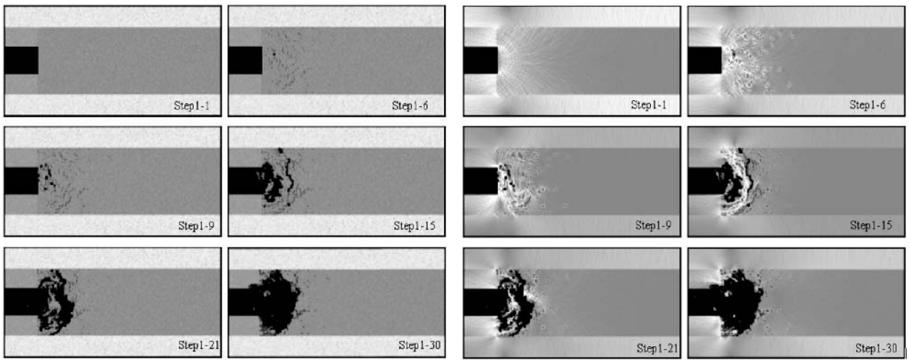
3. MODELING OF INSTANTANEOUS OUTBURSTS

3.1 Model set-up

A numerical model for an instantaneous outburst induced by cross-cut driving was performed in this paper. The numerical model with a dimension of 5 m × 10 m was discretized into 100 × 200 elements, as shown in Figure 1. The gas pressures on the coal face and in the deep coal seam are 0.1 and 4 MPa, respectively. The roof and floor of the coal seam are presumed to be impermeable, thus the seepage boundary conditions is set with no seepage flux. The model parameters are given in Table 1.

3.2 Modeling results

Figure 2 shows the numerical results of the entire outbursts process induced by cross-cut driving and the correspondingly associated evolution of shear stress fields in the coal seam and the roof and floor of strata. It can be seen from Figure 2, the cross-cut driving induced instantaneous outbursts can



(a) **(b)**

Figure 2. Simulated delay outburst (a) and associated shear stress fields (b).

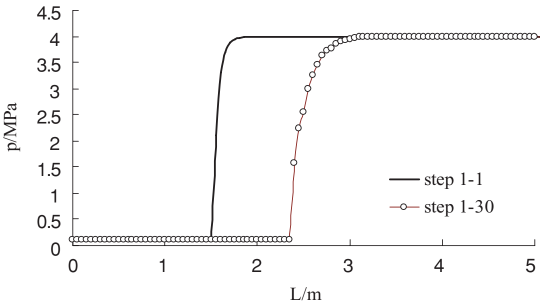


Figure 3. Changing curves of gas pressure pre- and post-instantaneous outbursts.

be subdivided into four major stages, that is, stress concentration stage, coal and rock fracturing and splitting induced by the combinations of gas pressure and abutment stress, cracks propagation driven by gas pressure and the abrupt ejection of coal and gas.

The changing curve of gas pressure pre- and post-outbursts is given in Figure 3. As seen from Figure 3, the gas pressure is on average at hydraulic pressure prior to cross-cut driving. During cross-cutting, mining removes the confining pressure on the coal face; nevertheless the gas pressure remains high even though the pores in the coal mass may expand some due to reduction in coal and rock stress. Thus the coal near working face sustain high gas pressure gradient and clusters of microcracks begin to form when the high gas pressure is beyond the tensile strength of coal. Meanwhile, the abutment stress acting on the coal face after mining excavations further deteriorate the properties of the coal and the coal parallel to free exposed face further fractured and split. The combinations of gas pressure in the coal front and the abutment stress on the coal face crushed and even pulverized the weak coal. Thus, the crushed and pulverized coal was abruptly ejected and a circular cross-section outburst cavity can be observed from Figure 2. With the continuance of outbursts process, the gas pressure progressively declines near the freshly exposed coal. The outbursts process ceased and an elongated hemispherical outburst cavity can be seen in the end.

4. CONCLUDING REMARKS

In this paper, the gas pressure incorporated Rock Failure Process Analysis for fluid flow code (RFPA^{2D}-Flow) on the basis of the theory of solid deformation and gas flow in porous media. Applying the RFPA^{2D}-Flow code, the numerical tests on cross-cut driving induced instantaneous outbursts were performed. The numerical simulations replicated the cross-cut driving induced outbursts process and the associated evolution of stress fields in coal seam and rock strata, and reveal that instantaneous outbursts generally experience four stages, i.e., stress concentration, coal and rock microcracking, cracks propagation and the ejection of coal and gas. Moreover, the numerical results not only trace the initiation, propagation and coalescence of cracks in coals, but also visualize the associated evolution of stress fields in rock formations and the observations in laboratory and in field, such as the hemispherical outburst cavity with nearly symmetrical dimensions in the coal seam. Thus, the information of this paper can provide a better understanding of the mining induced instantaneous outbursts and gain insight into the complex mechanisms.

ACKNOWLEDGEMENTS

The work was funded by NSFC (50374020, 50504003 and 50504005).

REFERENCES

1. C.A. Tang, L.G. Tham, P.K.K. Lee, T.H. Yang and L.C. Li (2002), Coupled analysis of flow, stress and damage (FSD) in rock failure. *International Journal of Rock Mechanics and Mining Science,* 39, pp. 477–489.
2. T. Xu (2004), Numerical tests on coupled gas flow in coal and rock failure. Ph.D. thesis, Northeastern University, Shenyang, China, pp. 1–164.
3. S.N. Zhou and B.Q. Lin (1998), *The Theory of Gas Flow and Storage in Coal Seams*. China Coal Industry Publishing Corporation, Beijing, pp. 69–70.
4. T.H. Yang, T. Xu, C.A. Tang and Q.Y. Feng (2003), Experimental investigation of permeability evolution of brittle rock in failure process. *Journal of Northeastern University,* 24, 10, pp. 919–923.

3D NUMERICAL SIMULATION OF A LARGE SPAN DOUBLE-ARCH TUNNEL CONSTRUCTION

C.C. Xia, H.J. Liu and H.H. Zhu
Geotechnical Department, Tongji University, Shanghai 200092, PR China

Abstract This paper presents a numerical simulation of double-arch tunnel excavation process by means of elasto-plastic finite element method. To perform a fully 3D numerical analysis of the whole excavating and supporting process, a FEM model including the surrounding rock and supporting structure was developed, and the whole process is divided into 20 construction steps according to the actual construction sequence. The results were satisfactory and were reasonably close to the displacement data monitored in the construction field. In the second part of the article, the redistribution of stress in shotcrete, lining and anchor against each excavation step was studied to understand the interaction between rock and the supporting structure. Furthermore, the load pattern in surrounding rock and the supporting structure were studied to analyze the influence of excavation in adjacent tunnel and find out the reasonable distance between adjacent excavation surfaces. The results obtained provide a better understanding of excavation process of double-arch tunnel, and is helpful for guiding design and construction.

Keywords: 3D numerical simulation, construction mechanics, double-arch tunnel.

1. INTRODUCTION

Due to its flexibility to adapt to different soil conditions and geological environment and its simplicity of equipment for tunnelling, the New Austrian Tunnelling Method (NATM) is regarded as the guiding philosophy for tunnel design and construction and is widely used in many countries.

In principle of NATM, the surrounding rock is considered an active part of the supporting structures, capable of sustaining the load due to excavation. Meanwhile the tunnel lining should be installed in an optimized way, capable of providing enough flexibility to allow a desirable degree rock deformation and

G. R. Liu et al. (eds.), Computational Methods, 1853–1858.

enough strength to control excessive deformation. The last and most important principle is that the excavation method including complimentary construction artifices should be adopted according to observational or measured data and geological condition during construction, so it is said to be observational and adjustable [1]. But the law of deformation and stress in surrounding rock and supporting structure against each excavation step is still not clear. Systematic understanding of the repeated loading and unloading process in the rock mass and its impact on the supporting structures still demand further research. Current research work on the complicated process of tunnelling mainly depend on the development of numerical model, most of which were carried out by means of finite element method (FEM) method [2]. According to a compilation of the documents, majority of the numerical simulation were still performed under 2D plane strain approach [3], but it has been widely recognized that the tunnel excavation process induces typically 3D stress effect, which can only be accurately described in 3D FEM model.

Double-arch tunnel is a special type of tunnel whose construction process is more complicate than that of single tunnel. Studies on the principle of its construction mechanics were mainly carried out by means of constructional monitoring or by local experience [4], and few numerical simulations have been carried out.

2. DESCRIPTION OF THE 3D FEM MODEL

The tunnel simulated here is a three-lane double-arch highway tunnel, and it is 87 m long, 26.5 m wide and 7.76m high. It is a shallow tunnel whose maximum rock cover is only 13 m.

Partial-face construction method was adopted in the construction and it was a complicated process with tens of steps and sub-steps. For simplicity, the whole construction process was reduced into 20 separate stages, details of which are not given here for the limit of the article. Planar schematic illustration of the reduced excavation sequences is shown in Figure 1.

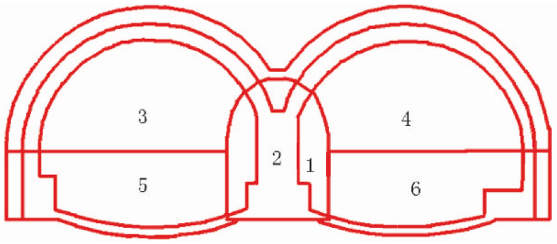


Figure 1. Schematic representation of excavation sequences.

Table 1. Mechanical parameters.

Material	E (MPa)	ν	γ (kN)	C (MPa)	φ (°)
Rock	250	0.35	20	0.1	35
Concrete	2.8e4	0.18	25	1.0	55
Anchor	2.1e5	0.3	78		

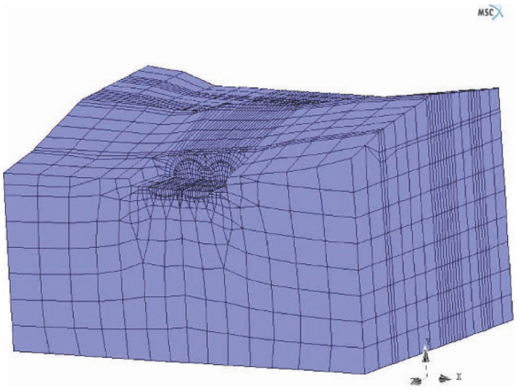


Figure 2. 3-D finite element mesh.

The rock encountered is primarily weak and soft rock, which was sensitive to blasting and water. In the numerical model, a linear Mohr–Coulomb material model was adopted for the rock, shotcrete and lining, and a linear-elastic model for the anchor. The mechanical parameters for the rock and the structures are given in Table 1.

The 3D FEM model developed consists of all the supporting structures and the surrounding rock, and it covers 107 m along the longitude axis of the tunnel, 67 m along the horizontal direction (perpendicular to the longitudinal axis), and about 80 m along the vertical direction (perpendicular to the horizontal plane). Eight-node hexahedral elements were introduced to simulate the surrounding rock and the lining, the middle wall and the invert, meanwhile four-node quadrangular thin shell elements and two-node beam elements were introduced to simulate the shotcrete and the anchor, respectively. Altogether 15,597 nodes and 20,212 elements are meshed and the element mesh of whole model is illustrated in Figure 2. Detail geometrical information is not given for the limit of the paper.

3. SUMMARY OF NUMERICAL RESULTS

According to experience, the most instructive information is that of deformation monitored, because the pressure cells do not always give reasonable results, especially when they monitor radial stress component. The results of

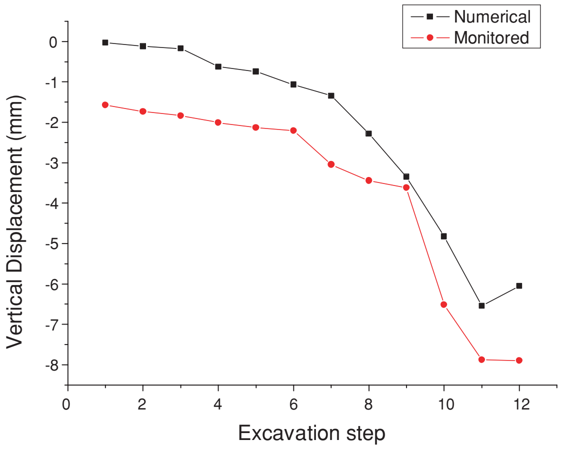


Figure 3. Comparison of the numerical and monitored result.

the numerical analysis were evaluated by comparing with the corresponding data monitored. The value of vertical displacement computed is compared with that obtained by the multiple-position borehole extensometers in the rock.

For the limit of the article, only the point above the left roof of the tunnel is selected, the evolution of its vertical displacement is shown in Figure 3. It can be seen that the numerical results are reasonably close to the corresponding data monitored in the construction field.

By surveying the distribution pattern and value of the stress, it can be seen that stress in the shotcrete, the lining, the invert and the anchors change considerably during the construction process.

As depicted in Figure 4, the value of compressive stress in the shotcrete roof decreases with drawing of the left excavation face, which is unfavourable for the stability of the tunnel and it increases gradually with the left excavation face advanced away. This indicates that the load caused by excavation was first taken up by the shotcrete and its adjacent rock, and later part of the load went into the surrounding rock gradually through the process of stress transfer. At last, the load is taken up by the rock and supporting structure together. The law of the stress change in the anchor, depicted in Figure 5, confirmed such proposition. Meanwhile the value of stress in the middle wall is considerably high and it changed more violent than other structures, so more attention is required.

4. IMPLICATIONS TO DESIGN AND CONSTRUCTION

A series of 3D numerical analysis were performed using finite element with Mohr–Coulomb model to simulate the double-arch tunnel construction in the principle of NATM. The 3D construction effect of double-arch tunnel can be

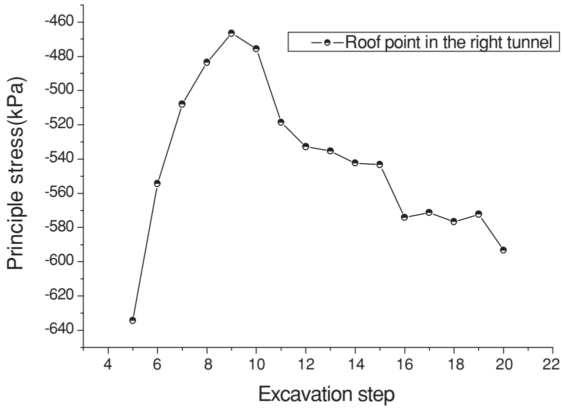


Figure 4. Principle stress in shotcrete roof.

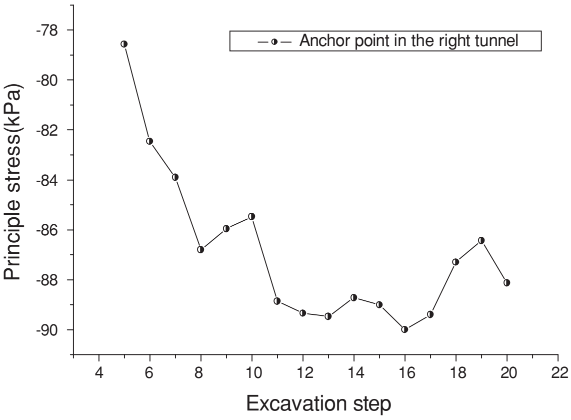


Figure 5. Principle stress in anchor blot.

represented by 3D numerical analysis and the result is satisfactory. The repeated loading and unloading process in the supporting structures and the surrounding rock is the effect of all construction activities in both tunnels. It is an interactive construction process and should and can be analyzed systematically by 3D numerical model.

The mechanical behaviour of the first tunnel construction is similar to that of single tunnel excavation except for the middle wall, but it is affected greatly by the construction activities in the second tunnel whose excavating face lagged. The distance between the two excavation faces is the key factor to the safety of the double-arch tunnel and its displacement control, meanwhile the span between the excavation face of the heading and the bench in both tunnels is also important. Insight into the construction effect and determination of the optimal distance between these excavation faces still demand further research.

The value of stress in the middle wall is considerably high and varies violently, and more attention is required. The construction of the lining and the invert is also of great importance to the ultimate value of the displacement and stress in the surrounding rock.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support of Communication Ministry of PR China.

REFERENCES

1. M.M. de Farias, A.H. Moraes Jr and A.P. de Assis (2004), Displacement control in tunnels excavated by the NATM: 3-D numerical simulations. *Tunneling and Underground Space Technology*, 19, pp. 283–293.
2. W.S. Zhu, S.C. Li and S.C. Li (2003), Systematic numerical simulation of rock tunnel stability considering different rock conditions and construction effects. *Tunneling and Underground Space Technology*, 18, pp. 531–536.
3. B.I.P. Queiroz and A.e. Negro (2000), Prediction and performance of soft ground tunnels. *Geotechnical Aspects of Underground Construction in Soft Ground*, pp. 409–418.
4. K.F. Bizjak and B. Petkovsek (2004), Displacement analysis of tunnel support in soft rock around a shallow highway tunnel at Golovec. *Engineering Geology*, 75, pp. 89–106.

AMPLITUDE FREQUENCY—LOAD CHARACTERISTIC RELATION OF CIRCULAR SANDWICH PLATES

Du Guojun and Hu Yuda
Yanshan University, Qinhuangdao 066004, P.R.China

Abstract The differential equations of the axisymmetric large amplitude free vibration for circular sandwich plates under static load are derived, and a set of nonlinearly coupled algebraic and differential eigenvalue equations of the problem are formulated following an assumed time mode approach suggested. The analytic solutions are presented and a relation for amplitude frequency-load of the plates with the edge clamped is derived by modified iteration method. The effects of static load on vibrations of plates are investigated.

Keywords: circular sandwich plate, uniformed load, nonlinear, amplitude frequency-load characteristic relation

1. INTRODUCTION

So far, only a few peoples have studied the large deflection problem of sandwich plates and shells, because off the difficulty of nonlinear mathematics. Liu Renhuai has done much to find a series of results with the value of application in engineering practice [1–5]. Du Guojun [6,7] have done the initial discussion to the large amplitude vibration of circular sandwich plate. But little has been done to study the effects of static load on the vibration of circular thin plate under a centre force, obtained the characteristic relation of fundamental frequency-load, based on the assumptions of space and time mode. Wang Jinying et al. [9] studied the nonlinear vibration of the flexure circular plate with initial deflection, deriving the control equation of the time mode by Galerkin method, based on the assumption of space mode, and the periodically solution of the nonlinear vibration is obtained by Lindstedt–Poincare perturbation. This

G. R. Liu et al. (eds.), Computational Methods, 1859–1872.

study is not perfectly because of the membrane stress under the static load is neglected.

In this paper, we studied the nonlinear vibration problem of circular sandwich plate under static load by an analysis method, the accurate solution is employed for the static problem. Based on the assumption of time mode and the method of calculus of variations, the deflection and stress function employs the simple harmonic function, the space function is unknown, the vibration mode assumed contains two unknown natural parameters, that is natural frequency and the nonlinear vibration 'drift' is brought by the static deformation [10,11], they are related to the loads, boundary conditions and the vibration amplitude. Substitution of the assumptive model function to the variations equation of this problem, the control equation of space mode and the algebraic equation for solving the 'drift' are derived, and the expression of the space mode are given by modify iteration method, thus the amplitude frequency-load characteristic relations of circular sandwich plate under the uniform load are derived, the numeric results are given and the effects of static load to the nonlinear vibration frequency and 'drift' are discussed, the method presented in this paper can be applied to the problem of the annular plates and the shells.

2. FUNDAMENTAL EQUATIONS

The nondimensional nonlinear static equation of circular sandwich plate under the uniformed load have been given in paper [1]

$$L(\rho(K\bar{S}_r^0 + 1)\phi^0) - \bar{S}_r^0\phi^0 - P_0\rho = 0 \tag{1}$$

$$L(\rho^2\bar{S}_r^0) + (\phi^0)^2/\rho = 0 \tag{2}$$

where the nondimensional variable are given by

$$\rho = r/a,\ \bar{W}^0 = \sqrt{2(1-\nu^2)}w^0/h_0,\ \phi^0 = \bar{W}^0_\rho$$

$$\bar{S}_r^0 = \frac{2h_1a^2}{D}\sigma_r^0,\ S_\theta^0 = \frac{2h_1a^2}{D}\sigma_\theta^0,\ K = \frac{D}{G_2h_0a^2},\ P_0 = \frac{\sqrt{2(1-\nu^2)}}{2h_0D}q_0 \tag{3}$$

where a is the radius of the circular sandwich plate, $r$is the radial coordinate, h_1 is the thick of the face plates, h_0 is the distance between the medium section of the top and the bottom face plates and q_0 is the density of the uniformed load, the detail note see paper [1], and $L(\cdots) = ((\cdots)_{,\rho}/\rho)_{,\rho}$.

The clamped boundary conditions are

$$\text{when } \rho = 1 : \bar{W}^0 = 0, (K\bar{S}_r^0 + 1)\phi^0 + KP_0 = 0, (\rho\bar{S}_r^0)_{,\rho} - \nu\bar{S}_r^0 = 0 \tag{4}$$

$$\text{when } \rho = 0 : (K\bar{S}_r^0 + 1)\phi^0 = 0, \bar{S}_r^0 < \infty \tag{5}$$

We study the large amplitude vibration under the static deformation, the variation formulation of the vibration differential equations are derived by Hamilton principle [6]

$$\int_{t1}^{t2}\int_{0}^{a}(rmw_{,tt}-2h_1(r\sigma_{r0}w_{,r})_{,r}-G_2h_0(r(\psi+w_{,r}))_{,r})\delta w\cdot 2\pi drdt=0 \tag{6}$$

$$D((r\psi)_{,r}/r)_{,r}-G_2h_0(\psi+w_{,r})=0 \tag{7}$$

The stress-displacement relation is given by

$$((r^2\sigma_r)_{,r}/r)_{,r}+Ew_{,r}^2/2r=0 \tag{8}$$

assumption w_0 is the deflection of the circular sandwich plate under the static load P_0, we study the free vibration base on the static deformation,$\overline{w}$ is the dynamic deflection, the total deflection w can be written as

$$w=w_0+\overline{w} \tag{9}$$

therefore, we obtain

$$\sigma_r=\sigma_r^0+\overline{\sigma}_r \tag{10}$$

$$\psi=\psi_0+\overline{\psi} \tag{11}$$

substitution of Equations (9)–(11) into (6)–(8), and the solutions are written as [10,11]

$$\begin{gathered}\bar{w}=\bar{W}(r)(\xi+\cos\omega t),\ \bar{\psi}=\bar{\phi}(r)(\xi+\cos\omega t)\\ \bar{\sigma}_r=\bar{S}(r)(\xi+\cos\omega t)+\bar{T}(r)(\xi+\cos\omega t)^2\end{gathered} \tag{12}$$

where ξ is a unknown variable, it is the nonlinear vibration 'drift' is brought by the static deformation that causing the plate has the different draw and press stiffen, ω is the nondimensional vibration frequency of thee large amplitude in the static configuration , $\bar{W}(r)$, $\bar{\phi}(r)$, $\bar{S}(r)$, $\bar{T}(r)$ are the unknown space functions, by a series of mathematics calculation, we obtain

$$(\tilde{L}^2+K\omega^2\tilde{L}-\omega^2)W=-KL^*(f(\rho))+f_{,\rho}(\rho)/\rho \tag{13}$$

$$a_0+a_1\xi+a_2\xi^2+a_3\xi^3=0 \tag{14}$$

where

$$\tilde{L}(\cdots)=(\rho(\cdots)_{,\rho})_{,\rho}/\rho,\ L^*=(\rho((\cdots)_{,\rho}/\rho)_{,\rho})_{,\rho}/\rho,$$

$$f(\rho)=\rho S_r^0W_{,\rho}+\rho SW_{,\rho}^0+f_1(\xi)\rho(W_{,\rho}^0T+W_{,\rho}S)+f_2(\xi)\rho W_{,\rho}T$$

$$f_1(\xi)=\frac{2\xi^3+3\xi}{2\xi^2+1},\quad f_2(\xi)=\frac{8\xi^4+24\xi^2+3}{8\xi^2+4}$$

and

$$
\begin{aligned}
a_0 &= -\int_0^1 (\rho W^0_{,\rho} T + \rho W_{,\rho} S)_{,\rho} W d\rho \\
a_1 &= -\int_0^1 \left(\rho \tilde{L}^2(W) - (\rho N^0_r W_{,\rho} + \rho S W^0_{,\rho} + \frac{3}{2}\rho T W_{,\rho})_{,\rho}\right) W d\rho \\
&\quad + \int_0^1 K(\rho((f(\rho))_{,\rho}/\rho)_{,\rho})_{,\rho} W d\rho + \frac{K\omega^2}{2\xi^2+1}\int_0^1 (\rho W_{,\rho})_{,\rho} W d\rho \qquad (15) \\
a_2 &= 2a_0 \\
a_3 &= -\int_0^1 (\rho T W_{,\rho})_{,\rho} W d\rho
\end{aligned}
$$

The stress-displacement relations are written as

$$L(\rho^2 S) = -2\beta W^0_{,\rho} W_{,\rho}/\rho \qquad (16)$$

$$L(\rho^2 T) = -\beta W^2_{,\rho}/\rho \qquad (17)$$

the clamped boundary condition is

$$
\begin{aligned}
&\text{when } \rho = 1\colon W = 0,\ K\omega^2 \frac{1}{\rho}\int_0^\rho \rho W d\rho + K(S^0_r \phi + S\phi^0 + S\phi) + \phi = 0, \\
&\quad (\rho S)_{,\rho} - \nu S = 0, \quad (\rho T)_{,\rho} - \nu T = 0 \qquad (18) \\
&\text{when } \rho = 0\colon W = W_0,\ K\omega^2 \frac{1}{\rho}\int_0^\rho \rho W d\rho + K(S^0_r \phi + S\phi^0 + S\phi) + \phi = 0, \\
&\quad S < \infty,\ T < \infty \qquad (19)
\end{aligned}
$$

where $W_0 = W(\rho)|_{\rho=0}$, and

$$W^+_m = W_0(1+\xi), \quad W^-_m = W_0(1-\xi)$$

W^+_m and W^-_m are the positive and negative vibration amplitude, respectively. Where, the nondimensional variable are

$$
\begin{aligned}
&\rho = r/a,\ \phi = W_{,\rho},\ W = \bar{W}/h_0,\ W^0 = w^0/h_0,\ \omega = \bar{\omega}(ma^4/D)^{\frac{1}{2}}, \\
&T = h_1 a^2 \bar{T}/D,\ S = h_1 a^2 \bar{S}/D,\ \bar{S}^0_r = h_1 a^2 \sigma^0_r/D,\ K = D/G_2 h_0 a^2, \qquad (20) \\
&\beta = 2(1-\nu^2)
\end{aligned}
$$

compared Equation (3) and (20), we obtain

$$W^0 = \bar{W}^0/\sqrt{2(1-\nu^2)},\ S^0_r = \bar{S}^0_r/2 \qquad (21)$$

3. ANALYTIC SOLUTION OF THE PROBLEM

The exact static solution can be obtain in the paper [2]

$$\bar{S}_r^0 = \sum_{i=0}^{\infty} a_{2i}\rho^{2i} \tag{22}$$

$$\phi^0 = \sum_{i=0}^{\infty} b_{2i+1}\rho^{2i+1} \tag{23}$$

where

$$\begin{aligned}
a_{2i} &= -\frac{1}{4i(i+1)}\sum_{m=0}^{i-1} b_{2m+1}b_{2i-2m-1}, (i = 1, 2, \cdots)\\
b_{2i+1} &= \frac{1}{Ka_0+1}\left(\frac{1}{4i(i+1)}\sum_{m=0}^{i-1} a_{2m}b_{2i-2m-1}\right.\\
&\quad\left. - K\sum_{m=1}^{i} a_{2m}b_{2i-2m+1}\right), (i = 2, 3, \cdots)\\
b_3 &= \frac{1}{8(Ka_0+1)}((a_0 - 8Ka_2)b_1 + P)
\end{aligned} \tag{24}$$

substitution of Equations (13) and (14) into the boundary conditions, and combining Equation (15), we obtain the nonlinear algebraic equation systems, and $a_0,\ b_1$ can be obtain to the given load P and the boundary conditions, the calculating course see paper [2]. We solve the boundary value problem (6)–(8), (10), (11) by the modify iteration method, in the first order approximate, we neglect the nonlinear terms and the terms of containing the static deformation in Equation (6), and let $\xi = 0$, we have

$$(\tilde{L}^2 + K\omega_0^2\tilde{L} - \omega_0^2)W_1 = 0 \tag{25}$$

$$L(\rho^2 S_1) = -2\beta W_{,\rho}^0 W_{1,\rho}/\rho \tag{26}$$

$$L(\rho^2 T_1) = -\beta W_{1,\rho}^2/\rho \tag{27}$$

The clamped boundary conditions are

$$\begin{aligned}
&\text{when } \rho = 1\text{: } W_1 = 0,\ K\omega_0^2\frac{1}{\rho}\int_0^{\rho} \rho W_1 d\rho + \phi_1 = 0,\\
&(\rho S_1)_{,\rho} - \nu S_1 = 0, \quad (\rho T_1)_{,\rho} - \nu T_1 = 0
\end{aligned} \tag{28a–d}$$

$$\text{when } \rho = 0: W_1 = W_0,\ K\omega_0^2\frac{1}{\rho}\int_0^{\rho} \rho W_1 d\rho + \phi_1 = 0,\ S_1 < \infty,\ T_1 < \infty \tag{29a–d}$$

the solution of Equation (25) can be written as

$$W_1 = W_0 \sum_{j=0}^{\infty} A_j^{(1)} \rho^{2j} \tag{30}$$

where

$$A_j^{(1)} = \mu_1 \frac{\omega_0^{2j}}{2^{2j}(j!)^2 c^j} + \mu_2 \frac{(-1)^j c^j}{2^{2j}(j!)^2}$$

$$c = \omega_0^2 (K + \sqrt{K^2 + 4/\omega_0^2})/2$$

where $\mu_i (i = 1, 2)$ is the coefficient which is to be determined by the boundary conditions. Substitution of Equation (30) into (28a,b) and (29a), we obtain

$$\mathrm{A}\mu = 0 \tag{31}$$

where

$$\mu = [\mu_1, \quad \mu_2, \quad 1]^{\mathrm{T}}$$

$$A = \begin{bmatrix} a_{01} & a_{02} & -1 \\ a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \end{bmatrix}$$

the expressions of elements of matrix A are omitted. Because μ is not a zero vector, the determinant of matrix A must equal zero, that is

$$\det \mathbf{A} = 0 \tag{32}$$

Solving Equation (32), we obtain ω_0, thus $a_{ij} (i = 0, 1, 2; j = 1, 2)$ is determined, μ_1 and μ_2 can also be determined in Equation (32), therefore, W_1 is determined. For determining S_1, T_1, substitution of (30) into Equations (26) and (27), doing integral, and applying the boundary conditions (28c,d) and (29c,d), we obtain

$$S^{(1)} = W_0 \sum_{j=0}^{\infty} B_j^{(1)} \rho^{2j} \tag{33}$$

$$T^{(1)} = W_0^2 \sum_{j=0}^{\infty} C_j^{(1)} \rho^{2j} \tag{34}$$

where

$$B_j^{(1)} = -\frac{\beta}{j(j+1)} \sum_{i=0}^{j} (j-1) b_{2i+1} A_{j-i}^{(1)}, \quad j = 1, 2, \cdots$$

$$
\begin{aligned}
B_0^{(1)} &= -\sum_{j=1}^{\infty} \frac{\eta(2j+1)-\nu}{\eta-\nu} B_j^{(1)} \\
C_j^{(1)} &= -\frac{\beta}{j(j+1)} \sum_{i=0}^{j} i(j-1) A_i^{(1)} A_{j-i}^{(1)}, \quad j = 1, 2, \cdots \\
C_0^{(1)} &= -\sum_{j=1}^{\infty} \frac{\eta(2j+1)-\nu}{\eta-\nu} C_j^{(1)}
\end{aligned}
\tag{35}
$$

For solving $a_0,\ a_1,\ a_2,\ a_3$, substitution of Equations (30), (33) and (34) into (14), we obtain

$$
\begin{aligned}
a_0 &= \varphi_1 W_0^3 \\
a_1 &= \varphi_2 W_0 + \varphi_3 W_0^2 + \varphi_4 W_0^3 + \frac{3}{2}\varphi_5 W_0^4 \\
a_2 &= 2\varphi_1 W_0^3 \\
a_3 &= \varphi_5 W_0^4
\end{aligned}
\tag{36}
$$

where

$$
\begin{aligned}
\varphi_1 &= -\sum_{j=0}^{\infty}\sum_{i=0}^{j} \frac{2i}{2j+1} A_i^{(1)} (E_{j-i} + F_{j-i}) \\
\varphi_2 &= \sum_{j=0}^{\infty}\sum_{i=0}^{j} \frac{8(j-i+1)^2(j-i+2)K}{2(j+1)} A_i^{(1)} (H_{j-i+1} + I_{j-i+1}) \\
\varphi_3 &= -\sum_{j=0}^{\infty}\sum_{i=0}^{j} 16(j+2)^2(j+1) A_{i+2}^{(1)} A_{j-i}^{(1)} + \sum_{j=0}^{\infty}\sum_{i=0}^{j} \frac{4i}{2j+1} A_i^{(1)} (H_{j-i} + I_{j-i}) \\
&\quad + \sum_{j=0}^{\infty}\sum_{i=0}^{j} \frac{8(j-i+1)^2(j-i+2)K}{2(j+1)} A_i^{(1)} f_1(\xi)(E_{j-i+1} + F_{j-i+1}) \\
&\quad + \frac{K\omega^2}{2\xi^2+1} \sum_{j=0}^{\infty}\sum_{i=0}^{j} \frac{2i^2}{j+1} A_i^{(1)} A_{j-i+1}^{(1)} \\
\varphi_4 &= \sum_{j=0}^{\infty}\sum_{i=0}^{j} \frac{8(j-i+1)^2(j-i+2)K}{2(j+1)} f_2(\xi) A_i^{(1)} G_{j-i+1} \\
\varphi_5 &= -\sum_{j=0}^{\infty}\sum_{i=0}^{j} \frac{4}{j+1} C_{j-i}^{(1)} G_i
\end{aligned}
\tag{37}
$$

and

$$\begin{aligned}
E_j &= \sum_{k=0}^{j} b_{2k+1} C_{j-k}^{(1)}, \quad F_j = \sum_{k=1}^{j+1} 2KA_k^{(1)} B_{j-k+1}^{(1)}, \\
H_j &= \sum_{k=0}^{j+1} 2(j-k+1) a_{2k} A_{j-k+1}^{(1)}, \quad I_j = \sum_{k=0}^{j} b_{2k+1} B_{j-k}^{(1)}, \\
G_j &= \sum_{k=1}^{j+1} k(j-k+1) A_k^{(1)} F_{j-k+1}, \quad j = 0, 1, 2, \cdots
\end{aligned} \tag{38}$$

thus, Equation (14) can be written as

$$\varphi_1 W_0^2 + (\varphi_2 + \varphi_3 W_0 + \varphi_4 W_0^2 + 1.5\varphi_5 W_0^3)\xi + 2\varphi_1 W_0^2 \xi^2 + \varphi_5 W_0^2 \xi^3 = 0 \tag{39}$$

where, ξ can be determined for the given W_0.

In the second order iteration, we have the following problem of modified eigenvalue

$$(\tilde{L}^2 + K\omega^2 \tilde{L} - \omega^2) W_2 = -KL^*(f(\rho)) + f_{,\rho}(\rho)/\rho \tag{40}$$

when $\rho = 1$: $W_2 = 0,\ K\omega^2 \frac{1}{\rho}\int_0^\rho \rho W_2 d\rho + K(S_r^0 \phi_2 + S_2 \phi^0 + S_2 \phi_2) + \phi_2 = 0,$

$$(\rho S_2)_{,\rho} - \nu S_2 = 0, \quad (\rho T_2)_{,\rho} - \nu T_2 = 0 \tag{41}$$

when $\rho = 0 : W_2 = W_0,$

$$K\omega^2 \frac{1}{\rho}\int_0^\rho \rho W_2 d\rho + K(S_r^0 \phi_2 + S_2 \phi^0 + S_2 \phi_2) + \phi_2 = 0, S_2 < \infty,\ T_2 < \infty \tag{42}$$

where

$$\begin{aligned}
f(\rho) = {} & \rho S_r^0 W_{,\rho}^{(1)} + \rho S^{(1)} W_{,\rho}^0 + f_1(\xi)\rho(W_{,\rho}^0 T^{(1)} + W_{,\rho}^{(1)} S^{(1)}) \\
& + f_2(\xi)\rho W_{,\rho}^{(1)} T^{(1)}
\end{aligned}$$

substitution of Equations (22), (23), (30), (33) and (34) into (40), we obtain the solution of Equation (40)

$$W^{(2)} = W_0 \sum_{j=0}^{\infty} (A_j^{(2)} + B_j^{(2)}) \rho^{2j} + W_0^2 f_1(\xi) \sum_{j=0}^{\infty} C_j^{(2)} \rho^{2j} + W_0^3 f_2(\xi) \sum_{j=0}^{\infty} D_j^{(2)} \rho^{2j} \tag{43}$$

where

$$
\begin{aligned}
A_j^{(2)} &= \delta_1 \frac{\omega^{2j}}{2^{2j}(j!)^2 c^j} + \delta_2 \frac{(-1)^j c^j}{2^{2j}(j!)^2} \\
c &= \omega^2 (K + \sqrt{K^2 + 4/\omega^2})/2 \\
B_0^{(2)} &= B_1^{(2)} = 0 \\
B_j^{(2)} &= \frac{E_{j-2}}{16j^2(j-1)^2} - \frac{\omega^2 K}{4j^2} B_{j-1}^{(2)} + \frac{\omega^2}{16j^2(j-1)^2} B_{j-2}^{(2)} \quad (j \geq 2) \\
E_j &= -16K(j+2)(j+1)^2 \sum_{i=0}^{j+2} i(A_i^{(1)} a_{2(j-i+2)} + B_{j-i+1}^{(1)} \tilde{b}_i) + \\
&\quad + 4(j+1) \sum_{i=1}^{j+1} i(A_i^{(1)} a_{2(j-i+1)} + B_{j-i+1}^{(1)} \tilde{b}_i) \\
\tilde{b}_i &= b_{2i-1}/2i \ (i = 1, 2, \cdots), \quad \tilde{b}_0 = \sum_{m=0}^{\infty} b_{2m+1}/2(m+1) \\
C_0^{(2)} &= C_1^{(2)} = 0 \\
C_j^{(2)} &= \frac{F_{j-2}}{16j^2(j-1)^2} - \frac{\omega^2 K}{4j^2} C_{j-1}^{(2)} + \frac{\omega^2}{16j^2(j-1)^2} C_{j-2}^{(2)} \quad (j \geq 2) \\
F_j &= -16K(j+2)(j+1)^2 \sum_{i=0}^{j+2} i(a_{2j} C_{j-i+2}^{(1)} + B_{j-i+1}^{(1)} A_j^{(1)}) + \\
&\quad + 4(j+1) \sum_{i=1}^{j+1} i(a_{2j} C_{j-i+1}^{(1)} + B_{j-i+1}^{(1)} A_j^{(1)}) \\
D_0^{(2)} &= D_1^{(2)} = 0 \\
D_j^{(2)} &= \frac{G_{j-2}}{16j^2(j-1)^2} - \frac{\omega^2 K}{4j^2} D_{j-1}^{(2)} + \frac{\omega^2}{16j^2(j-1)^2} D_{j-2}^{(2)} \quad (j \geq 2) \\
G_j &= -16K(j+2)(j+1)^2 \sum_{i=0}^{j+2} i A_j^{(1)} C_{j-i+2}^{(1)} + 4(j+1) \sum_{i=1}^{j+1} i A_j^{(1)} C_{j-i+1}^{(1)}
\end{aligned}
$$

substitution of Equation (43) into (41), (42), we obtain

$$\mathbf{H}\delta = 0 \tag{44}$$

where

$$\delta = [\delta_1, \delta_2, 1]^{\mathrm{T}}$$

$$H = \begin{bmatrix} 1 & 1 & -1 \\ h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \end{bmatrix}$$

all of the elements of the matrix H can be written as the functions of ω, the expression is omitted.

By Equation (44), we obtain the frequency equation

$$\det \mathbf{H} = 0 \tag{45}$$

Equation (45) is the characteristic relation for amplitude frequency-load of large amplitude vibration of circular sandwich late under the uniformed load, we can obtain the value of ω corresponding to any load P_0 and W_0 by solving Equation (45), thus, all of the elements of matrix H are determine, and we can determine the δ_1, δ_2 by Equation (44), so far, the solution of the second order modified iteration of the boundary value problem is determined completely.

4. NUMERIC RESULTS AND DISCUSSION

We completed the numeric calculation for discussed circular sandwich plate with edge clamped, let K=0.01,0.05,0.1, P_0=0,25,50,75,100, the results see Figure 1–6. The amplitude frequency-load curves corresponding to different sheer parameters are given in Figure 1–3. According to Figure 1–3, the static load not only can change the fundamental frequency of sandwich plate but also may change the nonlinear effects of the vibration, when the sheer parameter is smaller (K=0.01), the nonlinear effects are decreased gradually with the increasing of the load, when K is larger (K=0.05,0.1), the nonlinear effects are increased gradually with the increasing off the load, and larger of the K, larger of the increasing of the nonlinear effects.

The varies of the nonlinear vibration drift ξ with the load and amplitude are given in Figure 4–6. According to Figure 4–6, the drift of the nonlinear vibration is increased with the increasing of the load. When$K$=0.05, for the smaller amplitude, the varies of the drift of the nonlinear vibration is very small with the varies of the load, when the amplitude is increased to some value, the drift of the nonlinear vibration is increased with the increasing of the load, for the larger K value, the drift of the nonlinear vibration is decreased with the increasing of the load in begin, but when the amplitude is increased arriving some value, it is increased with increasing of the load. In the all, the drift is

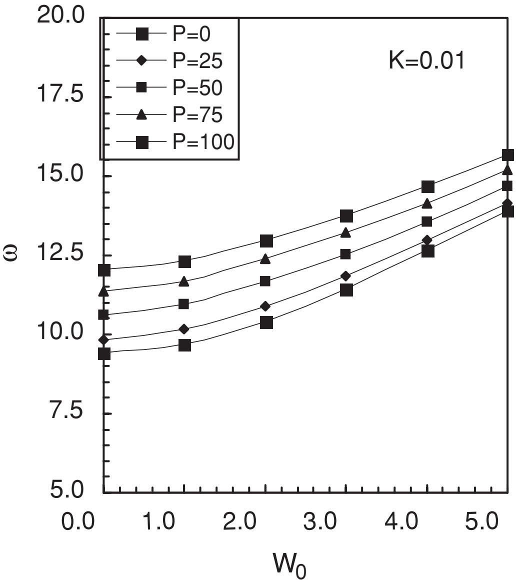


Figure 1. Amplification Frequency-load curve ($K = 0.01$).

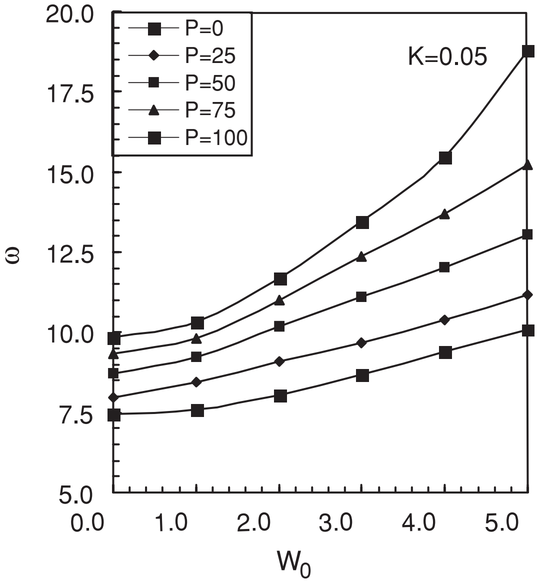


Figure 2. Amplification Frequency-load curve ($K = 0.05$).

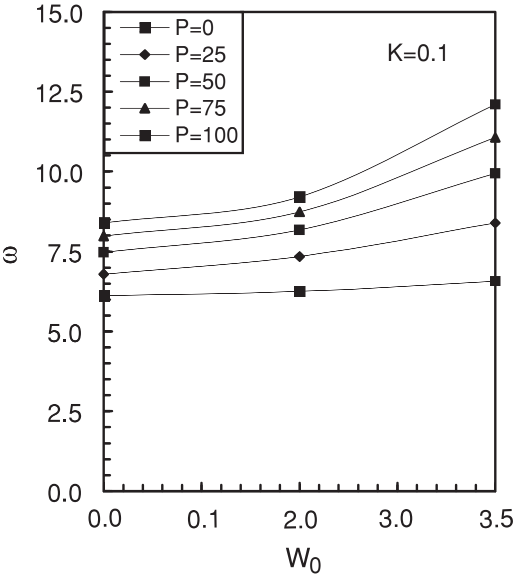


Figure 3. Amplification Frequency-load curve ($K = 0.10$).

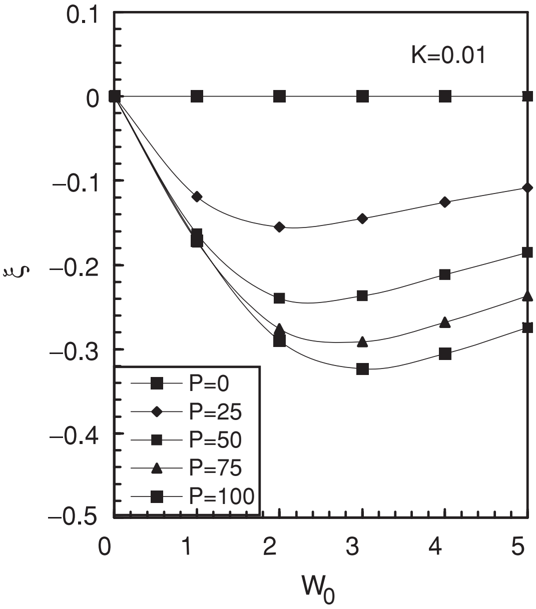


Figure 4. Effects of load to drift ($K = 0.01$).

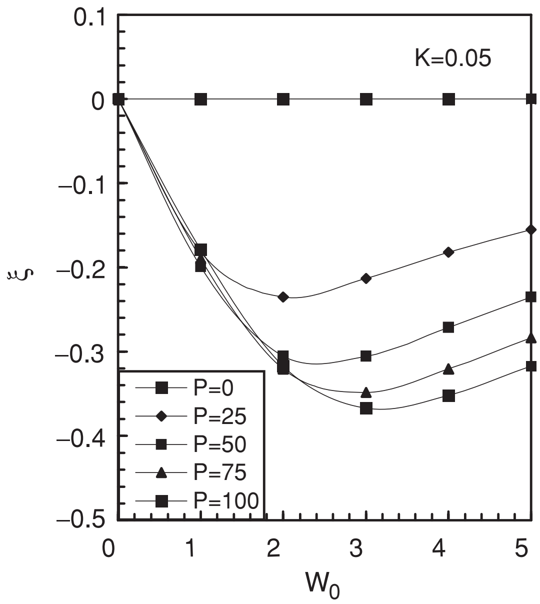


Figure 5. Effects of load to drift ($K = 0.05$).

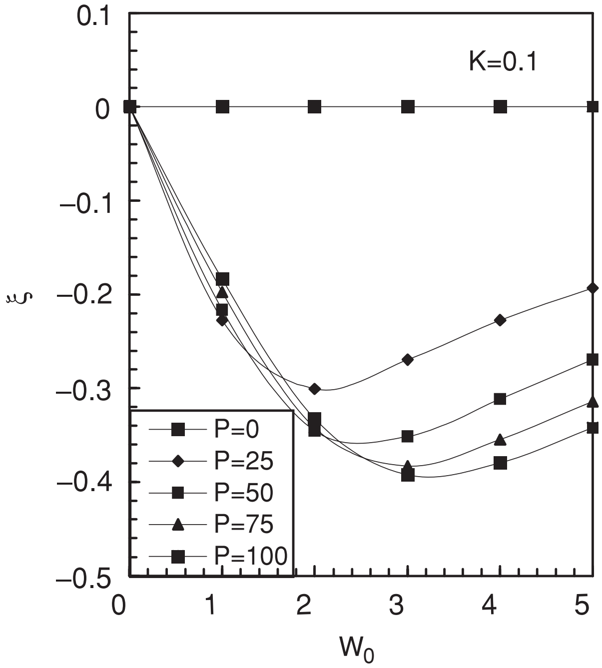


Figure 6. Effects of load to drift ($K = 0.10$).

increased with increasing of the amplitude in begin, when the drift arrived some limited value, it is decreased with increasing of the amplitude.

REFERENCES

1. Liu Renhuai (1981), Nonlinear bending of circular sandwich plates. *Applied Mathematics and Mechanics*, 2, 2, p. 173.
2. Liu Renhuai and Shi Yunfang (1982), Exact solution for circular sandwich plates with large deflection. *Applied Mathematics and Mechanics*, 3, 1, p. 11.
3. Liu Renhuai and Zhu Gaoqiu (1989), Further study on large deflection problem of circular sandwich plates. *Applied Mathematics and Mechanics*, 10, 12, p. 1041.
4. Liu Renhuai (1991), Nonlinear vibration of rectangular sandwich plates. *Scientia Sinica A*, 10, p. 1075.
5. Liu Renhuai (1993), Nonlinear bending of rectangular sandwich plates with edge suported. Applied Mathematics and Mechanics, 14, 3, p. 203.
6. Du Guojun (1994), Large amplitude vibration of circular sandwich plates. *Applied Mathematics and Mechanics*, 15, 5, p. 435.
7. Du Guojun (1996), Further study on large amplitude vibration of circular sandwich plates. *Applied Mathematics and Mechanics*, 17, 11, p. 1087.
8. Zhou Youhe (1992), Nature frequency-load characteristic relation of thin circular plates under a center force. *Chinese Journal of Applied Mechanics*, 1, 9, p. 119.
9. Wang Jinying and Chen Kejin (1993), Vibration problem of flexible circular plates with initial deflection. *Applied Mathematics and Mechanics*, 14, 2, p. 165.
10. Li Dong (1992), Nonlinear vibration of orthotropic shallow shells of revolution. *Applied Mathematics and Mechanics*, 13, 4, p. 313.
11. A.H. Nayfch and D.T. Mook (1979), *Nonlinear Oscillation*, John Wilcy and Sons Inc., New York.

ENERGY ABSORPTION CAPACITY OF LAYERED FOAM CLADDING

G.W. Ma and Z.Q. Ye
Department of Civil and Environmental Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798

Abstract In the present study, layered foam cladding is proposed to enhance the energy absorption capacity against blast load. The energy absorption of the layered foam cladding is investigated based on a rigid, perfectly plastic, locking foam model. The maximum blast impulses that can be resisted by different configurations of layered foam cladding are calculated. It is shown that a double-layer foam cladding can resist much higher impulse and is flexible for application of different purposes.

Keywords: layered foam cladding, blast mitigation, blast load, energy absorption.

1. INTRODUCTION

With the advantage of energy absorption and blast wave attenuation, metallic foams have been attached as sacrificial claddings to protect main structures against impacts and blast loads [1, 2]. However, it have been reported that the sacrificial foam claddings intended to attenuate blast loads may work unexpectedly as pressure amplifiers when the sacrificial claddings were not properly designed [3]. Thus, it is of importance to study the dynamical deformation mechanism in foam materials in order to optimize the design of the sacrificial layers for structural protection.

It has been found that a shock wave may rise in the foam materials when subjected to high strain rate compression. In order to describe the propagation of the shock wave in the metallic foam materials, the stress–strain curve of foam materials was idealized by Reid and Peng [4] as a rigid, perfectly plastic, locking (R-P-P-L) model as shown in Figure 1. By neglecting the short elastic regime of the foam material, the model has shown great advantages to estimate the blast resisting capacity of the foam structures [5, 6]. It is worth noting that in

G. R. Liu et al. (eds.), Computational Methods, 1873–1877.

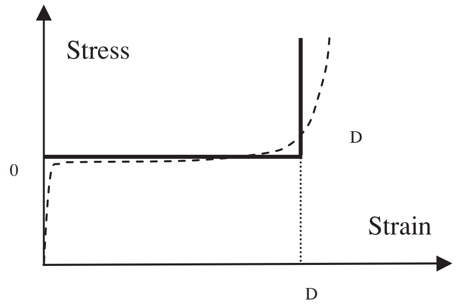


Figure 1. R-P-P-L model of foam.

all the existing investigations, only single-layer foam was used in the cladding structures.

In the present study, layered foam cladding is considered to enlarge the blast mitigation capacity. Different configurations of the layered foam claddings are analysed. The layered cladding can be used to fulfill different application purposes.

2. ENERGY ABSORPTION CAPACITY OF LAYERED FOAM CLADDING

For a cladding made of two foam layers, the front layer and the rear layer are assumed bonded perfectly. Each foam layer has a cover plate. Considering a unit strip along the thickness direction of the cladding, the initial density, total mass, plateau stress, and densification strain of the front layer foam are denoted respectively as ρ_1, m_{f1}, σ_{01}, ε_{D1}, the mass of the cover plate is m_1. The corresponding parameters of the rear layer foam are respectively ρ_2, m_{f2}, σ_{02}, ε_{D2} with a cover plate mass m_2.

As shown in Figure 2, there are two shock waves respectively in the two foam layers. For each shock front, the downstream stress is equal to the plateau

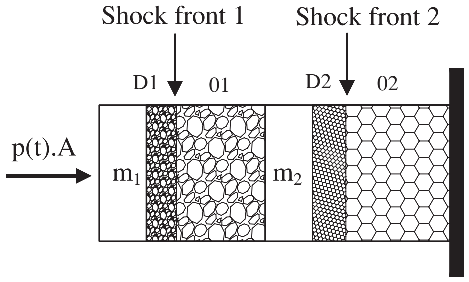


Figure 2. Double-layer foam cladding.

stress σ_0 of the foam, while the upstream stress increases to σ_D accordingly. The deformations of the foams are equal to the relative displacement of the two cover plates. The mass of the compacted foams of the two layers are respectively

$$\Delta m_1 = \frac{\rho_1 A}{\varepsilon_{D1}}(u_1 - u_2) \tag{1a}$$

$$\Delta m_2 = \frac{\rho_2 A}{\varepsilon_{D2}} u_2 \tag{1b}$$

At the upstream of the shock wave front in the two layers, the stresses are

$$\sigma_{D1} = \sigma_{01} + \frac{\rho_1 A}{\varepsilon_{D1}}(\dot{u}_1 - \dot{u}_2)^2 \tag{2a}$$

$$\sigma_{D2} = \sigma_{02} + \frac{\rho_2 A}{\varepsilon_{D2}} \dot{u}_2^2 \tag{2b}$$

Therefore, the equation of the motion of the front foam layer can be expressed as follows,

$$\left[m_1 + \frac{\rho_1 A}{\varepsilon_{D1}}(u_1 - u_2)\right]\ddot{u}_1 + \frac{\rho_1 A}{\varepsilon_{D1}}(\dot{u}_1 - \dot{u}_2)^2 + (\sigma_{01} - p(t))A = 0 \tag{3}$$

The undisturbed part of the front layer $(m_{f1} - \Delta m_1)$ moves together with the compacted foam $(m_2 + (\rho_2 A/\varepsilon_{D2})u_2)$ of the rear layer. Hence the motion equation of the rear foam layer can be expressed as follows,

$$\left[m_{f1} - \frac{\rho_1 A}{\varepsilon_{D1}}(u_1 - u_2) + m_2 + \frac{\rho_2 A}{\varepsilon_{D2}} u_2\right]\ddot{u}_2 + \frac{\rho_2 A}{\varepsilon_{D2}} \dot{u}_2^2 + (\sigma_{02} - \sigma_{01})A = 0 \tag{4}$$

The simultaneous nonlinear differential Equations (3) and (4) can be solved by using a Wilson-θ algorithm. For a given blast impulse, the deformation of the two foam layers can be derived from the above equations. When the rear foam layer is fully compacted, the pressure transferred to the protected structure could be larger than the plateau stress of the rear foam layer σ_{02} or even larger than the input peak pressure p_0. Therefore, the maximum blast impulse that the cladding can resist is the critical value when the rear foam layer becomes fully compacted.

3. BLAST RESISTANCE OF LAYERED FOAM CLADDING

As an example, in the present study, aluminum foam is chosen as the foam material, while steel is used for the cover plate. According to Andrews *et al.* [7], two foams are selected for the present study. Foam I has a density of

Table 1. Maximum blast impulse.

	Maximum blast impulse	
Configuration	Peak pressure (MPa)	Duration (ms)
Single layer (foam I)	24.6	0.3
Dual layers (foam I + I)	43.2	0.3
Two layers (foam I + II)	55.0	0.3
Two layers (foam II + I)	44.6	0.3

200 kg/m^3, plateau stress of 1.5 MPa and densification strain 0.85. Foam II has a density of 300 kg/m^3, plateau stress of 3.0 MPa and densification strain 0.78. The thickness of the foam layers is 10 cm, the thickness of the cover plates is 0.5 cm.

The blast resistances of four different configurations of the cladding have been investigated. Cladding-1 has a single-layer of foam I. Cladding-2 has two identical layers with each layer same as cladding-1. The only difference between cladding-3 and cladding-2 is that the rear foam layer of cladding-3 changes to foam II, which means that the rear layer foam is stronger than the front layer foam. In cladding-4, the foams of the two layers in cladding-3 exchange.

The maximum blast impulses that the four claddings can resist have been calculated, which is shown in Table 1. As can be seen, the double-layer foam claddings have much higher blast resistance than the single-layer foam cladding. The blast resistance of cladding-2 is 76% larger than that of cladding-1. Compared with cladding-2, the blast resistance of cladding-3 is further increased by 27%. On the other hand, the pressure transferred to the protected structure is twice as that of cladding-2 and the weight of cladding-3 is 8% higher. Cladding-4 is also expected to resist much higher blast impulse. However, the results indicate that the blast resistance of cladding-4 has no significant enhancement compared to cladding-2.

4. CONCLUSION

The analytical results show that the multiple layer foam cladding can resist much higher impulse induced by blast load. It can be used for different application purpose. The present study can also serve as design guidance to preliminarily predict the energy absorption capacity and the maximum resistance against the blast impulse of the layered foam claddings.

REFERENCES

1. L.J. Gibson and M.F. Ashby (1997), *Cellular Solids, Structure and Properties*, 2nd ed. Cambridge University Press, Cambridge, UK.
2. M.F. Ashby, A.G. Evans, N.A. Fleck, L.J. Gibson, J.W. Hutchinson and H.N.G. Wadley (2000), *Metal Foams: A Design Guide*. Butterworth Heinemann.
3. D. Karagiozova and N. Jones (2000), Energy absorption of a layered cladding under blast loading. In: *Structures Under Shock and Impact VI*. N. Jones and C.A. Brebbia, (ed.), Southampton, pp. 447–456.
4. S.R. Reid and C. Peng (1997), Dynamic uniaxial crushing of wood. *International Journal of Impact Engineering*, 19, pp. 531–570.
5. J.J. Harrigan, S.R. Reid and C. Peng (1999), Inertia effects in impact energy absorbing materials and structures. *International Journal of Impact Engineering*, 22, pp. 955–979.
6. A.G. Hanssen, L. Enstock and M. Langseth (2002), Close-range loading of aluminum foam panels. *International Journal of Impact Engineering*, 27, pp. 593–618.
7. E. Andrews, W. Sanders and L.J. Gibson (1999), Compressive and tensile behaviour of aluminum foams. *Materials Science and Engineering,* A270, pp. 113–124.

MODELLING BALLISTIC IMPACT ON WOVEN FABRIC WITH LS-DYNA

T.W. Ching and V.B.C. Tan
Department of Mechanical Engineering, National University of Singapore,
9 Engineering Drive 1, Singapore 117576

Abstract This paper presents a novel method of modelling ballistic impact on woven polymeric fabric commonly used in armour applications. This method incorporates the viscoelastic behaviour of the fabric yarns, yarn crimp, inter-yarn friction and friction between projectile and fabric. The yarns of the fabric are modelled as viscoelastic bar elements interwoven together. Excellent agreement between simulation and ballistic test data is obtained in the prediction of the energy absorbed by the fabric and the deformation of the fabric during impact. This is achieved despite the modest number of DOFs.

Keywords: ballistic impact, fabric, perforation, LS-DYNA.

1. INTRODUCTION

Woven fabric is commonly used in today's protective clothing due to their excellent impact resistance, high strength to weight ratio and drapability. The fibres used in the manufacture of such clothing are typically polymeric fibres. Kevlar® and Twaron® are two examples of aramid fibres that are commonly used. The design of such protective clothing is based mainly on extensive ballistic impact tests. There have also been various attempts, noticeably from the last decade, at numerically modelling the ballistic impact of woven fabric to study the mechanics involved in the deformation and perforation of such fabrics subjected to small projectile impacts.

The modelling of fabrics varies greatly in complexity, ranging from simple idealization of fabrics as shells or membranes as reported by Simons et al. [1] and Lim et al. [2], to detailed full scale discretization of the yarns of the fabric with solid finite elements as reported by Shockey [3] and Blankenhorn et al.

G. R. Liu et al. (eds.), Computational Methods, 1879–1884.

[4]. A popular approach to modelling fabric is to represent them as networks of one-dimensional elements pin-jointed at nodes [5–7].

Johnson et al. [8] and Shahkarami et al. [9] used bar elements to model the yarns of the fabric. Shell elements were used to provide contact surfaces for interactions with the projectile and different fabric layers. The difference between their models lies in the fabric material model. Cunniff et al. [10] also used bar elements to model the yarns of the fabric. However, the bar elements of this model are not co-planar, and the nodes of the warp and fill yarns are coupled together with spring elements, to better represent crimped yarns.

The software used by Lim et al., Simons et al., Shahkarami et al., and Blankenhorn et al. for their numerical models is LS-DYNA [11]. LS-DYNA is a non-linear, explicit, finite-element software. It has been used to successfully simulate various types of impact phenomena. The present study also uses LS-DYNA to model the ballistic impact of fabric.

The present study uses only bar elements in the modelling of the fabric yarns. These elements are tied together using spot-weld constraints, and are modelled in a non-co-planar manner. However, the warp and fill yarns of the fabric are not tied together, and can slide along one another. Details of the model can be found in the following section.

2. NUMERICAL MODEL

Actual ballistic tests on woven aramid fabric specimens (Twaron®CT716) were performed prior to the numerical simulations. Fabric specimens of dimensions 120 mm by 120 mm were fully clamped on all sides during the ballistic tests. The projectile was a steel sphere of diameter 12 mm and weight 7 g. The projectile was propelled normally onto the centre of the target by a high-pressure gas gun, with impact velocities ranging 80–520 m/s. The experimental set-up is similar to that of Tan et al. [7].

The fabric yarns were modelled using discrete elements with a viscoelastic material model. As the discrete elements are massless, truss elements with null material properties were used to add mass to the fabric yarns, as well as to provide contact surfaces. The fabric elements were modelled with a length consistent with the actual length of the fabric yarns, and were assigned a radius of 0.05 mm. The projectile was modelled as a rigid sphere and assigned the density of steel. Diagrams of the mesh can be found in Figure 1. Only one quarter of the actual set-up was modelled due to symmetry.

The viscoelastic material model used in this study to represent the aramid yarns is the three-element spring-dashpot model shown in Figure 2. This model

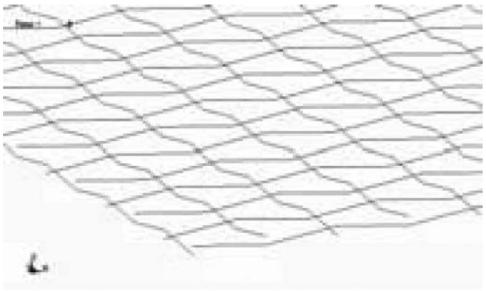


Figure 1. Mesh of fabric.

was also used by Shim et al. and Lim et al. The stress–strain response of this model can be described by the three parameters as

$$\left(1+\frac{K_2}{K_1}\right)\sigma+\frac{\mu}{K_2}\dot{\sigma}=K_2\varepsilon+\mu\dot{\varepsilon} \tag{1}$$

The effects of crimp (undulations in yarns of woven fabric) have been found by Shim et al. to affect the ballistic response of woven fabric and hence it is important to include this in the fabric model. Crimp effects were incorporated in this study by modelling the fabric yarns in a non-planar manner similar to Cunniff, as shown in Figure 1. At the cross-over points, the nodes of the warp and fill yarn elements were placed a distance of 0.1 mm apart in the thickness direction.

In order to allow for fabric perforation, the nodes of the bar elements of the yarns were joined together with spot-weld constraints. These constraints were defined to fail using the same failure criteria employed by Shim et al. [5].

Friction was introduced between the projectile and fabric, as well as between the warp and fill yarns of the fabric. The friction coefficient between yarns and between fabric and steel were tested to be 0.2.

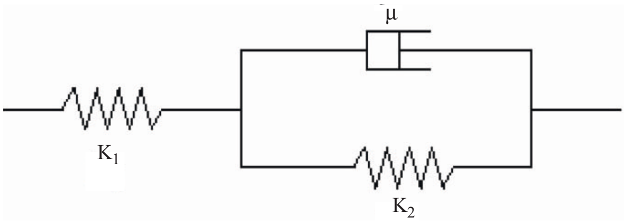


Figure 2. Three-element viscoelastic model.

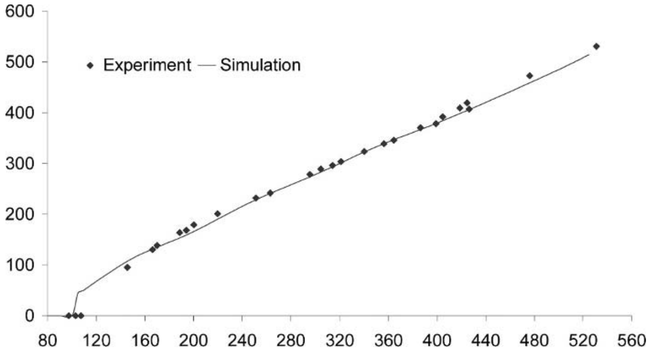


Figure 3. Residual vs. impact velocity (m/s).

3. RESULTS

Figure 3 shows a plot of the residual velocity of the projectile against impact velocity of the projectile. This plot includes the experimental data, as well as the simulation data. A similar plot, for energy absorbed by the fabric (calculated by the loss in kinetic energy of the projectile) against impact velocity of the projectile, can be found in Figure 4.

The deformation plots of the fabric models subjected to impacts at velocities of 110 and 400 m/s can be found in Figure 5.

The plots of residual velocity against impact velocity plot (Figure 3) and the energy absorbed against impact velocity (Figure 4) show that the numerical model is in excellent agreement with the experimental results. The residual velocity is seen to vary approximately linearly for the higher impact velocities. The experimental critical impact velocity for complete penetration of the fabric obtained is around 110 m/s. The numerical model predicts a slightly conservative critical impact velocity of 105 m/s.

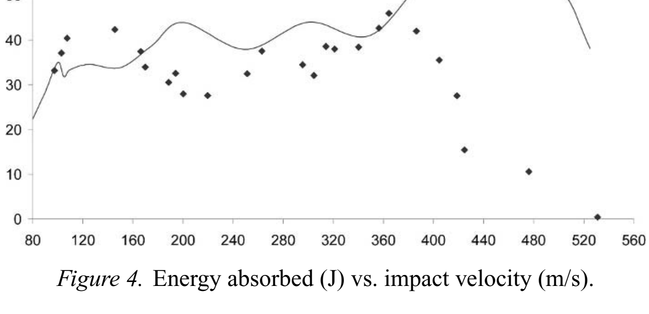
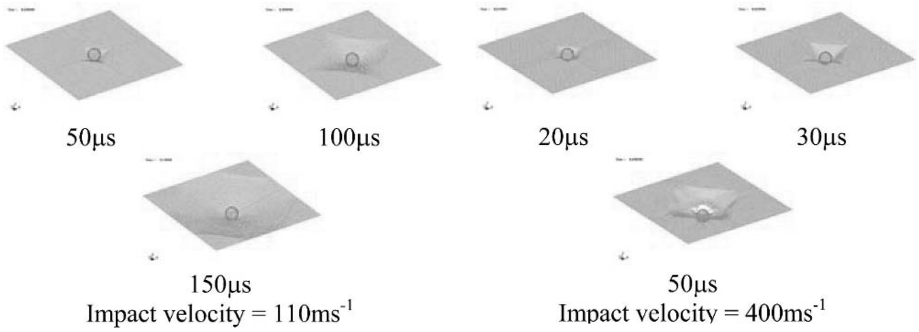


Figure 4. Energy absorbed (J) vs. impact velocity (m/s).



50μs 100μs 20μs 30μs

150μs
Impact velocity = 110ms^{-1}

50μs
Impact velocity = 400ms^{-1}

Figure 5. Sequence of fabric deformation.

The deformation plots of the fabric (Figure 5) show that the pyramidal shape deformation typically observed in high-speed photographs of ballistic impact experiments, was also obtained by the numerical model for both low and high impact velocities.

4. CONCLUSIONS

A numerical model of woven fabric comprising discrete elements interwoven together has been shown to give predictions of ballistic response of woven aramid fabric that are in good agreement with actual tests. The model is able to reproduce the deformation and damage that are observed in actual tests. Quantitative agreement in terms of residual projectile velocities was also obtained.

REFERENCES

1. J.W. Simons, D.C. Erlich and D.A. Shockey (2001), Finite element design model for ballistic response of woven fabrics. *Proceedings of the 19th International Symposium on Ballistics*.
2. C.T. Lim, V.P.W. Shim and Y.H. Ng (2003), Finite-element modelling of the ballistic impact of fabric armour. *International Journal of Impact Engineering*, 28, pp. 13—31.
3. D.A. Shockey, D.C. Erlich and J.W. Simons (1999), Lightweight fragment barriers for commercial aircraft. *Proceedings of the 18th International Symposium on Ballistics*.
4. G. Blankenhorn, K. Schweizerhof and H. Finckh (2003), Improved numerical investigation of a projectile impact on a textile structure. *Proceedings of the 4th European LS-DYNA Users Conference.*
5. V.P.W. Shim, V.B.C. Tan and T.E. Tay (1995), Modelling deformation and damage characteristics of woven fabric under small projectile impact. *International Journal of Impact Engineering*, 16, pp. 585–605.

6. C. Ting, J. Ting, P. Cunniff and D. Roylance (1998), Numerical characterization of the effects of transverse yarn interaction on textile ballistic response. *30th Int. SAMPE technical conference*.
7. V.B.C. Tan, V.P.W. Shim and T.E. Tay (2003), Experimental and numerical study of the response of flexible laminates to impact loading. *International Journal of Solids and Structures*, 40, pp. 6245–6266.
8. G.R. Johnson, S.R. Beissel and P.M. Cunniff (1999), A computational model for fabrics subjected to ballistic impact.*Proceedings of the 18th International Symposium on Ballistics.*
9. A. Shahkarami, R. Vaziri, A. Poursartip and K. Williams (2002), A numerical investigation of the effect of projectile mass on the energy absorption of fabric panels subjected to ballistic impact. *Proceedings of the 20th International Symposium on Ballistics*.
10. P.M. Cunniff and Ting J. (1999), Development of a numerical model to characterize the ballistic behaviour of fabrics. *Proceedings of the 18th International Symposium on Ballistics.*
11. J. Hallquist (1998), *LS-DYNA Theoretical Manual*. Livermore Software Technology Corporation.

RESEARCH FOR EXPLOSION OF HIGH EXPLOSIVE IN COMPLEX MEDIA

M. Otsuka[1], S. Tanaka[1] and S. Itoh[2]
[1] *Graduate School of Science and Technology, Kumamoto University, 2-39-1 Kurokami, Kumamoto City, Kumamoto 860-8555, Japan.*
[2] *Shock Wave and Condensed Matter Research Center, Kumamoto University, 2-39-1 Kurokami, Kumamoto City, Kumamoto 860-8555, Japan.*

Abstract Numerical analysis is an important tool in obtaining a detailed understanding about explosive phenomena. In this research, we perform the analysis of multi-layer model with the air layer. In our calculation code, the analysis of compressible substances, which cells transform greatly (like air) generally is very difficult. In this research, we used LS-DYNA3D, an analysis code using a finite-element method. By analyzing the stress state of the air hole circumference, multilayer models such as air, water and structure are analyzed and it aims to get the propagation process of the shock wave. The multi-material Arbitrary Lagrangian Eulerian (MMALE) [1] method improves upon pure Eulerian formulation by allowing the reference fluid meshes to translate, rotate and deform, thus minimize the amount of flux transport, and reduce mesh size of the reference fluid meshes.

Keywords: explosive, shock wave, air, the multi-material Arbitrary Lagrangian Eulerian (MMALE) method, LS-DYNA3D.

1. INTRODUCTION

Explosives can easily generate high energy and ultra-high pressure. In recent years, research on the advanced technological use of explosives is studied in various places. Here we focus on the Smooth Blasting Technique that is applied for tunnel blasting. This technique is performed to fracture concrete and to reduce the quantity of fragments, while avoiding stress concentration of ground pressure during tunnel blasting. An emulsion explosive is usually used for Smooth Blasting Technique. In this research we investigated how the shock wave generated from the explosive would act using a multilayer model. The

G. R. Liu et al. (eds.), Computational Methods, 1885–1890.

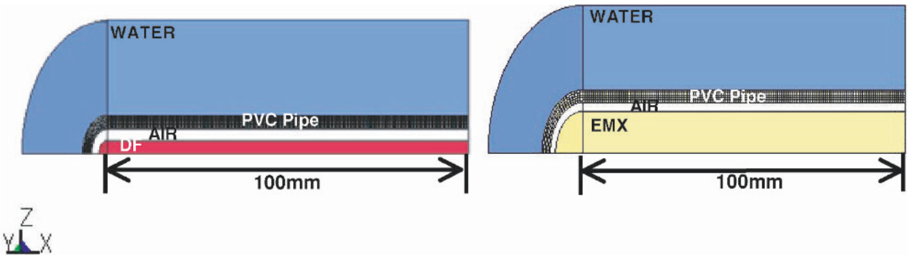


Figure 1. Numerical analysis model.

explosives used by this research are detonation cord (DF) and an emulsion explosive (EMX).

2. NUMERICAL ANALYSIS METHOD

The numerical analysis model is shown in Figure 1. This model has adopted 1/4 model of a pillar for calculating space reduction. The radial direction is set to z_e (explosive part), z_a (air part), z_v (poly vinyl chloride (PVC) pipe part = 3 mm) and z_w (water part = 20 mm). The used explosives are detonation cord (density 1.2 g/cm^3, detonation velocity 6300 m/s, detonation pressure 11.94 GPa, $z_e = 2.7$ mm) and emulsion explosive (density 0.9 g/cm^3, detonation velocity 2520 m/s, detonation pressure 1.4 GPa, $z_e = 10$ mm). The explosive detonated from the part at $x = 0$ which gave initial detonation. The air part was changed and numerical analysis was carried out. We evaluated the effect of the air layer by the pressure histories.

Numerical analysis was carried out using the Arbitrary Lagrangian Eulerian (ALE) approach. An ALE formulation contains both pure Lagrangian and pure Eulerian formulations. This has combined the Lagrange method (which observes change of the velocity, acceleration and so on paying attention to arbitrary fluid particles) and the Euler method (for dealing with change of physical quantity, such as velocity and pressure) as function of the position of the point and time in space. We used the multi-material ALE Formulation). The multi-material method uses the Euler element for fluid and the detonation products of the explosive, and uses the Lagrange element for a structure. Two levels of ALE technology exist. One allows ALE behaviour only within a material (forcing material boundaries to remain Lagrangian). The second level of ALE technology allows multi-material elements to form and is therefore more generally applicable.

About the handling of the detonation phenomenon in this analysis 'C-J Volume Burn' was used.

Table 1. JWL parameter of DF and EMX.

	A (GPa)	B (GPa)	R_1	R_2	ω
DF	452.35	8.85	5.485	1.425	0.28
EMX	1.51	4.43	5.8	1.7	0.24

The equation of state for explosives in this research made use of the Jones-Wilkins–Lee (JWL) equation [2]. This equation is shown below (1) and each coefficient is shown in Table 1.

$$\begin{aligned} P_{\overline{\mathrm{JWL}}} = {} & A\left[1-\frac{\omega}{VR_1}\right]\exp(-R_1V) \\ & +B\left[1-\frac{\omega}{VR_2}\right]\exp(-R_2V)+\frac{\omega e}{V} \end{aligned} \tag{1}$$

$V=\rho_0$ (Initial density of an explosive)/ρ(Density of detonation gas)
P_{JWL}: Pressure e: Specific internal energy
A, B, R_1, R_2, ω: JWL parameter.

The equation of state for water and PVC pipe in this research was used Mie–Grüneisen equation [3]. Mie–Grüneisen equation is shown below (2) and each coefficient is shown in Table 2.

$$P=\frac{\rho_0C_0^2\eta}{(1-s\eta)^2}\left[1-\frac{\Gamma_0\eta}{2}\right]+\Gamma_0\rho_0e \tag{2}$$

$\eta=1-\rho_0$ (initial density of the medium)/ρ(density of the medium)
P: Pressure e: Specific internal energy
$C_{0,S}$: Constant of material Γ_0: Grüneisen coefficient.

The equation of state for an air in this research was used linear polynomial equation. This equation is shown below and each coefficient is shown in Table 3.

$$P=C_0+C_1\mu+C_2\mu^2+C_3\mu^3+(C_4+C_5\mu+C_6\mu^2)E \tag{3}$$

Table 2. Mie–Grüneisen parameter of water and PVC pipe.

	ρ_0 (kg/m^3)	C_0 (m/s)	s	Γ_0
Water	1000	1490	1.79	1.65
PVC Pipe	1380	2300	1.47	0.40

Table 3. Linear polynomial parameter of air.

	ρ_0 (kg/m^3)	γ	C_4	C_5
Air	1.025	1.403	0.403	0.403

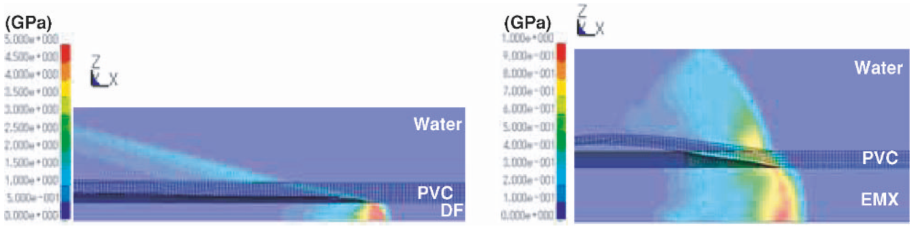


Figure 2. Relation of pressure contours using EMX and DF without an air-layer (detonation front at x = 70 mm).

The linear polynomial equation of state was used model gas with the gamma law equation of state. This may be achieved by setting.

$$C_0 = C_1 = C_2 = C_3 = C_6 = 0$$

and

$$C_4 = C_5 = \gamma - 1$$

Where γ is ratio of specific heats.

3. ANALYSIS RESULTS

The pressure contours of Figure 2 shows the shock wave which propagates into the PVC pipe and water when using DF and EMX as the explosive. There is no air layer in these models. Since the sound velocity (2300 m/s) of a PVC pipe is almost the same as the detonation velocity (2520 m/s) of EMX, the detonation front of EMX and the shock front of a PVC pipe are located in almost the same x-position. Incidentally, since destruction and deformation were taken into consideration to the PVC pipe, a failure strain term was incorporated in all the models used by this research.

Figure 3 shows the pressure contours of the numerical model with a 2.43 mm thick air layer using DF as the explosive. A shock wave and reflective wave can be seen in the water. Pressure histories are shown in Figures 4(a) and 4(b). When the air layer becomes thick, it turns out that $P_{\max}$ is low in the pipe and the water. However, with the air layer, the duration of pressure is increased by the effect of a reflective wave. Therefore, the impulse shows almost the same value in all cases.

4. CONCLUSIONS

In this research, we performed the numerical analysis of multilayer model with the air layer. Rather than the underwater shock wave produced by

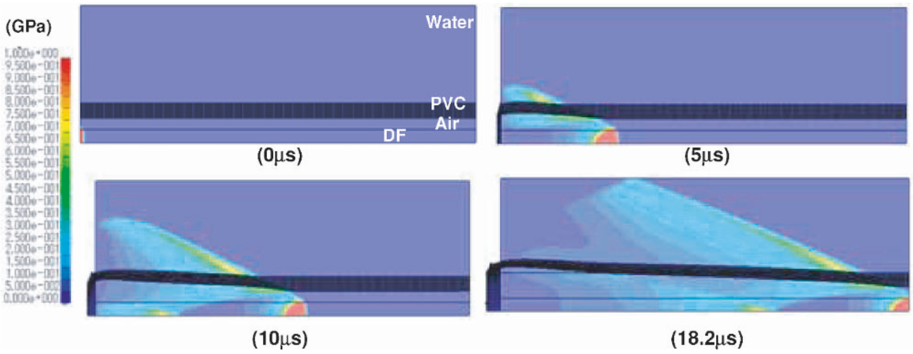


Figure 3. Pressure contours in the case of air layer thickness is 2.43 mm.

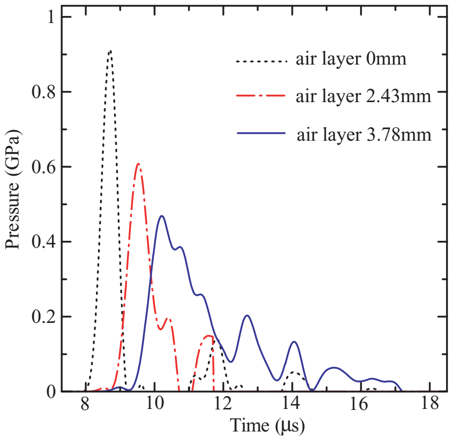


Figure 4.(a). Pressure histories in PVC pipe (x = 40 mm, y = 0 mm, z_v = 1.5 mm).

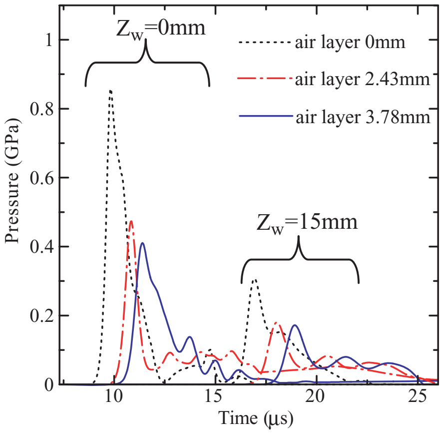


Figure 4.(b). Relation of pressure histories in water between and z_w = 0 mm and z_w = 15 mm (x = 40 mm, y = 0 mm).

detonation of an explosive, an air layer showed that a weaker shock wave was introduced, but also that a reflected pressure wave occurred. The duration of the pressure was increased by this, leaving the impulse practically unchanged.

REFERENCES

1. K. Mahmadi, N. Aquelet and M. Souli (2002), ALE multi-material formulation of high explosive detonation using LS-DYNA3D. In: *Emerging technologies in Fluids, Structures, and Fluid/Structure Interactions*, PVP Vol. 446–1.
2. E.L. Lee, M. Finger and W. Collins (1973), JWL equation of state coefficients for high explosives. Lawrence Livermore Laboratory, UCID-16189.
3. S.P. Marsh (1980), *LASL Shock Hugoniot Data*. University of California Press, Berkeley, CA.

NUMERICAL CALCULATION OF DETONATION PHENOMENON FOR EMULSION EXPLOSIVES

H. Hamashima[1], S. Itoh[1], F. Sumiya[2] and Y. Kato[2]
[1] *Shock Wave and Condensed Matter Research Center, Kumamoto University, 2-39-1 Kurokami, Kumamoto 860-8555, Japan*
[2] *NOF CORPORATION, 61-6 Kitakomatsudani, Taketoyo-cho, Chita-gun, Aichi 470-2398, Japan*

Abstract Usually, it is difficult to obtain the equation of state (EOS) for the non-ideal explosive, such as emulsion explosives (EMXs), that is most used in an industrial explosive. The reason is that the detonation performance of a non-ideal explosive changes with the charge diameter and confinement a lot. In this research, as for the EMXs, it asked for the parameters of the JWL EOS obtained from the shape of an underwater shock wave (Underwater Explosion Test). Numerical calculation was performed about the detonation phenomenon of EMXs using this equation, and it compared with the experiment. The result was well in agreement between them.

Keywords: JWL equation of state, explosive, underwater, ALE.

1. INTRODUCTION

The emulsion explosives (EMXs) show non-ideal detonation behaviour, and their detonation velocities are easily controlled by selecting the void size and adjusting the quantity of voids included. Their detonation waves steadily propagate. However, their characteristics are significantly affected by the conditions such as charge diameter or confinement. JWL EOS [1] is widely used because of its simplicity in hydrodynamic calculations. JWL EOS contains parameters that may be determined by the metal cylinder expansion test. However, the real expansion of detonation products for EMXs is not estimated by the metal cylinder expansion because of using them confined in the metal cylinder. In order to study the expanding process of detonation products of EMXs, the optical observation of the underwater explosion of cylindrical EMXs was carried out. Using a method of characteristics applied to the configurations of underwater

G. R. Liu et al. (eds.), Computational Methods, 1891–1896.

Table 1. Compositions of emulsion matrix (wt %).

	Ammonium nitrate	Sodium nitrate	Hydrazine nitrate	Water	EDTA EDTA	Wax and emulsifier	Balloon
EMXA	72.29	6.22	5.52	11.04	0.10	4.82	RMB
EMXB	72.29	6.22	5.52	11.04	0.10	4.82	EPS
EMXC	70.44	6.22	7.37	11.05	0.10	4.82	EPS

Table 2. Structures of balloons.

	Average diameter (mm)	Bulk density (kg/m^3)	Structure	Material
RMB	0.05	27	Mono-cell	Acrylonitrile/vinylidene chloride
EPS	2.2	43	Multi-cell	Polystyrene

shock waves and applying one-dimensional hydrodynamic analysis for the axis symmetric flow, the expanding process of detonation products is made clear in all stages. Therefore, the pressure and density of detonation products can be determined. The parameters of JWL EOS are obtained by using this technique.

2. EXPERIMENTAL PROCEDURE [2]

Three kinds of sample EMXs are used in this study. Their compositions of emulsion matrix are shown in Table 1. The structures of balloons mixed in them are shown in Table 2. The detonation properties of sample explosives are shown in Table 3. Sample explosives were set in the aquarium made of Polymethylmethacrylate (PMMA) and initiated by No. 6 electric detonator. Streak photographs and framing photographs are taken by a high-speed camera (IMACON468, HADLAND PHOTONICS) using a conventional shadowgraph system. The configurations of underwater shock waves were obtained from the streak photographs. The experimental device for the cylindrical sample

Table 3. Detonation properties of emulsion explosive.

	Detonation velocity (m/s)	Initial density (kg/m^3)
EMXA	4390	900
EMXB	2520	900
EMXC	2540	900

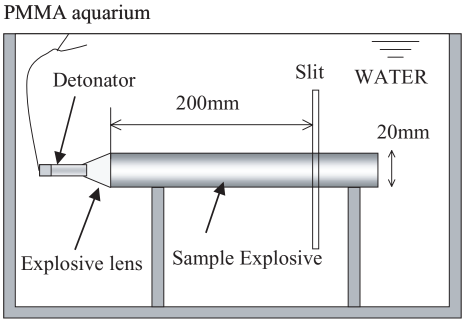


Figure 1. Experimental configuration.

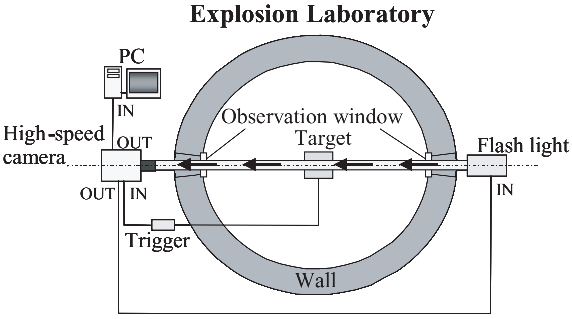


Figure 2. Experimental set-up.

explosive is shown in Figure 1. In this figure, the slit shows the optical slit for taking the streak photograph. The experimental set-up in the Explosion Laboratory is shown in Figure 2.

3. METHOD OF CHARACTERISTICS

The underwater shock wave system described in stationary coordinate system fixed to detonation front is shown in Figure 3. Boundary between detonation products and water is shown by curve AB. Curve of Characteristics, such as S_1B_1, is described between underwater shock wave AS and this boundary. If the configuration of underwater shock wave is given, physical quantities of a range between AS and AB are obtained by using some equations for underwater shock wave, streamline and Curve of Characteristics. Using one-dimensional hydrodynamic analysis for the axis symmetric flow, the pressure and density of products are found by making the underwater expanding process of the detonation products clear. Hence, if the configurations of underwater shock waves are

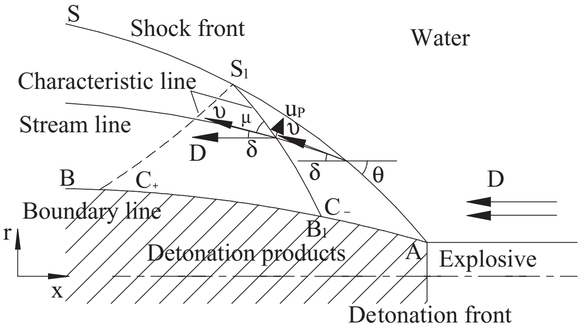


Figure 3. Stationary coordinate system.

known, the expanding process of the detonation products is made clear, even if the composition of explosive is unknown.

4. NUMERICAL SIMULATION

The numerical simulation of the underwater explosion of cylindrical explosive was conducted by Arbitrary Lagrangian–Eulerian (ALE) method [3], by using C-J Volume Burn Technique [4] and by using the laws of conservation of mass, momentum, energy and EOS. Mie–Grüneisen EOS is used for water.

$$P = \frac{\rho_0 C_0^2 \eta}{(1 - s\eta)^2}\left(1 - \frac{\Gamma\eta}{2}\right) + \Gamma\rho_0 e, \quad \eta = 1 - \frac{\rho_0}{\rho}. \tag{1}$$

The constants of Mie–Grüneisen EOS for water are shown in Table 4.

JWL EOS was used for the detonation products.

$$P = A\left(1 - \frac{\omega}{R_1 V}\right)\exp(-R_1 V) + B\left(1 - \frac{\omega}{R_2 V}\right)\exp(-R_2 V) + \frac{\omega\rho_e e}{V}, \tag{2}$$

where A, B, R_1, R_2, ω are JWL parameters. V is ρ_e (density of explosive)/gρ (density of detonation products). JWL parameters of EMXs obtained from underwater explosion test are shown in Table 5.

Table 4. Constants of Mie–Grüneisen EOS.

Material	ρ_0 (kg/m^3)	C_0 (m/s)	s	Γ
Water	1000	1489	1.79	1.65

Table 5. JWL parameters of sample emulsion explosive.

	A (GPa)	B (GPa)	R_1	R_2	ω
EMXA	169.0	3.67	5.69	1.19	0.33
EMXB	61.5	0.89	6.01	2.01	0.19
EMXC	72.9	0.92	6.35	2.15	0.22

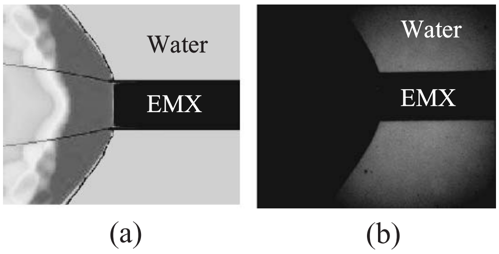
5. RESULTS

The configurations of underwater shock wave for cylindrical explosive EMXB obtained from the numerical and experimental results are shown in Figure 4. The cylindrical explosive has 20 mm in diameter and 250 mm long. Good agreement is obtained between the numerical and experimental results.

The configurations of underwater shock wave for all sample explosives obtained from the numerical and experimental results are shown in Figure 5. The vertical axis is the distance from the outer surface of the explosive in the direction of radius. The horizontal axis is the distance measured from the detonation front. Good agreements between the numerical and experimental results are obtained for all samples.

6. CONCLUSIONS

A new technique in determining the JWL parameters of detonation products for emulsion explosives is proposed in this paper. This technique developed the method of characteristics in the relation between underwater shock wave and the expansion wave of detonation products. Using this theory, we can estimate the relation between the pressure and volume in the expanded region of detonation products. Then finally we can get the parameters of JWL EOS. It is concluded



(a) (b)

Figure 4. Configurations of underwater shock wave for EMXB: (a) numerical result (b) experimental result.

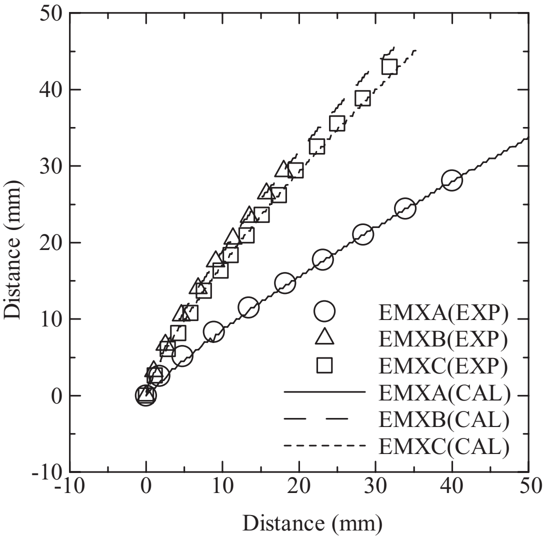


Figure 5. Configurations of underwater shock wave for numerical and experimental results.

that for cylindrical charge, the configuration of the underwater shock wave is well estimated by the numerical calculation.

REFERENCES

1. J.W. Kury, H.C. Hornig, E.L. Lee, J.L. McDonnel, D.L. Ornellas, M. Finger, F.M. Strange and M.L. Wilkins (1965), Metal acceleration by chemical explosive. In: *Fourth Symposium on Detonation*, p. 3.
2. S. Itoh, Z. Liu and Y. Nadamitsu (1997), *Transactions of the ASME*, 119, p. 498.
3. A.A. Amsden, H.M. Ruppel and C.W. Hirt (1980), SALE: A Simplified Ale Computer Program for Fluid Flow at All Speeds, LA-8095, UC-32.
4. C.L. Mader (1979), *Numerical Modeling of Detonations*. University of California Press, Berkeley and Los Angeles, CA.

NUMERICAL SIMULATION OF EXPLOSIVE FORMING

H. Iyama[1], T. Hinata[2], M. Otsuka[2] and S. Itoh[2]
[1] *Department of Mechanical and Electrical Engineering, Yatsushiro National College of Technology, 2627 Hirayamashin-machi, Yatsushiro, Kumamoto 866-8501, Japan*
[2] *Shock Wave and Condensed Matter Research Center, Kumamoto Univ., 2-39-1 Kurokami, Kumamoto 860-8555, Japan*

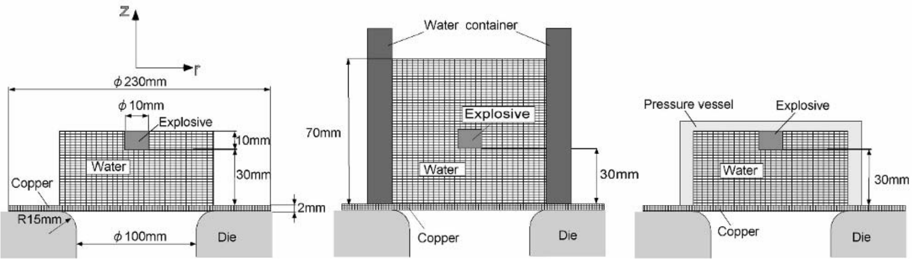
Abstract Explosive forming is one of the unconventional techniques, in which, most commonly, the water is used as the pressure transmission medium. The explosive is set at the top of the pressure vessel filled with water, and is detonated by an electric detonator. The underwater shock wave propagates through the water medium and impinges on the metal plate, which in turn, deforms. There is another pressure pulse acting on the metal plate as the secondary by product of the expansion of the gas generated by detonation of explosive. The secondary pressure pulse duration is longer and the peak pressure is lower than the primary shock pressure. However, the intensity of these pressure pulse is based also on the conditions of a pressure vessel. In order to understand the influence of the configuration of the pressure vessel on the deformation of a metal plate, numerical analysis was performed. This paper reports those results.

Keywords: explosive forming, underwater shock wave, finite difference method, high strain rate, high explosive.

1. INTRODUCTION

When an explosive is detonated underwater, an underwater shock wave is generated which is primarily responsible for the plastic deformation of the metal plate. After the impingement of this primary underwater shock wave on the metal plate, a secondary pressure pulse, generated from the expanding gaseous products of the detonation, impinges on the metal plate. These pressure states of acting on a metal plate change with the type of the container to be used. The shock pressure and gas pressure acting on a metal plate in the case of the sealed

G. R. Liu et al. (eds.), Computational Methods, 1897–1901.



Model 1 Model 2 Model 3

Figure 1. Simulation models.

pressure vessel can be predicted easily. However, in this method, the required forming configuration was obtained with a lidless pressure vessel. Nevertheless, in the case of employing a simple water filled container the influence of reflected pressure from sidewall of container is significant. If no container is used then there is almost no influence of gas pressure. Therefore, an attempt was made to investigate the influence of the pressure vessels on the deformation of a metal plate.

2. SIMULATION METHOD

Figure 1 shows three simulated models of underwater explosive forming. All models were for axi-symmetric free forming operations. Free forming indicates that the metal plate deforms by overhanging. In model 1, the pressure acts on a metal plate without the confinement of pressure vessel. In model 2, there was no lid for the container. The pressure pulse released from the upper part of explosive is unrestricted because of the lidless container. In model 3, an airtight container was used. In all models, the copper plate with 2 mm thick and of diameter 230 mm. The size of an explosive was determined as in model 1. A distance between the explosive and metal plate was set to 30 mm. High explosive, SEP was used. The density of this explosive is 1.31 kg/m^3 and detonation pressure is 15.9 GPa. The copper plate was placed on the die. The die shoulder is of 15 mm radius of curvature. The pressure vessel and die holder are assumed to be rigid body.

The r–z coordinates are set as shown in Figure 1. The simulation method was finite difference method (FDM) using Lagrangian coordinates. The copper plate is treated as the viscous elastic-plastic material. The pressure is solved from the following Mie–Grüneisen equation of state [1],

$$P = \frac{\rho_0 c_0^2 \eta}{(1 - s\eta)^2}\left[1 - \frac{\Gamma_0 \eta}{2}\right] + \Gamma_0 \rho_0 e \tag{1}$$

Table 1. Material constants in Mie–Grüneisen EOS.

	ρ_0(kg/m^3)	C_0(m/s)	S	Γ_0
Water	1000	1490	1.79	1.65
Copper	8930	3940	1.48	1.96

where, ρ_0 is initial density, e is energy, Γ_0 is Grüneisen parameter, $\eta = 1 - \rho_0/\rho$, c_0 and s are material constants. For the related materials, the values of those constants are given in Table 1.

The pressure in detonation products of explosive is calculated by using the JWL (Jones–Wilkins–Lee) equation of state [2]. The equation has the following expression,

$$P = A\left(1 - \frac{\omega}{R_1 V}\right)\exp(-R_1 V) + B\left(1 - \frac{\omega}{R_2 V}\right)\exp(-R_2 V) + \frac{\omega E}{V} \qquad (2)$$

where, A, B, R_1, R_2, C and ω are constants (or JWL parameters). V is the ratio of the volume of the product gases to initial volume of the undetonated explosive. For the explosive of SEP, those constants were obtained from cylindrical expansion test and are given in Table 2.

3. RESULTS AND DISCUSSION

Figure 2 shows the pressure profiles of the water element on the copper plate in the position of $r = 0$–50 mm. In all the models, the pressure pulse is seen to act on the metal plate from the central part, and the pressure peak is also observed to be decreasing gradually from central part. Figure 3 shows deformation process of each model. In model 1, the quantity of deformation of copper plate is small in comparison with the other two models and the propagation of the plastic bending wave of metal plate [3] does not occur. In model 2, the propagation of the plastic bending wave occurs and the deformation of the copper plate at the central region is seen in the last stage. In model 3, as the pressure pulse is of larger magnitude in comparison with other models, the central region of the copper plate is conical.

Table 2. JWL parameters for SEP explosive.

A(GPa)	B(GPa)	R_1	R_2	ω
365	2.31	4.30	1.10	0.28

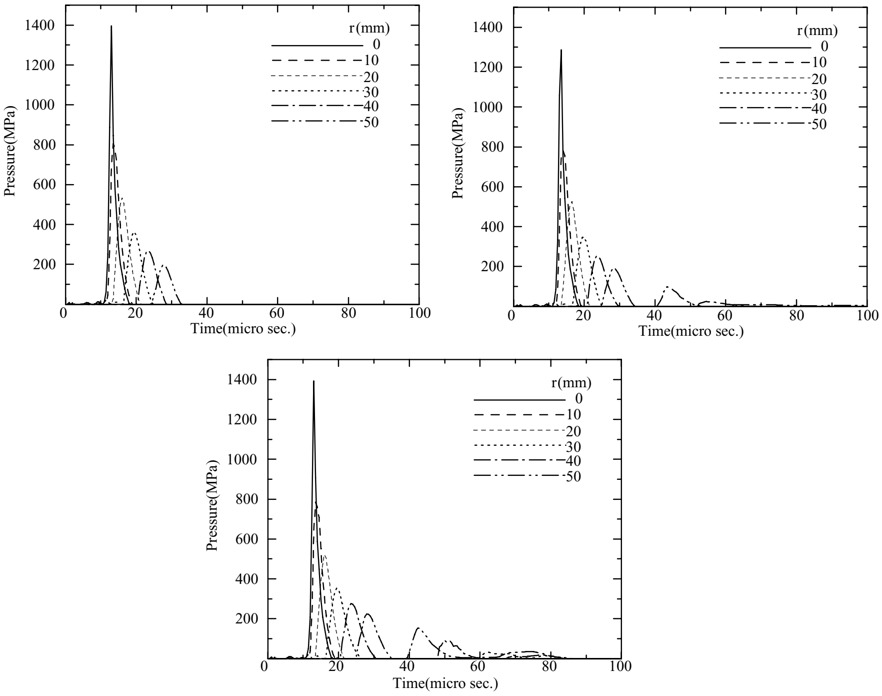


Figure 2. Pressure profiles of water element on the metal plate from 0 to 100 μs.

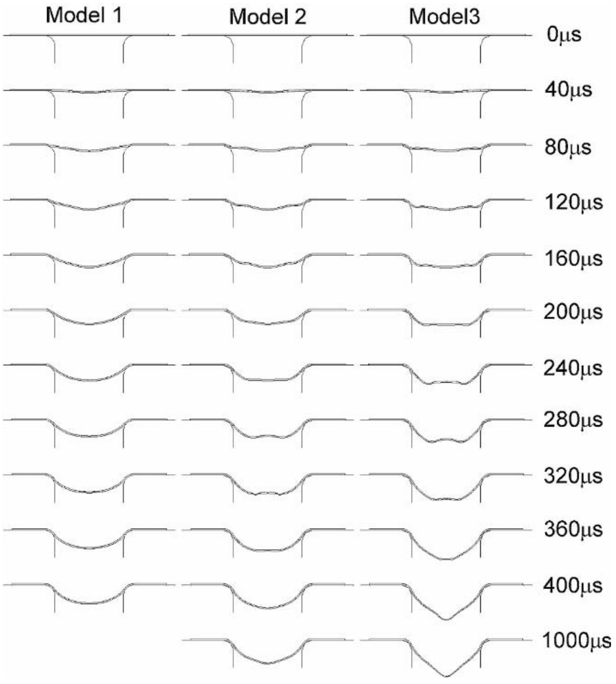


Figure 3. Deformation processes in three models.

4. CONCLUSION

In the first model l, there was no pressure vessel on the metal plate, enveloping the explosive charge. Consequently, the underwater shock wave generated by detonation of explosive, was exposed all around. In the model 2, a lidless container was employed. In the final case, model 3, an airtight pressure vessel with lid was used. The pressure history acting on the metal plate varied. It is seen that the changes in the pressure vessel influence the deformation process of the metal plate. In the case the use of an airtight pressure vessel, model 3, as the pressure generated inside the vessel was completely used, without getting vented to atmosphere, for the deformation of the metal plate, the deformation is significantly large and therefore, the efficiency of the process.

REFERENCES

1. G. McQueen, S.P. Marsh, J.W. Taylor, J.N. Fritz and W.J. Carter (1970), The equation of state of solids from shock wave studies. *High-Velocity-Impact Phenomena*. Academic Press, NY, p. 230.
2. E. Lee, M. Finger and W. Collins (1973), JWL equation of state coefficients for high explosives. *Lawrence Livermore National Laboratory Report*, UCID-16189, Livermore, CA.
3. H. Iyama, Z.Y. Liu and S. Itoh (2001), Numerical simulation of free forming using underwater shock wave. In: *Proceedings, ASME PVP*, Vol. 431, pp. 171–176.

FLEXIBLE AND INCOMPRESSIVE GOAL NETS IN SOCCER

F. Fujii[1], H. Noguchi[2] and K. Oda[3]
[1] *Department of Mathematical and Computational Engineering, Gifu University, Gifu 501-1193, Japan*
[2] *Department of System Design Engineering, Keio University, Yokohama 223-8522, Japan*
[3] *Advanced Structures R&D Department, Taiyo Kogyo Corporation, Hirakata 573-1132, Japan*

Abstract Sports nets and hammocks are usually not pre-tensioned and too slack and flexible to resist compressive forces, which the high pre-stressing would counterbalance in tension structures. Due to this incompressibility, the low-tension nets may exhibit extreme sags in the deformed configuration, which may be one computational problem in analysis. The present study is concerned with this particular, but interesting issue and proposes some modelling ideas to simulate the static and dynamic response of flexible and incompressible nets, which are not frequently treated in engineering applications. Special attention is paid to dynamics of hexagon-mesh goal nets used in FIFA World Cup 2002 in Japan and South Korea. In numerical examples, it is illustrated that the hexagon-mesh net is much better than the square-mesh net, especially for more enthusiasm in stadium and also for better visual effects in TV.

Keywords: flexible nets, incompressibility, large displacement, dynamics, sags, slack.

1. INTRODUCTION (MOTIVATION)

In the last FIFA World Cup 2002 in Japan and South Korea, the hexagon-mesh goal nets were used for the first time in the world cup history. There was and is a very clear reason, why the hexagon-mesh net is better than the square-mesh net: For all enthusiastic soccer funs watching the world event in stadium and in TV, the hexagon nets will present the most exciting and dramatic images at the moment of highlights. This ICCM paper is concerned with the static

G. R. Liu et al. (eds.), Computational Methods, 1903–1909.

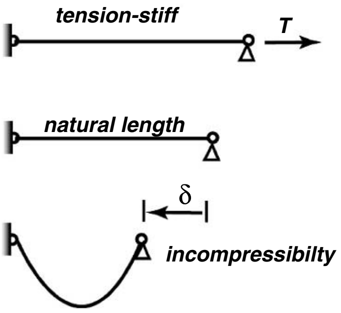


Figure 1. Incompressibility.

and dynamic behaviour of extremely flexible and incompressive goal nets for soccer, which have been so far not frequently treated for engineering purposes. Modelling and computational techniques are described for net mechanics in sports entertainment [1, 2]. The special emphasis is placed upon the different behaviour of hexagon- and square-mesh goal nets.

2. NONLINEAR EFFECTS

The static and dynamic behaviour of flexible nets is in fact a geometrically highly nonlinear problem including large nodal displacements. The singular stiffness matrix, for example, may be avoided by the geometric stiffness matrix activated by initial stresses. The stability problem and dynamic instability may also be observed in the net behaviour, when the equilibrium solution will jump from unstable equilibria subject to compression to the tension-only stable equilibrium solution. For all these computational issues, innovated computational techniques are nowadays well established [3–8].

One computational problem inherent in flexible nets is, however, their incompressibility, as illustrated in Figure 1. Starting from natural length L, the net segment in Figure 1 has no stiffness on compression side and the distance of two edges (nodes) will simply be shortened without any resistance and represent a sagging segment. On tension side in the meantime, a number of braided fibres in zigzag lines (Figure 2) stay initially slack in the textile structure and be stretched by $\Delta_o = \varepsilon_o L$ without being stressed. Beyond this limit of initial slackness, the relationship between axial force N and elongation Δ are linear elastic, as all fibres in straight lines are ready to resist the stretching force. The modelling of this material property of flexible nets is the major concern of the present study [1, 2].

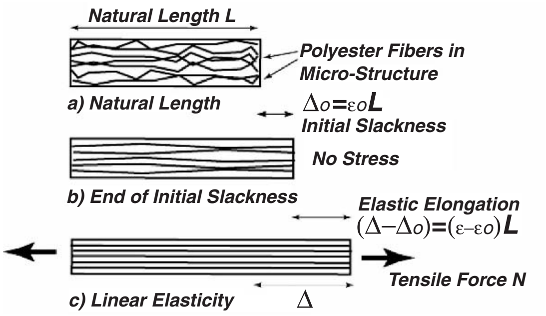


Figure 2. Initial slackness.

3. COMPRESSION-FREE MATERIAL LAW WITH INITIAL SLACKNESS (FIGURE 3)

The incompressibility of flexible nets may directly be considered in the constitutive equation. For the relationship between axial force N and elongation Δ for an element, we assume an analytical function for $N(\Delta)$ as;

$$\frac{N(\Delta)}{EA} = \left(\frac{1}{2}\right)\left\{\left(\frac{\Delta}{L}-\varepsilon_o\right)+\sqrt{\left(\frac{\Delta}{L}-\varepsilon_o\right)^2+4\left(\frac{T_0}{EA}\right)^2}\right\} \tag{1}$$

in dimensionless form with axial stiffnessEA for element cross-section, natural element length L and small initial tension $T_0 = N(0)$ for $\Delta = 0$, respectively $\varepsilon_o$designates the initial slackness in Figure 3, in which the function curves

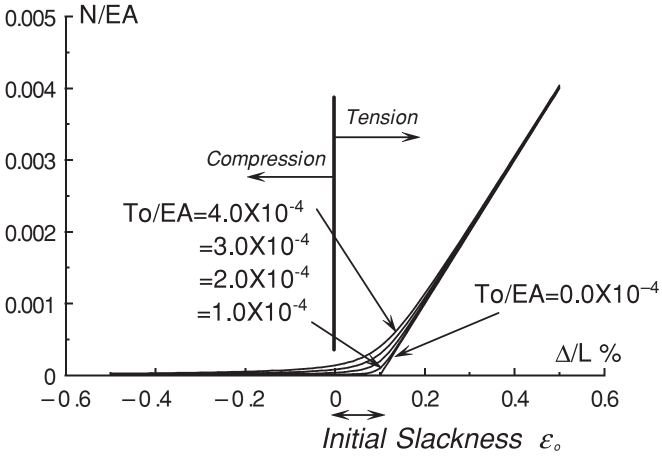


Figure 3. Compression-free Material Law.

for different values of T_0 with a fixed value for $\varepsilon_o = 0.001$ are illustrated. The constitutive curve for $T_0 = 0$ is bi-linear and the first derivative of the function$N(\Delta)$ with respect to Δ is discontinuous at $\varepsilon = \varepsilon_o$. This may cause a numerical trouble during equilibrium iteration in the vicinity of $\varepsilon = \varepsilon_o$. A small initial stress $T_0 \neq 0$ is, therefore, necessary to make the equilibrium iteration stable near the limit of initial slackness.

4. STATIC AND GEOMETRIC IMPERFECTION FOR STATICS (FIGURE 4)

The second idea is possible in the structural level. We may discretize a single net segment into more than one truss element, which is equally stiff against compression and tension, and introduce a small static imperfection. The static nonlinear equilibrium equations will then be rewritten as follows:

$$\boldsymbol{R}(\boldsymbol{u}) - p\boldsymbol{e} - \boldsymbol{f} = 0 \tag{2}$$

Here, $\boldsymbol{R}(\boldsymbol{u})$ is the internal resistance depending upon nodal displacement vector $\boldsymbol{u}$ and the load parameter p will represent, together with reference load vector $\boldsymbol{e}$, proportional external loads. The static imperfection may be considered in the constant vector $\boldsymbol{f}$. Figure 4a depicts an illustrative example, which will be subject to tension (left) and compression (right) for a simple two-element discretization, when a horizontal design load is applied in the middle. To avoid this unacceptable solution, the refined discretization in Figure 4b and the small load imperfection in Figure 4c are effective to make the iteration convergent to the sagging and compression-free configuration in Figure 4d. To activate the geometric stiffness matrix, at the beginning of iteration, it is recommended to assume an initial displacement vector $\boldsymbol{u}_0$ such that;

$$\boldsymbol{u}_0\boldsymbol{f}^T \cong |\boldsymbol{u}_0|\ |\boldsymbol{f}\,| \tag{3}$$

and to superpose $\boldsymbol{u}_0$ on input geometry, as shown in Figure 4c. This is due to the reason, that the assumed configuration in Figure 4c should stay compression-free for two vectors $\boldsymbol{u}_0$ and $\boldsymbol{f}$ in almost same directions. Care must be taken, however, that the static imperfection in $\boldsymbol{f}$ need not always coincide with dead loads of the net material. In the computed equilibrium configuration, as shown in Figure 4d, all static imperfection is iterated away in the tension-stiff left half, while the right half will be slack in a local catenary due to the assume static imperfection.

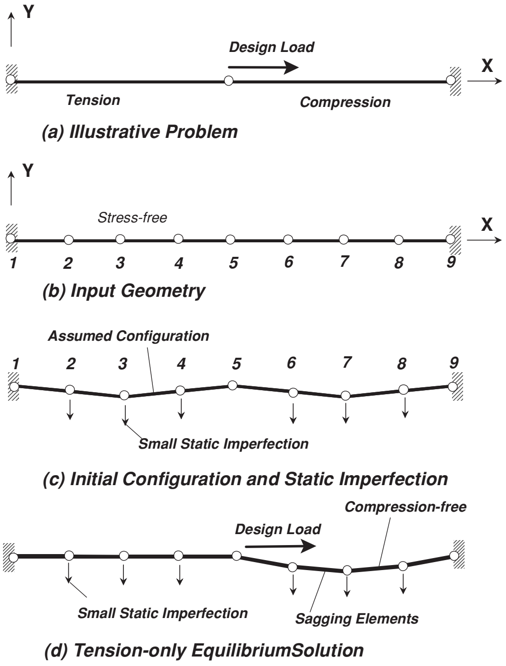


Figure 4. Static and geometric imperfection.

5. EQUATIONS OF MOTION

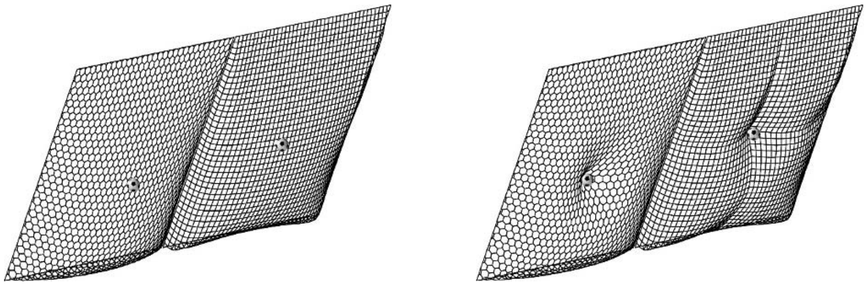
The static solution may also be computed by time-integrating the equations of motion;

$$\boldsymbol{M}\ddot{\boldsymbol{u}} + \boldsymbol{C}\dot{\boldsymbol{u}} + \boldsymbol{R}(\boldsymbol{u}) - \boldsymbol{f} = \boldsymbol{Q}(t) \tag{4}$$

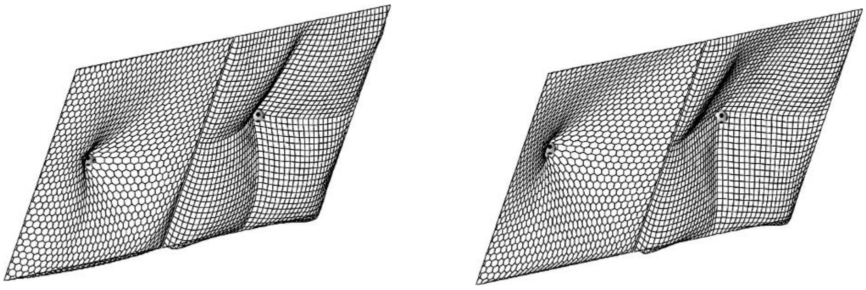
with mass matrix $\boldsymbol{M}$, acceleration vector $\ddot{u}$, damping matrix $\boldsymbol{C}$, velocity vector $\dot{\boldsymbol{u}}$. The static loads may be considered as a step-function in dynamic external load vector $\boldsymbol{Q}(t)$. Compared to the arc-length control method to solve Equation (2), the dynamic approach based upon Equation (4) has the advantage that we may exclude unacceptable equilibrium points subject to compression to attain a tension-only stable equilibrium solution, when the motion is sufficiently damped. Equation (4) is also used to study the dynamic behaviour of the goal net at the moment of ball shooting.

6. NUMERICAL EXAMPLE (FIGURE 5)

The hexagon- and square-mesh nets, spanned over a mini-goal (steel frame: 4.75 m × 4.75 m) inclined to the ground surface at an angle of 67°, are compared



(a) *t = 0.00 sec* (b) *t = 0.05 sec*



(c) *t = 0.10 sec* (d) *t =0.15 sec*

Figure 5. Shooting a ball into the goal net.

to each other. The material parameters are $EA = 7539.82$ N, mass density $\rho = 960$ kg/m^3, $T_0 = EA \times 10^{-6}$ and $\varepsilon_o = 0.003$. The each side of the hexagon and square are 0.095 and 0.125 m in length, respectively.

We first apply the static dead weight to both types of nets and the deformed hexagon net surface for $t = 0.0$ sec is smother and softer, compared to the deformed configuration of the square net. The significant difference in the deformed surfaces may be identified especially near two corners at bottom. When the net catches the ball, the shock wave propagates concentrically around the ball in the hexagon net and the isotropic property of the hexagon net is clearly illustrated in the sequence of dynamically deformed configuration for each time step. The square net, on the other hand, transfers the dynamic loads in horizontal and vertical directions, namely in directions of net strings and we clearly see the orthotropic property and stiff behaviour of the square net.

7. CONCLUSIONS

The static and dynamic behaviour of flexible and incompressible goal nets has been studied from the viewpoint of computational mechanics and sports entertainment. The special emphasis is placed upon the best image of the highlights of soccer games, which will receive much attention in the stadium and in TV from all over the world. Compared to the conventional stiff square-mesh nets, the hexagon-mesh nets are softer and more flexible and their larger dynamic deformation will more fever the soccer funs at the moment of goals. "The hexagon mesh is the best for the goal net to make a big world game more enthusiastic and dramatic!"—This is the conclusion of the present ICCM paper.

REFERENCES

1. F. Fujii, H. Noguchi and P. Osterrieder (2003), *Static and Dynamic Large Displacement Response of Flexible Nets*, IASS-APCS, Taipei.
2. F. Fujii and H. Noguchi (2004), *Flexible Nets in Statics and Dynamics*, ECCOMAS2004, 24–28 July, Jyväskylä, Finland, CD-ROM, Proceedings Volume I, ISBN 951-39-1868-8.
3. P. Broughton and P. Ndumbaro (1994), *The Analysis of Cable & Catenary Structures*. Thomas Telford, 1994.
4. Architectural Institute of Japan (2001), *Guideline for Numerical Analysis of Spatial Structures*, Tokyo.
5. Japan Society for Civil Engineers (1999), *Basis and Application of Cable and Space Structures,* Steel Structure Series 11, Tokyo.
6. S. Palkowski (1990), *Statik der Seilkonstruktionen – Theorie und Zahlenbeispiele.* Springer-Verlag, Berlin.
7. M. Irvine (1992), *Cable Structures*. Dover Publication, Inc., New York.
8. I. Fried (1982), Large deformation static and dynamic finite element analysis of extensible cables. *Computers and Structures*, 15, 3, 315–319.

AERODYNAMIC PROPERTIES OF SOCCER BALL

Y. Nakabayashi[1] and S. Yoshimura[2]
[1] *Department of Computational Science and Engineering, Toyo University, Japan*
[2] *Institute of Environmental Studies, The University of Tokyo, Tokyo, Japan*

Abstract In this paper, the authors analyse the aerodynamic properties of soccer ball. Using ADVENTURE System, the finite element flow analysis around soccer ball is performed to figure out the coefficient of drag of soccer ball, the critical Reynolds number of drag crisis, the track of curveball.

Keywords: incompressible viscous flow analysis, ADVENTURE system, large-scale parallel processing, soccer ball, aerodynamic properties.

1. INTRODUCTION

The Soccer ball has hexagonal and pentagonal surfaces and grooves between them. Therefore, it is supposed to show different aerodynamic properties from the smooth sphere. The target of this study is to figure out (i) the coefficient of drag, lift and swing in various Reynolds numbers, (ii) the critical Reynolds number of drag crisis, (iii) the tracks of curveball. Our final target is to build the database which contains the information about various aerodynamic properties of soccer ball, and apply it to free kick support system in real soccer game.

In this paper, we show the process of finite element mesh generation of soccer ball. Then flow analyses around soccer ball are performed in various Reynolds number. After evaluating the critical Reynolds number, we estimate the track of soccer ball when it is kicked in certain direction, velocity and rotation.

G. R. Liu et al. (eds.), Computational Methods, 1911–1917.

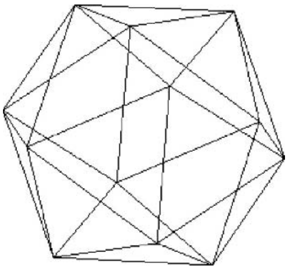


Figure 1. Regular icosahedron.

2. MODELLING AND MESH GENERATION

For generating finite element analysis model of the flow around soccer ball, the authors used pre-processing modules of ADVENTURE System. The process of mesh generation has the following six steps.

Step 1. Calculate the coordinates of all (12) vertices of regular icosahedron using Pythagoras's theorem (Figure 1).

Step 2. Generate new vertices on all the edges, which vertices divide the edges into three same length parts. The Fullerene C_{60} type structure is obtained by connecting the new vertices (Figure 2).

Step 3. The Fullerene C_{60} type structure has 20 hexagons and 12 pentagons. Divide hexagons into six triangle patches, and pentagons into five triangle patches (Figure 3).

Step 4. Generate more fine triangle patches using the quad-tree algorithm. In this study we use the quad-tree algorithm four times, then we get 23,042 vertices and 46,080 triangle patches (Figure 4).

Step 5. The radius of soccer ball is 110 mm. Shrink the radius of edges, which correspond the Fullerene C_{60}, into 108.3 mm for generating the grooves (Figure 5).

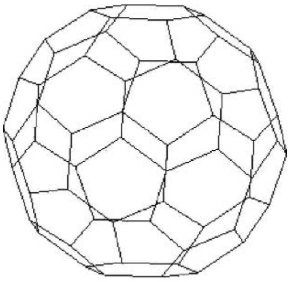


Figure 2. Fullerene (C_{60}).

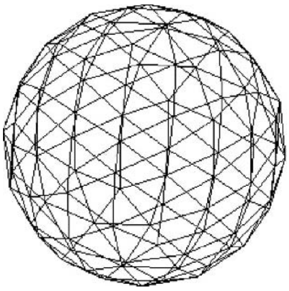


Figure 3. Triangle patch.

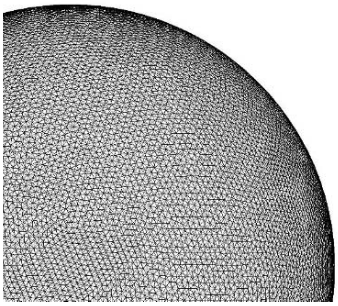


Figure 4. Fine patch.

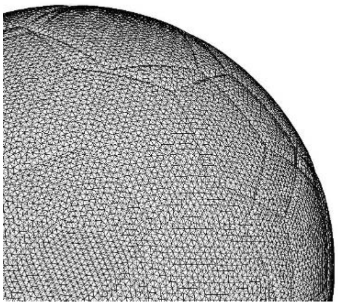


Figure 5. Patch with grooves.

Step 6. Using ADVENTURE TetMesh (tetrahedral mesh generator), the finite element model of flow around soccer ball is obtained (Figure 6). This model has 237,292 nodes and 1,410,488 elements.

3. ADVENTURE FLUID

ADVENTURE Fluid, which has been developed in the ADVENTURE project, is a general purpose parallel 3D finite element solver for the

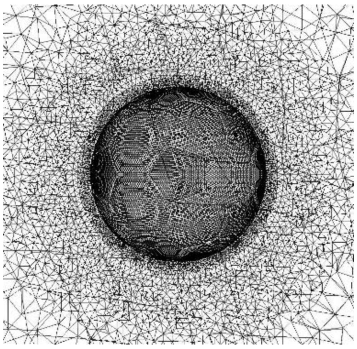


Figure 6. FE model.

incompressible thermal-fluid analysis (a manual in the homepage of the ADVENTURE project). Some realistic problems have been solved by using this code [1].

In the tetrahedral P1–P1 element version of the code, which will be used in this study, the SUPG (Streamline-upwind/Petrov–Galerkin) method and PSPG (Pressure-stabilized/Petrov–Galerkin) method are implemented as below:

$$\begin{aligned}&\int_{\Omega} w_i \rho \left(\frac{\partial u_i}{\partial t} + u_j u_{i,j} - f_i\right) d\Omega + \int_{\Omega} w_{i,j}\sigma_{ij}\, d\Omega + \int_{\Omega} q u_{i,i}\, d\Omega \\ &+ \sum_e \int_{\Omega^e} \tau_{\mathrm{SUPG}}\, u_j w_{i,j} \left[\rho\left(\frac{\partial u_i}{\partial t} + u_k u_{i,k} - f_i\right) - \sigma_{ik,k}\right] d\Omega \\ &+ \sum_e \int_{\Omega^e} \tau_{\mathrm{PSPG}}\, \frac{1}{\rho} q_i \left[\rho\left(\frac{\partial u_i}{\partial t} + u_k u_{i,k} - f_i\right) - \sigma_{ik,k}\right] d\Omega = \int_{\Gamma_h} w_i h_i\, d\Gamma \qquad (1)\end{aligned}$$

The Crank–Nicolson method is employed for the implicit time integration. A system of linear equations for the velocity field and pressure field is solved by using diagonal scaling Bi-CG STAB method.

4. RESULTS OF FLOW ANALYSIS

Finite element flow analyses are performed using PC cluster (dual Xeon 2.0 GHz, 64 nodes). The Reynolds numbers of these analyses are set to 10, 100, 1000, 10,000 and 100,000. Time increment of these analyses is set to 0.002. Total analysis time is 10.0 (i.e., 5000 time steps). Figure 7 shows the streamline around the soccer ball in the case of Reynolds number 1000. The variation of coefficient of drag by Reynolds number is shown in Figure 8, which indicates

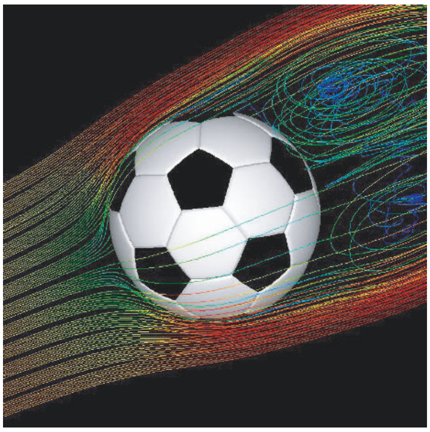


Figure 7. Streamline (Re = 1000).

that the coefficient of drag of soccer ball is slightly lower than that of sphere especially in the high Reynolds number.

5. TRACK OF SOCCER BALL

After the flow analysis, the track of soccer ball is simulated using developed free kick simulator (Figure 9). The simulator requires many parameters, which are shown in Figure 10. The coefficient of drag, lift and swing are the most important parameters in this simulator for evaluating the exact track of

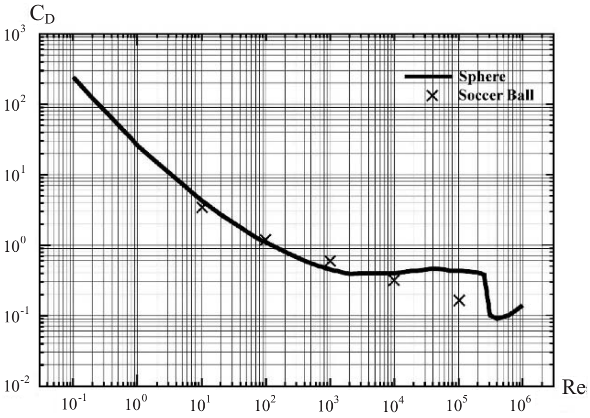


Figure 8. Coefficient of drag.

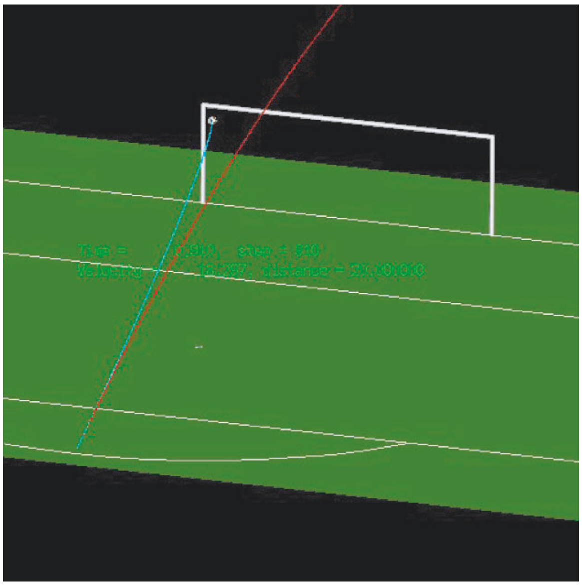


Figure 9. Free kick simulator.

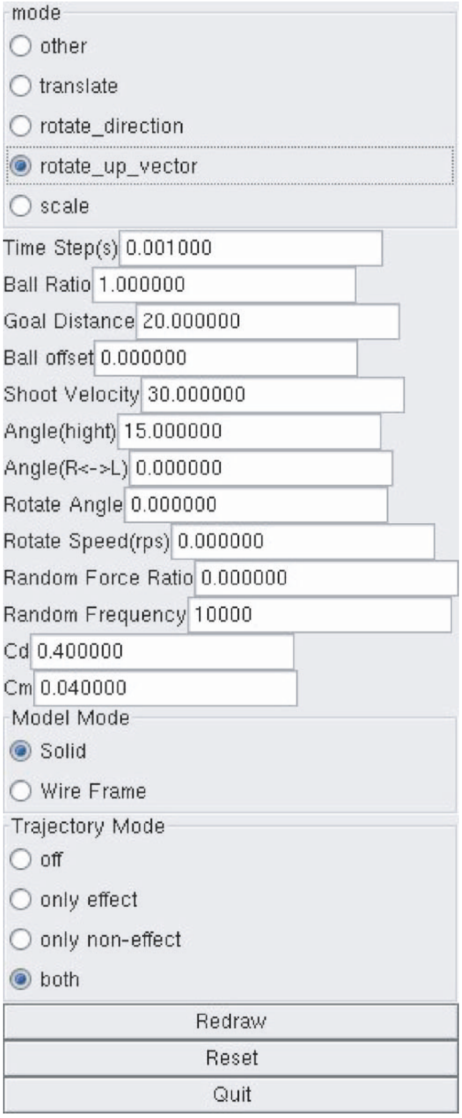
soccer ball. At this time, there are not enough database concerning such parameters, therefore the approximated simulations are performed as the test case.

6. CONCLUSIONS

The authors generated the finite element model around the soccer ball for evaluating aerodynamic properties of soccer ball. Large scale flow analyses were performed for estimating the coefficient of drag, lift and swing using ADVENTURE Fluid. The result showed that the coefficient of drag of soccer ball was slightly lower than that of sphere. The authors also simulated the track of soccer ball using computed parameters. For applying this simulator to real soccer game, many more flow analyses will be required.

ACKNOWLEDGEMENT

The free kick simulator was developed by Dr. Y. Mimura. Software modules used in this study have been developed in the ADVENTURE project.



mode
○ other
○ translate
○ rotate_direction
● rotate_up_vector
○ scale

Time Step(s) 0.001000
Ball Ratio 1.000000
Goal Distance 20.000000
Ball offset 0.000000
Shoot Velocity 30.000000
Angle(hight) 15.000000
Angle(R<->L) 0.000000
Rotate Angle 0.000000
Rotate Speed(rps) 0.000000
Random Force Ratio 0.000000
Random Frequency 10000
Cd 0.400000
Cm 0.040000

Model Mode
● Solid
○ Wire Frame

Trajectory Mode
○ off
○ only effect
○ only non-effect
● both

Figure 10. Controller.

REFERENCES

1. T. Miyamura, Y. Nakabayashi, H. Kawai and S. Yoshimura (2004), Parallel finite element analysis of Lake Inawashiro. The International Conference on Computational Methods (ICCM2004), Singapore.
2. Home page of ADVENTURE project, http://adventure.q.t.u-tokyo.ac.jp.
3. Manual of ADVENTURE Fluid, http://adventure.q.t.u-tokyo.ac.jp.
4. G. Yagawa, Y. Nakabayashi and H. Okuda (1997), *Parallel Computing*, 23, pp. 1365–1377.

THE LEAKAGE ANALYSIS IN A TWIN-SCREW SUPERCHARGER BY USING AN INTEGRATED CAD/CFD THREE-DIMENSIONAL MODEL

H.F. Li, J.Y. Tu, A. Subic and B. Abu-Hijleh
School of Aerospace, Mechanical and Manufacturing Engineering, RMIT University,
Bundoora East Campus, Melbourne, VIC 3083, Australia

Abstract The performance of a twin-screw supercharger is influenced more by the internal air leakages than by any other thermo-fluid aspect of its behaviour. The distribution of leakage through the various leakage paths within the machine is, thus, important for the improvement of the performance. The numerical analysis is conducted using a Computational Fluid Dynamics (CFD) software commercially available, FLUENT. The main innovation in this research is to integrate this advanced CFD modelling technique, which is used for modelling complex airflow behaviour in the complex supercharger passages, with an advanced CAD module, which is used for design of new supercharger rotors and manifolds. This has allowed creating a new innovative fully integrated CAD/CFD supercharger design environment, which can address all aspects of fluid and geometry requirements for the future supercharger designs. This approach will insure the production of the highest quality designs at the lowest manufacturing cost and in the shortest turn around time.

Keywords: twin-screw supercharger, optimization, CAD/CFD, leakage.

1. INTRODUCTION

A twin-screw supercharger is a positive displacement rotary compressor, the working cavity of which is enclosed by the casing bores, casing end plates and the helical surfaces of the male and female rotors. Due to the geometry of the meshing parts and the need for clearances between them, the supercharger has several leakage paths as follows: across the contact line between the male and female rotors, across the sealing lines between the rotor tips and casing, through the blow hole, and through the clearance between the end plate and the rotor end face at the discharge end [1]. As the performance of a twin-screw

G. R. Liu et al. (eds.), Computational Methods, 1919–1923.

supercharger is influenced more by the internal air leakages than by any other thermo-fluid aspect of its behaviour, it is important to quantify the leakage rate through each leakage pathway, and, the percentage by which it can reduce the efficiencies for the purpose of prioritizing design procedures in general and for improving the rotor lobe profile.

The field of engineering design optimization using computational fluid dynamics (CFD) can improve the performance and quality of existing products and potentially lead to novel designs, which can be crucial in maintaining competitiveness in world markets. However, engineering design optimization can be very time-consuming when high-fidelity analysis models are involved. In this study, two essential elements are required in the optimization process: these are geometry design using computer-aided design (CAD), and CFD simulation and analysis [2]. These two elements construct a cycle for optimization design. Virtually, all computer-based design tasks commence with the use of CAD systems to create detailed geometrical models. These models serve as the point of departure for analysis tool, CFD. By integrating these essential elements, the optimization design process allows easy access to remote computational resources and data archive capabilities. By repeating this cycle, the optimization design is achieved.

2. ANALYSIS PROCEDURE

As Figure 1 shows, this study effectively uses such an innovative CAD/CFD interaction approach to achieve engineering optimization design.

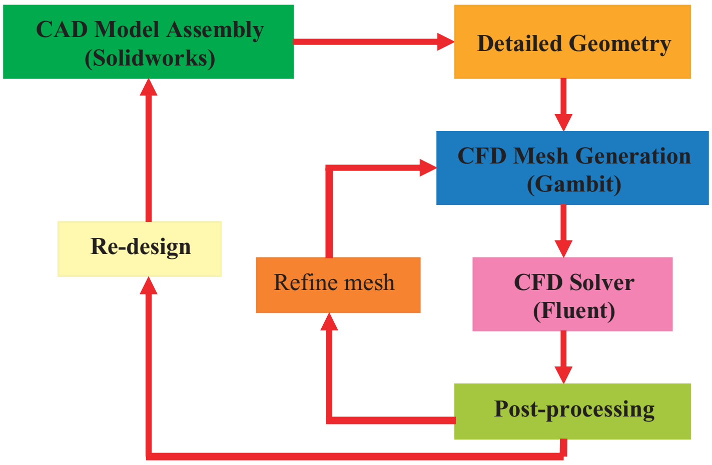
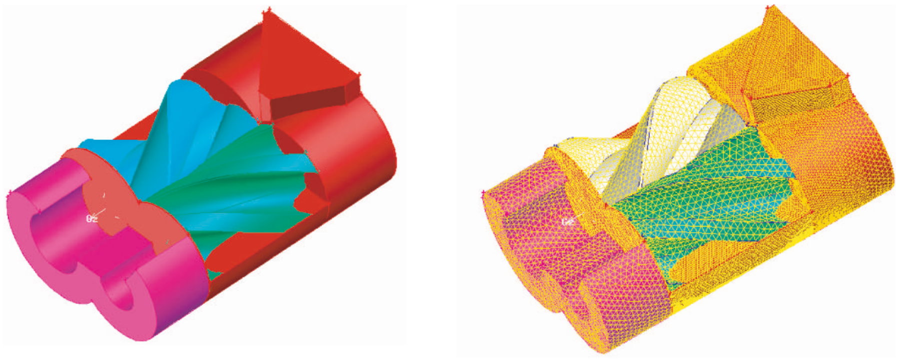


Figure 1. CAD/CFD integration and optimization process.



(a) The 3D solid model **(b) The 3D tetra-mesh**

Figure 2. The 3D model of a supercharger.

2.1 Creation of a 3D solid model

As shown in Figure 2a, in the development process, the author creates a 3D detailed solid model that accurately represents the geometry of the supercharger. By making quite accurate solid model in Solidworks, few works have been done during the procedure that CAD data can be easily converted for CFD mesh generation, which thus accelerates the whole optimization procedure. Currently, CAD and CFD are frequently interfaced via standardized interface file formats, IGES (Initial Graphics Exchange Specification) for 2D model. Since the IGES format contains surfaces only, the 3D solid data cannot be exchanged completely [3]. For 3D modelling, ACIS files would be the most effective format.

2.2 Mesh generation

With the new method, the time required for mesh generation for fluid analysis has been dramatically reduced through the use of tetra-mesh technology. Hexahedral mash generally provides a higher degree of analysis accuracy than tetra-mesh for fluid analysis. However, analyses with the new method are not independent procedures like those seen with the conventional procedures, but a tool completely integrated in the entire development flow as evaluated from a macroscopic viewpoint. In order to make the tool adequate for the entire process, and because the emphasis is on making the generation of meshes much easier, a tetra-mesh is employed, as shown in Figure 2b.

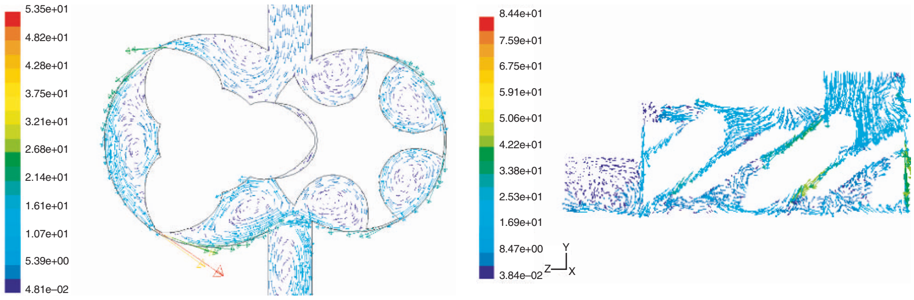


Figure 3. Leakage flow behaviour in the supercharger.

2.3 Leakage analysis

Leakage analysis is performed using the Fluent CFD package, which is employed to simulate the flow passage through the supercharger. The flow is simulated as a 3D incompressible air using the Realizable k-ε turbulence model [4].

Figure 3 shows the leakage flow behaviour inside the supercharger. Table 1 gives the leakage flow rates through each leakage path in this model. The results show that Path 2 and Path 3 have a very important influence on the leakage distribution. Consequently, the improvement of a profile should concentrate on the following two aspects: reducing the blowhole area by an asymmetric profile and reducing the leaking rates through Path 2 by optimizing the wrap angle of rotors.

2.4 Optimization

Making use of the above ideas and the optimization process, a preliminary improved profile has been developed. Compared with the original design, the total leakage has reduced by 1.5%. Thus in this case, the change of rotor profile does reduce the leakage rate; 1.2 of the 1.5% improvement is obtained by

Table 1. Leakage flow rates through each leakage pathway in a supercharger.

Leakage pathways	Leakage flow rate (kg/s)	%
Contact line (Path 1)	0.0015528	7.1
Sealing lines (Path 2)	0.0082927	37.8
Blowhole (Path 3)	0.0065058	29.6
End face (Path 4)	0.0055924	25.5

reducing the blowhole area and 0.3 is obtained by increasing the rotor warp angle. Repeated applications of optimization procedure can further improve the performance of the supercharger.

3. CONCLUSIONS

In summary, the main innovation in this research is to integrate advanced CFD modelling techniques for the modelling of the complex airflow passages throughout the complex supercharger system with an advanced CAD module for supercharger rotors and manifolds designs that can be used to directly interface with the CFD program. This will create a new innovative fully integrated CAD/CFD supercharger design environment, which addressed all aspects of fluid and geometry requirements for the twin-screw supercharger. The success of the research program will be demonstrated by an optimized design of the supercharger. This approach will insure the production of the highest quality designs at the lowest manufacturing cost and in the shortest turn around time.

REFERENCES

1. J.S. Fleming, Y. Tang and H. Anderson (1994), Optimization techniques applied to the design of a refrigeration twin screw compressor. In: *Proceedings of International Compressor Engineering Conference*, pp. 641–646.
2. M. Obi, N. Takei and N. Matsuura (2000), Integration of CFD and thermal stress analysis for turbochargers. *IHI Engineering Review*, 33, 2, pp. 31–35.
3. M. Obi, A. Ohkita, N. Takei and N. Matsuura (1999), Application of 3D-CAD to turbocharger development. *IHI Engineering Review*, 32, 4, pp. 113–117.
4. Fluent Inc. (2001), FLUENT 6.0 User's Guide.

A MULTI-AGENT TRAFFIC AND ENVIRONMENTAL SIMULATOR AND ITS APPLICATION TO THE ANALYSIS OF TRAFFIC CONGESTION IN KASHIWA CITY

Yutaka Nakama, Shinobu Yoshimura and Hideki Fujii
University of Toyko, 7-3-1 Hongo, Bunkyo-ku, Tokyo 1113-8586, Japan

Abstract Traffic problems are now very serious in various cities. Traffic simulators are expected to be useful for traffic system analysis. This paper describes research and development of the advanced traffic simulator using a multi-agent approach which is named MATES. Here each car is modelled as an intelligent agent with memory and preference. They can reflect individual characteristics to their driving. The environment surrounding a car agent consists of other cars, roads, traffic signals and pedestrians. We've implemented some algorithms needed for real world problems, and proposed a framework to define some types of car agent, sedan, track and bus. These cars show different characteristic of acceleration and deceleration performance, route inquiry and the behaviour when getting close to bus stop. The MATES is applied to analyse traffic congestion in Kashiwa City in Japan.

Keywords: traffic engineering, traffic simulator, multi-agent system, traffic congestion.

1. INTRODUCTION

Traffic simulators are expected to be useful for traffic system analysis. They are classified into two models. One is a macro simulator, and the other is a micro model simulator. The macro model is usually expressed by macro equations describing traffic flow. In the micro model, the action of each vehicle is modelled more precisely, and the traffic flow is expressed as their aggregation.

The macro model simulation is often used when an analysis area is large. It is useful for some particular cases, but it is not useful for some other cases, such that many types of vehicles exist. For example, we have to treat different types of decision-making for driving route selection, and each car has different

G. R. Liu et al. (eds.), Computational Methods, 1925–1830.

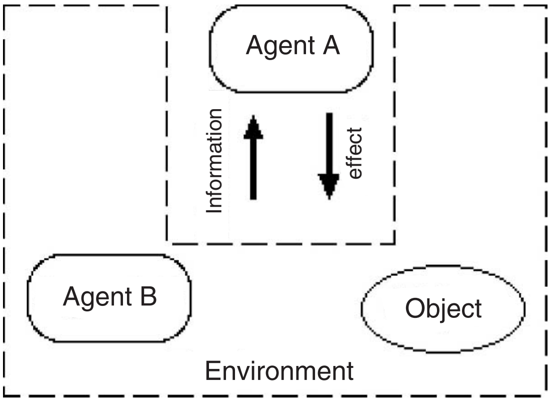


Figure 1. Environment and agent.

characteristic about acceleration and deceleration performance. Moreover, it is difficult to use traffic information changing dynamically.

In the micro model simulation, each car is able to have individual characteristics to their driving. Therefore, simulations by the micro model have been used in recent years. The present authors are developing a kind of micro model traffic simulator called MATES (Multi-Agent based Traffic and Environment Simulator) based on the agent with autonomy, and moreover we are trying to apply MATES to analyse the traffic phenomenon (Kashiwa City in Japan), as an example of real world problems.

2. MULTI-AGENT SYSTEM

A multi-agent system is a system in which some distributed programs perform problem solution in cooperation in an environment. Here an agent is a composition element. The agent perceives the environment by its own sensors, and affects the environment through its own effector. In the multi-agent approach, many agents are released into the environment, and a whole complex phenomenon is evaluated as the consequence of their interaction.

A conceptual figure of the multi-agent system developed in this research is shown in Figure 1. Each agent takes the action considered to be locally the best based on the information acquired in its state. For each agent, other agents are regarded as a part of environment.

Figure 2 shows a conceptual model of the agent and the environment implemented in MATES. Here each car and pedestrian perceive various kinds of information from the environment, and determine their action. The pile of the interaction of them is regarded as the whole traffic phenomenon. We design these (car and pedestrian) as an agent. On the contrary, neither a road nor an

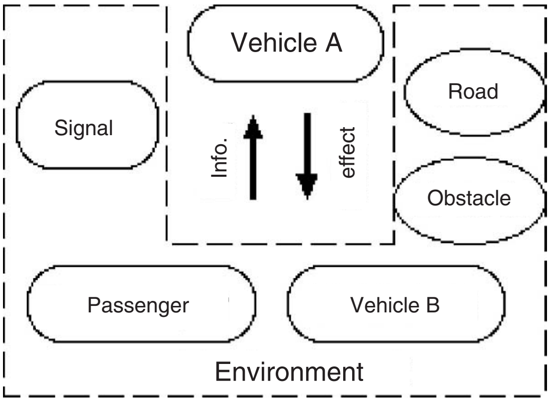


Figure 2. Environment and agent in MATES.

obstacle is modelled as an agent in order not to have active influence in the environment. Then they are simply regarded as objects.

The advantage of applying the multi-agent approach to a traffic simulator is summarized below.

- *Precision:* A situation can be expressed more precisely than a macro model simulator.
- *Diversity:* The characteristic of a driver or a car can be treated reasonably well.
- *Dynamic:* Decision-making portions such as route selection can be treated dynamically and automatically.
- *Existence of information use:* A driver can use various kind of information retrieved from the environment.
- *Extensibility:* The reusability if the agent itself is high. It is also easy and flexible to add a new function into the agent and to change data structure.

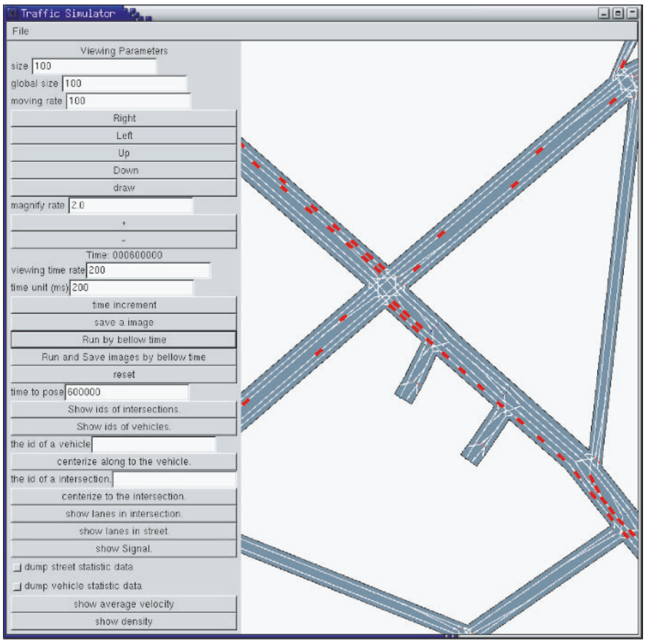
3. CALCULATION MODEL

Taking one of typical medium-sized cities in Japan, i.e. Kashiwa City in Chiba Prefecture as an example, MATES is applied to perform the following analyses:

1. Influence in traffic congestion by the difference in route selection behaviour of drivers.
2. Traffic congestion relief by signal control.
3. Influence on traffic congestion by the different characteristics of vehicles.

The road map used in this analysis is shown in Figure 3. Some general views of MATES are shown in Figures 4 and 5.

Figure 3. Road map of Kashiwa City.



size 100
global size 100
moving rate 100
magnify rate 2.0
Time: 000600000
viewing time rate 200
time unit (ms) 200
time to pose 600000
the id of a vehicle
the id of a intersection.
☐ dump street statistic data
☐ dump vehicle statistic data

Figure 4. General view of MATES.

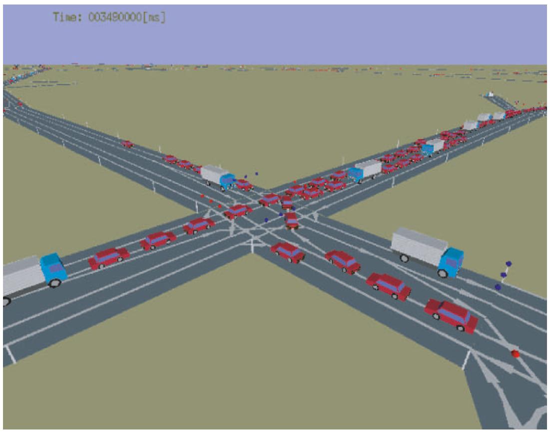


Figure 5. 3D visualization.

4. INFLUENCE ON TRAFFIC CONGESTION BY THE DIFFERENT BEHAVIORS OF ROUTE SELECTION

In MATES, each vehicle has a start point and a destination, and route selection is performed independently. We investigated the influence on traffic congestion by the different behaviours of route selection. Here we consider the following two different route selection methods.

1. The route with the shortest distance is chosen.
2. The route with the shortest travel time is chosen.

5. TRAFFIC CONGESTION RELIEF BY SIGNAL CONTROL

Next, we conducted a numerical experiment varying signal control. Here the data of the signal control actually performed in Kashiwa City is employed.

Influence on the traffic congestion by the different characteristics of vehicles. Unlike the environment which consists only of usual passenger cars, buses on regular routes and large-sized tracks are intermingled in an actual traffic network. Such a heterogeneous situation may influence on certain traffic congestion, because of the difference in acceleration or a slowdown performance of each vehicle. Moreover, when a bus on a regular route stops regularly at a stop, it may also influence traffic congestion.

6. CONCLUSIONS

We have developed a multi-agent traffic and environment simulator MATES. Taking Kashiwa City, some numerical experiments were conducted. We show that the micro model traffic simulator could be adapted for practical problems.

REFERENCES

1. S. Yoshimura, H. Nishikawa and S. Moriyasu (2004), Development of multi-agent traffic and environmental simulator MATES. *Journal of the Japan Society of Simulation Technology*, 23, 3 (in print).

NUMERICAL INVESTIGATION OF COUPLED TRANSPORT OF IONS AND IONIC SOLUTION IN CORNEA AND THEIR INFLUENCE ON CORNEAL HYDRATION

Long-Yuan Li and Brian Tighe
School of Engineering and Applied Science, Aston University, Birmingham, UK

Abstract This paper presents a numerical study on the transport of ions and ionic solution in human corneas and their influences on the corneal hydration. The transport equations for ions and ionic solution within the cornea are derived based on transport processes in electrolytic solutions, whereas the transport across epithelial and endothelial membranes is described by using phenomenological equations derived on the thermodynamics of irreversible processes. Numerical examples are given for a typical human cornea from which some important features are highlighted.

Keywords: corneal hydration, electrolyte, ionic transport, flow, thermodynamics.

1. INTRODUCTION

Corneal transparency is dependent on the regulation of normal stromal hydration. Isolated stromal tissues swell freely in isotonic saline and become opaque. However, in vivo corneas remain thin and clear. How the limiting membranes control the stromal hydration and thus maintain the hydrophilic stroma in a state of relative deturgescence is one of the most perplexing questions in corneal physiology.

Modelling of corneal swelling and its interaction with environment is critical to our understanding of corneal function and the potential effect of ocular therapies on that function. Considerable efforts have been made in past decades for developing models that can simulate the whole corneal transport system and describe how the stroma and limiting layers interact to control stromal hydration [1]. These models can be generally divided into two categories. One is the flow model [2–5] in which the increase or decrease of corneal thickness at a particular

G. R. Liu et al. (eds.), Computational Methods, 1931–1936.

point was purely due to the volume difference between the inflow and outflow of water at that point. The driving force of water flow in the corneal stroma was assumed to be the gradients of hydrostatic and osmotic pressures; the former was expressed in terms of the swelling pressure of corneal stroma and the latter was linked to the concentrations of solutes involved. The transport of the solutes in the corneal stroma was treated in a similar manner. The flows of water and solutes across epithelial and endothelial membranes were described by using the phenomenological equations derived on the thermodynamics of irreversible processes. The other is the triphasic model based on the poroelastic theory [6] in which the solid matrix is regarded as a linear isotropic elastic material and the corneal swelling is considered to be governed by the transport of ions in the electrolyte within the poroelastic medium controlled by the coupled electrochemical and mechanical processes [7]. The volume change of the cornea is thus related to both the fluid flow and the elastic deformation of the corneal matrix.

2. TRANSPORT MODEL

The transport model to be developed here is similar to that presented by Klyce and Russell [4] and Li et al. [5], which is a one-dimensional model, with stroma in the middle and epithelium and endothelium on the two sides (Figure 1). However, since the solution considered here is a multi-component one, the governing equations for the transport of ions and ionic solution are derived on the transport processes in electrolytic solutions. The governing equations describing the transport of ionic solution and ionic species in corneal stroma are expressed as [5, 8]

$$\frac{\partial W_f}{\partial t} = -\frac{\partial}{\partial \xi}\left(\frac{\rho_f J_f}{\rho_d}\right) \quad \text{and} \quad \frac{\partial}{\partial t}\left(\frac{W_f C_k}{\rho_f}\right) = -\frac{\partial}{\partial \xi}\left(\frac{J_k}{\rho_d}\right) \tag{1}$$

where W_f is the mass of the solution in unit mass of dry tissue, t is the time, ρ_f

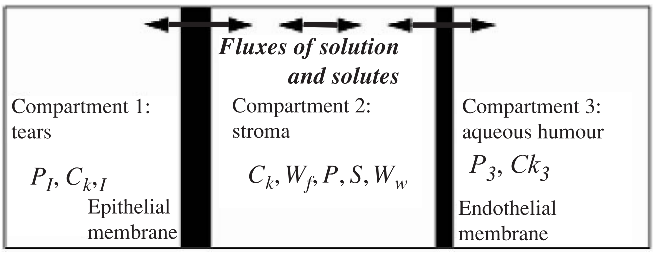


Figure 1. A two membrane-three compartment one-dimensional flow model for the cornea.

is the density of the solution, J_f is the volume flux of the solution flowing in unit time through unit area in the thickness direction, ξ is the thickness coordinate of the dry tissue, ρ_d is the density of the dry tissue, C_k is the concentration of species k and J_k is the molar flux of species k flowing in unit time through unit area in the thickness direction. Equation (1) represents the well-known mass conservation of the solution and species k in the unit mass of dry tissue. The fluxes of the solution and species k can be expressed as

$$J_f = v \quad \text{and} \quad J_k = vC_k - z_k D_k \left(\frac{F}{RT} \frac{\partial \phi}{\partial x} \right) C_k - D_k \frac{\partial C_k}{\partial x} \tag{2}$$

where v is the velocity of convective flow, z_k is the charge number of species k, D_k is the diffusion coefficient of species k, F is the Faraday's constant, R is the universal gas constant, T is the absolute temperature, ϕ is the electrostatic potential, x is the transient thickness coordinate, which is expressed in terms of the dry tissue thickness coordinate ξ

$$dx = \left(1 + \frac{W_f \rho_d}{\rho_f} \right) d\xi \tag{3}$$

The electrostatic potential gradient can be determined in terms of the conservation of current density $F \sum z_j J_j = 0$ and the condition of electro-neutrality $\sum z_j C_j = 0$.

The velocity of the convective flow in the corneal stroma can be determined using the Darcy's law as follows

$$v = -K_\mu \frac{dP}{dx} = -\frac{dP}{d\xi} \frac{K_\mu}{1 + W_f \rho_d / \rho_f} \tag{4}$$

where K_μ is the flow conductivity coefficient and $P = IOP - \gamma \exp(-W_w)$ is the local hydrostatic pressure, IOP is the intraocular pressure, γ is an empirical constant and W_w is the stromal hydration which is defined as the mass of water in unit mass of dry tissue. Note that for dilute solutions $C_w \upsilon_w \approx 1$. Thus the following equations can be obtained

$$\frac{W_f}{W_w} = 1 + \frac{\sum C_j m_j}{C_w m_w} \approx \frac{\rho_f}{\rho_w} \tag{5}$$

where C_w is the concentration of water (solvent), υ_w is the partial molar volume of water, m_w is the molar mass of water, m_j is the molar mass of species j, ρ_w is the density of water.

Note that the flux expressions given by Equation (2) applies only to the corneal stroma. To the epithelium and endothelium the flux expressions have to be obtained from the theory of irreversible thermodynamics [9], which can be expressed as

$$\begin{aligned} J_f &= L_p \Delta P - L_p \sum \sigma_j (RT\Delta C_j + z_j C_j F \Delta\phi) \\ J_k &= (1-\sigma_k) J_f C_k + \omega_k (RT\Delta C_k + z_k C_k F\Delta\phi) + J_{ak} \end{aligned} \tag{6}$$

where L_p is the hydraulic conductivity, ΔP, $\Delta\phi$ and ΔC_j are the differences of hydrostatic pressure, electrostaic potential and concentration of species j in the two sides of the membrane, σ_k is the reflection coefficient, ω_k is the permeability coefficient of species k, J_{ak} is the active pump rate of species k. Again, the fluxes of species across the membrane should satisfy the conservation of current density and the condition of electro-neutrality, from which the electrostatic potential gradient across the membrane can be determined. Equation (6) provides the flux boundary conditions for the mass transport Equation (1). The corneal thickness changes can be calculated using Equation (3) [10] after the stromal hydration is calculated from Equation (5).

3. NUMERICAL RESULTS AND CONCLUSIONS

For given initial conditions (such as ionic concentrations and the hydration in stroma) and boundary conditions (ionic concentrations and hydrostatic pressures in tears and aqueous humour) one can solve the partial differential Equation (1) using various existing codes (such as FEA and Matlab) to obtain the hydration distribution, W_w and then use Equation (3) to obtain the thickness change. Figure 2 shows the stromal thickness response to 30 mOsm NaCl

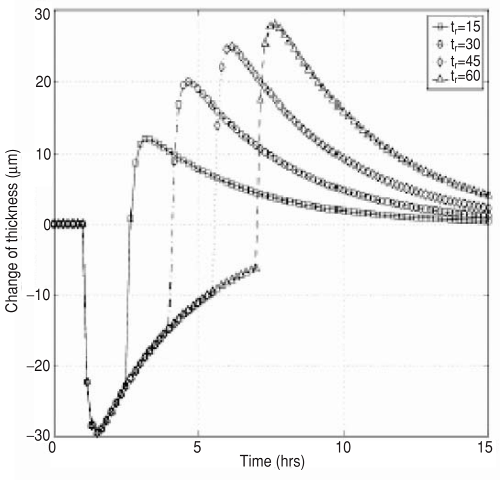


Figure 2. Stromal thickness response to 30 mOsm NaCl shook and reversal at various different times (hours).

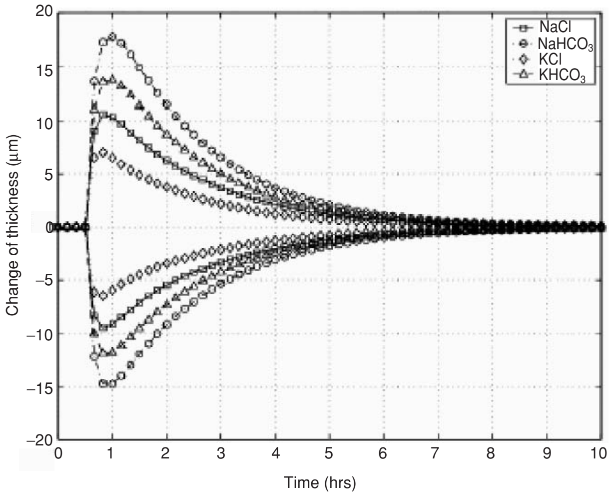


Figure 3. Stromal thickness response to the profusion with 15 mOsm solute hypertonic and hypotonic solutions applied to endothelial surface.

shook and reversals at various different times (applied to the endothelial surface while the epithelium is blocked). The results provide the information on how the initial hydration affects the response of the stromal thickness when there is a sudden change in osmotic pressure created by the change in ionic concentration. It shows that, for the same hypertonic reversal, the higher the initial hydration, the larger the change found in the stromal thickness. Figure 3 shows typical responses of the stromal thickness to the profusion with 15 mOsm hypotonic and hypertonic solutions of four different single solutes. The results show that the responses of the stromal thickness are different to different solutions although all solutions have the same ionic concentration. The greatest change in thickness is found to the solution of $NaHCO_3$, while the least change is to the solution of KCl. It is interesting to notice that the thickness change is greater in hypotonic shock than in hypertonic shock.

REFERENCES

1. J.W. Ruberti and S.D. Klyce (2002), Physiology system models of the cornea. In: *Models of the Visual System*, G.K. Hung and K.J. Ciuffreda (eds.), Kluwer Academic/Plenum Publishers, New York.
2. J. Stanley, S. Mishima and S.D. Klyce (1966), In vivo determination of endothelial permeability to water. *Investigative Ophthalmology*, 7, pp. 371–377.
3. M.H. Friedman (1973), Unsteady aspects of corneal thickness control. *Experimental Eye Research*, 15, pp. 645–658.

4. S.D. Klyce and S.R. Russell (1979) Numerical solution of coupled transport equations applied to corneal hydration dynamics. *Journal of Physiology (London)*, 292, pp. 107–134.
5. L.Y. Li, B.J. Tighe and J.W. Ruberti (2004), Mathematical modeling of corneal swelling. *Biomechanics and Modeling in Mechanobiology*, 3, 2, pp. 114–123.
6. W.M. Lai, J.S. Hou and V.C. Mow (1991), A triphasic theory for the swelling and deformation behaviours of articular cartilage. *Journal of Biomechanics Engineering,* 113, pp. 245–258.
7. M.R. Bryant and P.J. McDonnell (1998), A triphasic analysis of corneal swelling and hydration control. *Journal of Biomechanics Engineering,* 120, 3, pp. 370–381.
8. I. Fatt and T. Goldstick (1965), Dynamics of water transport in swelling membranes. *Journal of Colloid Science,* 20, 9, pp. 962–989.
9. O. Kedem and A. Katchalsky (1958) Thermodynamic analysis of the permeability of biological membranes to non-electrolytes. *Biochimica et Biophysica Acta*, 27, 229–246.
10. B.O. Hedbys and S. Mishima (1966) The thickness-hydration relationship of the cornea. *Experimental Eye Research,* 5, 221–228.

A QUASI-BUBBLE FINITE ELEMENT FORMULATION FOR THE SHALLOW WATER EQUATIONS WITH A DISCONTINUOUS BOUNDARY IMPLEMENTATION

Shintaro Bunya[1], Shinobu Yoshimura[1] and Joannes J. Westerink[2]
[1] *Institute of Environmental Studies, The University of Tokyo, 731 Hongo, Bunkyo, Tokyo 1130033, Japan*
[2] *Department of Civil Engineering and Geological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA*

Abstract A finite element model for the shallow water equations based on the quasi-bubble-function approximation is discussed focusing on an alternative implementation of surface elevation boundary conditions. Numerical experiments demonstrate that the conventional implementation of the boundary conditions originate spurious oscillations. In this work, the alternative implementation, which we refer to as a discontinuous boundary implementation (DBI), is proposed with validation via numerical testing. The results indicate that the DBI improves accuracy of numerical solutions and eliminates the spurious modes.

Keywords: shallow water equations, quasi-bubble function, spurious oscillation, surface elevation boundary condition, discontinuous boundary implementation.

1. INTRODUCTION

The shallow water equations (SWE) are used to describe a number of important environmental problems including storm surges, tidal fluctuations and tsunami waves. Many finite element models have been investigated with unstructured grids, which are preferable for intricate natural geography. The quasi-bubble function approximation is one of such finite element models that eliminate instabilities originated from the well-known $2\Delta x$-scale spurious oscillations [1–3]. In addition bubble-function schemes are known to be efficient because the static condensation procedure allows us to extensively reduce the number of unknowns [3].

G. R. Liu et al. (eds.), Computational Methods, 1937–1941.

However, our numerical experiments, some of which are presented later, showed that the quasi-bubble approach with the conventional implementation of surface elevation boundary conditions degrades the accuracy of velocity solutions. In the next section we propose an alternative implementation which remedies deteriorated velocity solutions. The alternative implementation is tested by comparing numerical solutions with the exact analytical solutions.

2. GOVERNING EQUATIONS AND DISCRETIZATION

2.1 A Quasi-Bubble Finite Element Model

The weighted residual form of the depth-averaged two-dimensional shallow water equations in a non-conservative form are written as:

$$\int_\Omega \left(\frac{\partial \zeta}{\partial t} + \nabla \cdot (H\mathbf{u}) \right) \zeta^* \, d\Omega = 0 \tag{1}$$

$$\int_\Omega \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}.\nabla \mathbf{u} + g\nabla\zeta + \tau_{\mathrm{b}}\mathbf{u} \right) u^* \, d\Omega = 0 \tag{2}$$

where ζ is the surface elevation, $\mathbf{u}$ is the vertically averaged lateral velocity, $H(= h + \zeta)$ is the total depth of the water column denoting the bathymetry by h, g is the gravitational acceleration, and τ_{b} is the bottom friction coefficient. ζ^* and u^* are test functions and Ω is the computational domain. The Coriolis effect and diffusive terms are omitted for the sake of simplification.

In order to obtain finite element equations, basis functions are to be defined. In this work, the standard linear triangular element is used for ζ and ζ^*, while the triangular quasi-bubble element is used for u and u^* [3].

2.2 A Discontinuous Boundary Implementation

We propose an alternative implementation derived by extending the idea of Westerink et al. [4]. Applying Green's Theorem to the third term of Equation (2), we obtain

$$\int_\Omega \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}.\nabla \mathbf{u} - g\nabla\zeta + \tau_{\mathrm{b}}\mathbf{u} \right) u^* \, d\Omega + \int_{\Gamma_s} g\hat{\zeta} u^* \, d\Gamma + \int_{\Gamma_n} g\zeta u^* \, d\Gamma = 0 \tag{3}$$

where surface elevation boundary conditions are imposed by substituting given values to $\hat{\zeta}$. Γ_s denotes segments of the domain boundary where the surface elevation boundary conditions are specified, and Γ_n is the complement of Γ_s.

Because this formulation weakly imposes the surface elevation boundary conditions, ζ on Γ_s may differ from $\hat{\zeta}$. Due to the existence of such discontinuity on Γ_s, we call the alternative implementation as a discontinuous boundary implementation (DBI).

In addition, in order to control the jump $\overline{\zeta}(= \zeta - \hat{\zeta})$ on Γ_s, an additional term proportional to $|\zeta|$ is added to Equation (1). Hence we have

$$\int_\Omega \left(\frac{\partial \zeta}{\partial t} + \nabla \cdot (H\mathbf{u}) + \tau_d \overline{\zeta} \right) \zeta^* \, d\Omega = 0 \tag{4}$$

Approximating the unknown functions with nodal values and the prescribed basis functions, we finally derive a finite element form of the governing equations. As a time marching scheme, two-step explicit method is employed.

3. NUMERICAL EXPERIMENTS

In order to evaluate the performance of the DBI, the quarter annulus problem [5] is solved, adopting the linearized shallow water equations (LSWE) as the governing equations. The exact analytical solutions of LSWE were developed by Lynch and Gray [5]. The discretization described above is applied to the LSWE as well. The geometry is depicted in Figure 1 (a). The inner radius r_1 is set to 2×10^5 ft (60.96 km), and the outer radius r_2 is 5×10^5 ft (152.4 km). Bathymetry varies quadratically between $h = 10$ ft (3.048 m) at r_1 and $h = 62.5$ ft (19.05 m) at r_2. The linear bottom friction is adopted with the coefficient τ_b of 0.0001 s^{-1}. Δt and τ_d are set to 12.5 s and 0.01, respectively. An M_2 tidal wave with the amplitude of 0.1 ft (0.03048 m) is imposed on the

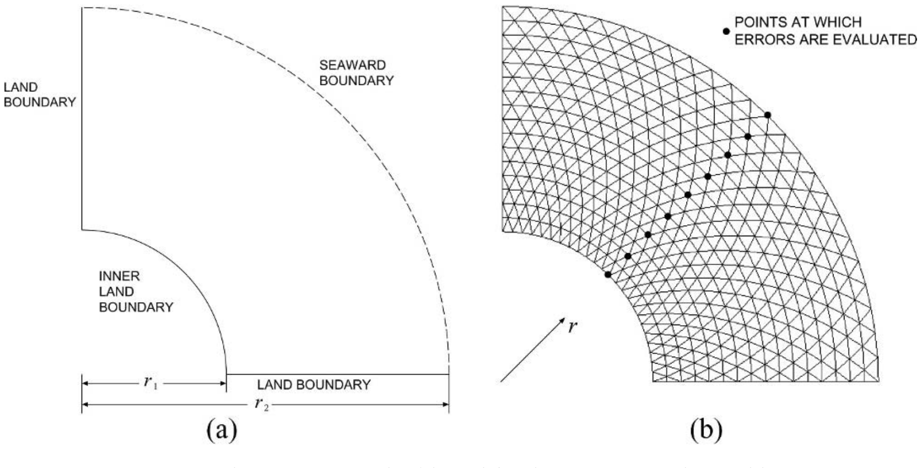
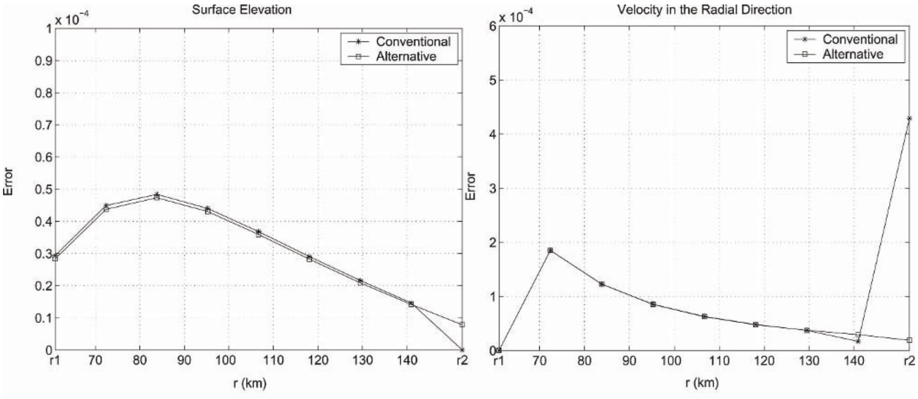


Figure 1. The geometry and grid used for the quarter annulus problem.



(a) (b)

Figure 2. Errors in numerical solutions at each node relative to the exact analytical solutions.

seaward boundary as a surface elevation boundary condition. The grid used in this test is shown in Figure 1 (b). Errors in numerical solutions relative to the exact analytical solutions are evaluated at the nodes indicated in Figure 1 (b). Maxima of the absolute values of the error at each node over the 10th period are shown in Figure 2.

It is observed in Figure 2 (b) that the conventional implementation yields significant error in the velocity solution at the seaward boundary node. Our supplemental numerical experiments indicated that, when the advection term is added to the momentum equation, the deteriorated velocity solutions at the seaward boundary lead to critical non-physical oscillations. On the other hand, as shown in Figure 2 (b), the DBI shows smooth error profiles, which significantly improves stability of the quasi-bubble scheme.

4. CONCLUSION

In this paper, quasi-bubble finite element models of the SWE have been discussed. It was demonstrated by some numerical experiments that the conventional implementation of surface elevation boundary degrades accuracy of velocity solutions near the surface elevation boundary. As an effective remedy of the deteriorated solutions, the DBI was proposed with numerical verifications. Considering that the quasi-bubble-function scheme eliminates $2\Delta x$-scale oscillation without requiring any inconsistent damping, and that the static condensation procedure enhances its computational efficiency, the quasi-bubble scheme adopting the DBI is expected to be an advantageous choice for solving large-scale shallow water problems.

REFERENCES

1. P. Mewis and K.-P. Holz (1993), A quasi-bubble-function approach for shallow water waves. *Advances in Hydro-Science and Engineering*, 1, pp. 768–774.
2. J.M. Hervouet (2000), On spurious oscillations in primitive shallow water equations. In: *the XIII International Conference on Computational Methods in Water Resources*, pp. 929–936.
3. J.H. Atkinson, J.J. Westerink and J.M. Hervouet (2004), Similarities between the quasi-bubble and the generalized wave continuity equation to the shallow water equations. *International Journal for Numerical Methods in Fluids*, 45, pp. 689–714.
4. J.J., Westerink, J.K. Wu and R.L. Kolar (1994), The influence of normal flow boundary conditions on spurious modes in finite element solutions to the shallow water equations. *International Journal for Numerical Methods in Fluids*, 18, pp. 1021–1060.
5. D.R. Lynch and G. Gray (1978), Analytic solutions for computer flow model testing. *Journal of the Hydraulics Division*, 104, pp. 1409–1428.

FREE VIBRATION ANALYSIS OF MULTIPLY CONNECTED PLATES USING THE METHOD OF FUNDAMENTAL SOLUTIONS

Ying-Te Lee[1], Jeng-Tzong Chen[1] and I-Lin Chen[2]
[1]*Department of Harbor and River Engineering, National Taiwan Ocean University, Keelung 20224, Taiwan*
[2]*Department of Naval Architecture, National Kaohsiung Marine University, Kaohsiung 81443, Taiwan*

Abstract In this paper, the method of fundamental solutions (MFS) for solving the eigenfrequencies of multiply connected plates is proposed. The coefficients of influence matrices are easily determined when the fundamental solution is known. True and spurious eigensolutions appear at the same time. It is found that the spurious eigensolution using the MFS depends on the location of the inner boundary where the fictitious sources are distributed. To verify this finding, mathematical analysis for the appearance of spurious eigenequations using degenerate kernels and circulants is done by demonstrating an annular plate with a discrete model. In order to obtain the true eigensolution, the Burton & Miller method is utilized to filter out the spurious eigensolutions. One example is demonstrated analytically and numerically to see the validity of the present method.

Keywords: method of fundamental solutions, biharmonic equation; circulant; degenerate kernel; SVD updating document; Burton & Miller method

1. INTRODUCTION

The method of fundamental solutions (MFS) is a numerical approach as well as finite difference method (FDM), finite element method (FEM) and boundary element method (BEM). This method was attributed to Kupradze in 1964 [1]. The MFS was applied to many problems (Refs.[2–4]), and can be regarded as one kind of meshless method. It has several advantages over boundary element

G. R. Liu et al. (eds.), Computational Methods, 1943–1949.

method, e.g., no boundary integrals, no singularity and mesh-free model. Although MFS has been applied to solve many engineering problems, most of them are for cases of simply connected domains. Chen *et al*. have tried to solve the eigenproblem of multiply connected membrane and found that spurious eigenvalues also appear (Ref. [5]) as well as BEM (Ref. [6]). We may wonder whether spurious solutions also occur for the plate case rather than membrane.

In this paper, the MFS for solving the eigenfrequencies of annular plate is proposed. The occurring mechanism of the spurious eigensolution of an annular plate is studied analytically. The degenerate kernels and circulants are employed to determine the spurious eigensolution. In order to filter out the spurious eigenvalues, singular value decomposition updating technique and Burton & Miller method are utilized. An annular case is demonstrated analytically to see the validity of the present method.

2. ANALYTICAL DERIVATION OF FREE VIBRATION FOR ANNULAR PLATE USING THE METHOD OF FUNDAMENTAL SOLUTIONS

The governing equation for an annular plate vibration in Figure 1 is the biharmonic equation as follows:

$$\nabla^4 u(x) = \lambda^4 u(x), \quad x \in \Omega$$

where ∇^4 is the biharmonic operator, u is the lateral displacement, $\lambda^4 = \omega^2 \rho_0 h/D$, λ is the frequency parameter, ω is the angular frequency, ρ_0 is the surface density, D is the flexural rigidity expressed as $D = Eh^3/12(1-\nu^2)$ in terms of Young's modulus E, the Poisson ratio ν and the plate thickness h, Ω

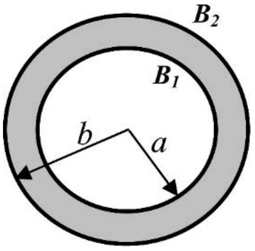


Figure 1. An annular problem.

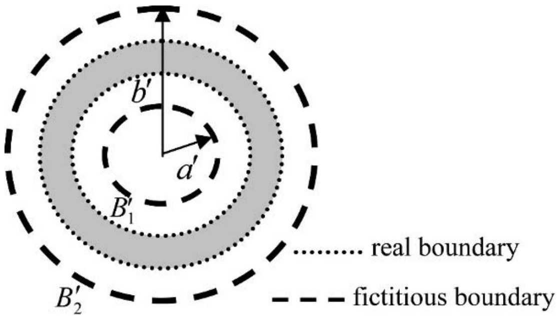


Figure 2. Figure sketch for source distribution.

is the domain of the thin plate. The fundamental solution is chosen as

$$U(s,x) = \frac{1}{8\lambda^2}\left[Y_0(\lambda r) - iJ_0(\lambda r) + \frac{2}{\pi}(K_0(\lambda r) - iI_0(\lambda r))\right] \tag{2}$$

where $r \equiv |s-x|$, $i^2 = -1$, $J_0(\lambda r)$ and $Y_0(\lambda r)$ are the first kind and second kind zeroth-order Bessel functions, respectively, $I_0(\lambda r)$ and $K_0(\lambda r)$ are the first and second kind zeroth-order modified Bessel functions, respectively. Based on the MFS, we can represent the displacement field of plate vibration by

$$u(x_i) = \sum_{j=1}^{2N} P(s_j, x_i)\phi_j + \sum_{j=1}^{2N} Q(s_j, x_i)\varphi_j, \tag{3}$$

where $2N$ is the number of fictitious source nodes. ϕ_j and φ_j are the known densities with respect to P and Q. The two kernels (P and Q) are obtained from either the two of the kernel $U(s,x)$ and the other three kernels, $\Theta(s,x)$, $M(s,x)$ and $V(s,x)$ (Ref. [7]). The slope (θ), normal moment (m) and effective shear force (v), are also obtained as reference (Ref. [7]). In order to derive the exact eigensolution, degenerate kernel and circulant are considered for an annular plate. The field and source points are distributed as shown in Figure 2. Here, we consider the clamped case ($u = 0$ and $\theta = 0$) by using U and Θ kernels. We distributed $2N$ field points on the real boundary, and the same $2N$ sources are distributed on the fictitious boundary. By matching the boundary condition, we obtain

$$[SM^{cc}]\begin{Bmatrix}\phi\\ \phi2\\ \varphi1\\ \varphi2\end{Bmatrix} = \begin{bmatrix} U11 & U12 & \Theta11 & \Theta12\\ U21 & U22 & \Theta21 & \Theta22\\ U11_\theta & U12_\theta & \Theta11_\theta & \Theta12_\theta\\ U21_\theta & U22_\theta & \Theta21_\theta & \Theta22_\theta \end{bmatrix}_{8N\times8N}\begin{Bmatrix}\phi1\\ \phi2\\ \varphi1\\ \varphi2\end{Bmatrix} = \{0\} \tag{4}$$

where $\{\phi 1\}$, $\{\phi 2\}$, $\{\varphi 1\}$ and $\{\varphi 2\}$ are the generalized coefficients for B_1 and B_2 with a dimension of $2N \times 1$, the matrices $[Uij]$, $[\Theta ij]$, $[Uij_\theta]$ and $[\Theta ij_\theta]$ mean the influence matrices of U, Θ, U_θ and Θ_θ kernels which are obtained by collocating the field and source points on B_i and B'_j with a dimension of $2N \times 2N$, respectively. For the existence of nontrivial solution, the determinant of the matrix vs. the eigenvlaue must be zero, i.e.,

$$\det[SM^{cc}] = \prod_{m=-(N-1)}^{N} \det([T_m^{cc}][S_m^{U\Theta}]) = 0, \tag{5}$$

where

$$[T_m^{cc}] = \begin{bmatrix} J_m(\lambda a) & Y_m(\lambda a) & I_m(\lambda a) & K_m(\lambda a) \\ J_m(\lambda b) & Y_m(\lambda b) & I_m(\lambda b) & K_m(\lambda b) \\ J'_m(\lambda a) & Y'_m(\lambda a) & I'_m(\lambda a) & K'_m(\lambda a) \\ J'_m(\lambda b) & Y'_m(\lambda b) & I'_m(\lambda b) & K'_m(\lambda b) \end{bmatrix} \tag{6}$$

and

$$[S_m^{U\Theta}] = \begin{bmatrix} -iJ_m(\lambda a') & Y_m(\lambda b') - iJ_m(\lambda b') & -iJ'_m(\lambda a') & Y'_m(\lambda b') - iJ'_m(\lambda b') \\ J_m(\lambda a') & 0 & J'_m(\lambda a') & 0 \\ -(-1)^m i\frac{2}{\pi}I_m(\lambda a') & \frac{2}{\pi}[k_m(\lambda b') - (-1)^m i I_m(\lambda b')] & -(-1)^m i\frac{2}{\pi}I'_m(\lambda a') & \frac{2}{\pi}[k'_m(\lambda b') - (-1)^m i I'_m(\lambda b')] \\ \frac{2}{\pi}I_m(\lambda a') & 0 & \frac{2}{\pi}I'_m(\lambda a') & 0 \end{bmatrix}. \tag{7}$$

It is noted that the matrix $[T_m^{cc}]$ denotes the matrix of true eigenequation for the C-C case and the matrix $[S_m^{U\Theta}]$ denotes the matrix of spurious eignequation in the U-Θ formulation after comparing with the analytical solution for the annular plate (Ref. [8]). The matrix in Equation (7) can be further decomposed into

$$\det[S_m^{U\Theta}] = \begin{vmatrix} J_m(\lambda a') & J'_m(\lambda a') \\ I_m(\lambda a') & I'_m(\lambda a') \end{vmatrix} \times \begin{vmatrix} Y_m(\lambda b') - iJ_m(\lambda b') & Y'_m(\lambda b') - iJ'_m(\lambda b') \\ K_m(\lambda b') - i(-1)^m I_m(\lambda b') & K'_m(\lambda b') - i(-1)^m I'_m(\lambda b') \end{vmatrix} = 0 \tag{8}$$

Since the latter part of Equation (8) is never zero, the spurious eigenequation depends on a'. It is noted that the spurious eigensolution happens to be true

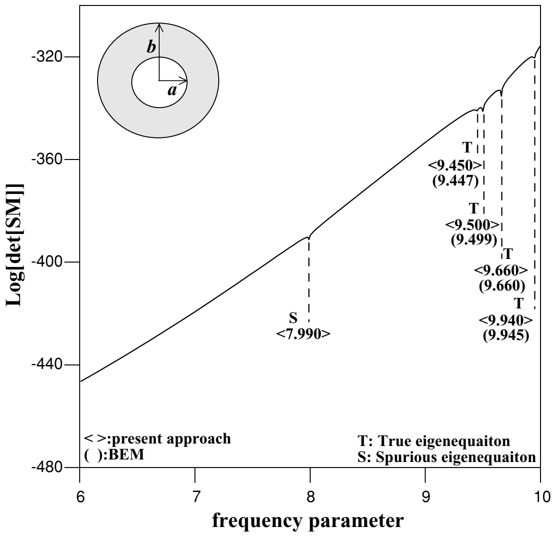


Figure 3. The determinant vs. frequency parameter by using the U-Θ formulation.

eigensolution of the clamped circular plate with a radius a'. Therefore, the positions of spurious eigenvalues for the annular problem depend on the location of inner fictitious boundary a' where the sources are distributed.

3. A NUMERICAL EXAMPLE

An annular plate with the inner radius of 0.5 meter and the outer radius of 1 meter are considered, respectively. The source points are distributed at $a' = 0.4$ meter and $b' = 1.2$ meter. Forty-six nodes are uniformly distributed on the inner and outer fictitious boundaries. Figure 3 shows the determinant vs. frequency parameter by using the U-Θ formulation. The drop location indicates the possible eigenvalues. Figure 4 shows the determinant vs. frequency parameter by using the Burton & Miller method for the annular plate. It is found that the appearance of spurious eigenvalues is suppressed. After comparing the result with the analytical solution, good agreement is made.

4. CONCLUSIONS

The mathematical analysis has shown that spurious eigenvalues occur by using degenerate kernels and circulants when the method of fundamental solutions is used to solve the eigenvalue of annular plates. The positions of

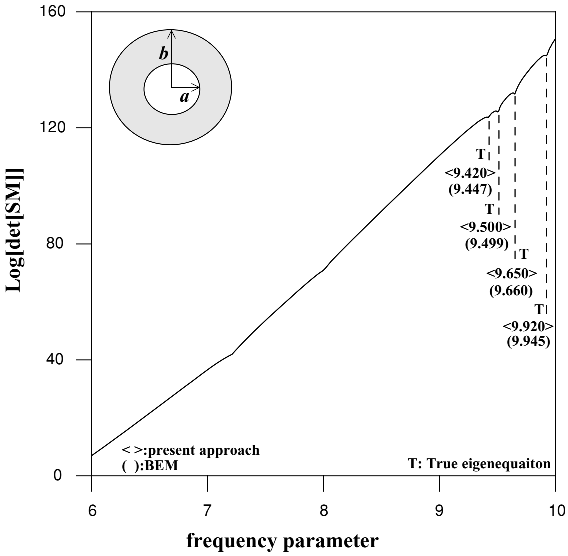


Figure 4. The determinant vs. frequency parameter by using the U-Θ formulation in conjunction with Burton & Miller method.

spurious eigenvalues for the annular problem depend on the location of inner fictitious boundary where the sources are distributed. The spurious eigenvalues in the annular problem are found to be the true eigenvalues of the associated simply connected problem bounded by the inner sources. We have employed the Burton & Miller method to filter out the spurious eigenvalues successfully.

REFERENCES

1. V.D. Kupradze (1964), A method for the approximate solution of limiting problems in mathematical physics. *Computational Mathematics and Mathematical Physics*, 4, pp. 199–205.
2. G. Fairweather and A. Karageorghis (1998), The method of fundamental solutions for elliptic boundary value problems. *Advances in Computational Mathematics*, 9, pp. 69–95.
3. C.S. Chen, M.A. Golberg and Y.C. Hon (1998), The method of fundamental solutions and quasi-Monte-Carlo method for diffusion equations. *International Journal for Numerical Methods in Engineering*, 43, pp. 1421–35.
4. A. Karageorghis (2001), The method of fundamental solutions for the calculation of the eigenvalues of the Helmholtz equation. *Applied Mathematics Letters*, 14, pp. 837–42.

5. J.T. Chen, S.Y. Lin, K.H. Chen and I.L. Chen (2004), Mathematical analysis and numerical study of true and spurious eigenequations for free vibration of plates using real-part BEM. *Computational Mechanics*, 34, pp. 165–180.
6. J.T. Chen, L.W. Liu and H.-K. Hong (2003), Spurious and true eigensolutions of Helmholtz BIEs and BEMs for a multiply-connected problem. In *Proceedings of the Royal Society London Series A*, 459, pp.1891–925.
7. J.T. Chen, I.L. Chen and Y.T. Lee (2004), Eigensolutions of multiply-connected membranes using the method of fundamental solutions. *Engineering Analysis with Boundary Elements*, Revised.
8. W. Leissa (1969), *Vibration of Plates*. NASA SP-160.

A TWO-GRID FINITE ELEMENT DISCRETIZATION SCHEME FOR NONLINEAR EIGENVALUE PROBLEMS

C.-S. Chien and B.-W. Jeng
Department of Applied Mathematics, National Chung-Hsing University,
Taichung, Taiwan 402 China

Abstract We describe a two-grid finite element discretization scheme which can be used to trace solution branches as well as to detect bifurcation points of certain second order semilinear elliptic eigenvalue problems. Sample numerical results are reported.

Keywords: two-grid scheme, continuation method, finite element method, nonlinear eigenvalue problem.

1. INTRODUCTION

Recently Chien and Jeng [1] proposed some two-grid discretization schemes for both linear and nonlinear eigenvalue problems, where both finite elements and centered differences were used to discretize the PDEs. The modifications and numerical implementations of these algorithms can be found, e.g., in [2–4].

In this paper we modify the two-grid finite element discretization scheme proposed by [5] for second order linear elliptic eigenvalue problem, and combine it with the two-grid finite element discretization scheme proposed by the authors [1] for nonlinear eigenvalue problems. The resultant algorithm can be used to trace solution branches as well as to detect simple and multiple bifurcation point on the solution branches. Moreover, the two-grid finite element discretization scheme for linear eigenvalue problems can be exploited to compute eigenpairs of the Maxwell equations. The details will be given elsewhere.

G. R. Liu et al. (eds.), Computational Methods, 1951–1955.

This paper is organized as follows. In Section 2 we describe the main algorithm. The numerical results are given in Section 3.

2. A TWO-GRID FINITE ELEMENT DISCRETIZATION SCHEME

In this section we describe a two-grid finite element discretization scheme for nonlinear eigenvalue problems. First we consider the following linear eigenvalue problem

$$\begin{aligned} F(u,\lambda) = -\Delta u - \lambda u = 0 &\quad \text{in } \Omega = [0,1]^2, \\ u = 0 &\quad \text{on } \partial\Omega. \end{aligned} \tag{1}$$

Let $V_{\tilde{h}}(\Omega)$ and $V_h(\Omega)$ be any two finite dimensional subspaces with $0 < h < \tilde{h} < 1$ which satisfy $V_{\tilde{h}}(\Omega) \subset V_h(\Omega) \subset H_0^1(\Omega)$. The finite element approximation of Equation (1) on the coarse space $V_{\tilde{h}}(\Omega)$ is to find $(u_{\tilde{h}}, \lambda_{\tilde{h}}) \in V_{\tilde{h}}(\Omega) \times \mathbb{R}$ such that

$$\int_\Omega \nabla u_{\tilde{h}} \cdot \nabla v \, dxdy = \lambda_{\tilde{h}} \int_\Omega u_{\tilde{h}} v \, dxdy \quad \forall v \in V_{\tilde{h}}(\Omega). \tag{2}$$

The algebraic version of Equation (2) is to find $(x_{\tilde{h}}, \lambda_{\tilde{h}}) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}$ such that

$$A_{\tilde{h}} x_{\tilde{h}} = \lambda_{\tilde{h}} B_{\tilde{h}} x_{\tilde{h}},$$

where $\widetilde{N} = \dim V_{\tilde{h}}(\Omega)$, $x_{\tilde{h}} = [x_{\tilde{h},1}, x_{\tilde{h},2}, \ldots, x_{\tilde{h},\widetilde{N}}]^T$, and $A_{\tilde{h}}, B_{\tilde{h}} \in \mathrm{R}^{\widetilde{N} \times \widetilde{N}}$ corresponding to the integrals $\int_\Omega \nabla u_{\tilde{h}} \cdot \nabla v$ and $\int_\Omega u_{\tilde{h}} v$, respectively.

Now we consider the following nonlinear eigenvalue problem

$$\begin{aligned} F(u,\lambda) = -\Delta u + \lambda f(u) = 0 &\quad \text{in } \Omega = [0,1]^2, \\ u = 0 &\quad \text{on } \partial\Omega, \end{aligned} \tag{3}$$

where $f,\ g : \mathbb{R} \to \mathbb{R}$ are smooth maps. The two-grid finite element discretization scheme for Equation (3) is based on the predictor-corrector continuation method [6] for tracing solution branches. The computations of extremum eigenvalues determine the locations of the first few bifurcation points.

Algorithm 2.1. A two-grid finite element continuation algorithm for Equation (3).

1. Use the RQI to find the first few eigenvalues $\lambda_{\tilde{h}}^1, \cdots, \lambda_{\tilde{h}}^k$ and the corresponding eigenvectors $x_{\tilde{h}}^1, \cdots, x_{\tilde{h}}^k$ *with* $||x_{\tilde{h}}^1||_2 = \cdots = ||x_{\tilde{h}}^k||_2 = 1$ on the coarse grid such that

$$A_{\tilde{h}} x_{\tilde{h}}^i = \lambda_{\tilde{h}}^i B_{\tilde{h}} x_{\tilde{h}}^i, \quad i = 1, 2, \ldots, k.$$

2. Solve k linear systems on the fine grid: Find $x_h^1, \dots, x_h^k$, such that

$$A_h x_h^i = \lambda_{\tilde{h}}^i B_h (I_{\tilde{h}}^h x_{\tilde{h}}^i), \quad i = 1, 2, \dots, k.$$

3. Compute the Rayleigh quotient

$$\lambda_h^i = \frac{(x_h^i)^T A_h (x_h^i)}{(x_h^i)^T B_h (x_h^i)}, \quad i = 1, 2, \dots, k.$$

4. Perform the RQI:
   (i) Normalize the approximate eigenvectors obtained in Step 2,

$$x_0^i = x_h^i \big/ ||x_h^i||_2 \quad \text{and} \quad \text{set } \mu_0^i = \lambda_h^i, \quad i = 1, 2, \dots, k.$$

```
(ii) For $i = 1 : k$
         For $j = 0, 1, 2, \dots$
             Solve $(A_h - \mu_j^i B_h) z_{j+1}^i = B_h x_j^i$ for $z_{j+1}^i$
             $x_{j+1}^i = z_{j+1}^i \big/ ||z_{j+1}^i||_2$
             $\mu_{j+1}^i = \frac{(x_{j+1}^i)^T A_h (x_{j+1}^i)}{(x_{j+1}^i)^T B_h (x_{j+1}^i)}$
         End
     End
```

5. Use the predictor-corrector continuation algorithm to find the approximate solution $(u_{\tilde{h}}, \lambda_{\tilde{h}})$on the coarse space $V_{\tilde{h}}(\Omega)$: Find $(u_{\tilde{h}}, \lambda_{\tilde{h}}) \in V_{\tilde{h}}(\Omega) \times \mathbb{R}$ such that

$$\int_\Omega \nabla u_{\tilde{h}} \cdot \nabla v + \lambda_{\tilde{h}} \int_\Omega f(u_{\tilde{h}}) v = 0 \quad \forall v \in V_{\tilde{h}}(\Omega).$$

6. For each continuation step, correct the approximate solution $(u_{\tilde{h}}, \lambda_{\tilde{h}})$by the following steps:
   (i) Make a correction on the fine space $V_h(\Omega)$*: Find* $e_h \in V_h(\Omega)$ such that

$$\int_\Omega \nabla e_h \cdot \nabla v + \lambda_{\tilde{h}} \int_\Omega f'(u_{\tilde{h}}) e_h v = -\int_\Omega \nabla u_{\tilde{h}} \cdot \nabla v - \lambda_{\tilde{h}} \int_\Omega f(u_{\tilde{h}}) v$$
$$\forall v \in V_h(\Omega).$$

   (ii) Make a further correction on the coarse space $V_{\tilde{h}}(\Omega)$: Find $e_{\tilde{h}} \in V_{\tilde{h}}(\Omega)$ such that

$$\int_\Omega \nabla e_{\tilde{h}} \cdot \nabla v + \lambda_{\tilde{h}} \int_\Omega f'(u_{\tilde{h}}) e_{\tilde{h}} v = -\frac{1}{2} \lambda_{\tilde{h}} \int_\Omega f''(u_{\tilde{h}}) e_h^2 v \quad \forall v \in V_{\tilde{h}}(\Omega).$$

   (iii) Set $u_h = u_{\tilde{h}} + e_h + e_{\tilde{h}}$.

Table 1. The first six eigenvalues and residuals of the corresponding eigenvectors of Equation (1).

	Coarse grid eigenvalue	Fine grid eigenvalue		RQI	
	$\lambda_{\widetilde{h}}$	λ_h	Residual	μ_1	Residual
1	0.19805118628[+2]	0.19739262864[+2]	0.10[−4]	0.19739209915[+2]	0.24[−9]
2	0.50383506082[+2]	0.49351779995[+2]	0.95[−4]	0.49348041011[+2]	0.17[−7]
3	0.50383506082[+2]	0.49351779995[+2]	0.95[−4]	0.49348041011[+2]	0.17[−7]
4	0.82142640415[+2]	0.79076580633[+2]	0.27[−3]	0.78956910426[+2]	0.85[−6]
5	0.103696686497[+3]	0.98547414332[+2]	0.42[−3]	0.98696132374[+2]	0.13[−5]
6	0.103696686497[+3]	0.98547414332[+2]	0.42[−3]	0.98696132374[+2]	0.13[−5]

The first six exact eigenvalues:

19.739208802 49.348022005 49.348022005 78.956835208 98.696044010 98.696044010

(iv) Compute $\lambda_h = -\dfrac{\int_\Omega \nabla u_h \cdot \nabla u_h}{\int_\Omega f(u_h)u_h}$.

(v) If necessary, use (u_h, λ_h) as the predicted point, perform Newton's method on the fine space $V_h(\Omega)$.

3. NUMERICAL RESULTS

We report some numerical results concerning the implementations of Algorithm 2.1. All of our computations were executed on a Pentium 4 computer using FORTRAN 95 with double precision arithmetic. The notation $[\pm n]$ stands for multiplication by $10^{\pm n}$.

Example. We consider Equation (1). The domain Ω was divided into two families $T_{\widetilde{h}}$ and T_h of quadratic Lagrange triangles, and $V_{\widetilde{h}}(\Omega),\ V_h(\Omega) \subset H_0^1(\Omega)$ are subspaces of piecewise quadratic functions defined on $T_{\widetilde{h}}$ and T_h, respectively. We took $h = 1\big/16\widetilde{h}$ with $\widetilde{h} = 1/4$. The steps 1-4 of Algorithm 2.1 were implemented to compute the first six eigenvalues of Equation (1), where the MINRES was used to solve the associated linear systems. Table 1 shows that the first six eigenvalues consist of two simple eigenvalues and two clusters of eigenvalues.

REFERENCES

1. C.-S. Chien and B.-W. Jeng (2006), A Two-grid discretization scheme for semilinear elliptic eigenvalue problems. *SIAM Journal on Scientific Computing*, 27, pp. 1287–1304.

2. S.-L. Chang, C.-S. Chien, and B.-W. Jeng (2005), An efficient algorithm for the Schrödinger-Poisson eigenvalue problem. Submitted to *Journal of Computational and Applied Mathematics*.
3. C.-S. Chien and B.-W. Jeng (2005), Symmetry reductions and a posterior finite element error estimators for bifurcation problems, *International Journal of Bifurcation and Chaos*, 15, pp. 2091–2107.
4. S.-L. Chang, C.-S. Chien and B.-W. Jeng (2004), Implementing two-grid centered difference discretization schemes with Lanczos type algorithms, *Technical Report, National Chung-Hsing University*.
5. J. Xu and A. Zhou (1999), A two-grid discretization scheme for eigenvalue problems, *Mathematics of Computation*, 70. pp. 17–25.
6. E.L. Allgower and K. Georg (2003), Introduction to Numerical Continuation Methods, *SIAM Publications*, *USA*.

IMPLEMENTING MINRES AND SYMMLQ FOR EIGENVALUE PROBLEMS

C.-S. Chien and B.-W. Jeng
Department of Applied Mathematics, National Chung-Hsing University,
Taichung, Taiwan 402, China

Abstract We study two grid centered-difference discretization schemes for both linear and nonlinear eigenvalue problems, where the Lanczos type algorithms, namely MINRES and SYMMLQ are used as linear solvers. Numerical experiments on some test problems are reported.

Keywords: two-grid scheme, finite difference method, linear eigenvalue problems, nonlinear eigenvalue problems.

1. INTRODUCTION

In 1975, Paige and Saunders [1] proposed the MINRES and SYMMLQ algorithms for solving large sparse symmetric indefinite linear systems

$$Ax = b, \tag{1}$$

where $A \in \mathbb{R}^{N \times N}$ is nonsingular and symmetric indefinite, and $b \in \mathbb{R}^N$. Since then MINRES and SYMMLQ became popular for solving (1) and large sparse matrix eigenvalue problems. The implementations of the preconditioned versions of these two algorithms can be found in [2]. It is pointed out therein that MINRES is better than SYMMLQ in eigenpair computations. Recently, Chang *et al*. [4] give a comprensive study of the Lanczos type algorithms for both linear and nonlinear eigenvalue problems. It is shown therein that the preconditioned Lanczos algorithm is still very competive compared to MINRES and SYMMLQ. Moreover, Chien and Jeng [4] proposed some two-grid discretization schemes for solving linear and nonlinear eigenvalue problems. Our aim here is to modify and implement the two-grid centered difference discretization

G. R. Liu et al. (eds.), Computational Methods, 1957–1961.

schemes described therein for solving linear and nonlinear eigenvalue problems defined on an L-shaped domain.

This paper is organized as follows. In Section 2 we describe the main algorithm. Our numerical results are reported in Section 3.

2. A TWO-GRID CENTERED DIFFERENCE DISCRETIZATION SCHEME

We consider the following nonlinear boundary value problems

$$\begin{aligned} F(u,\lambda) = -\Delta u - \lambda f(u) - g(u) = 0 \qquad & \text{in } \Omega, \\ u = 0 \qquad & \text{on } \partial\Omega, \end{aligned} \tag{2}$$

where $f, g : \mathbb{R} \to \mathbb{R}$ are smooth maps and Ω is an L-shaped domain which is obtained by cutting away the upper right corner of the unit square. We discretize Equation (2) via centered differences on the coarse grid with uniform meshsize $\tilde{h}$ on the x- and y-axis. The corresponding matrix nonlinear eigenvalue problem is

$$\text{Find } (u_{\tilde{h}}, \lambda_{\tilde{h}}) \in \mathbb{R}^{\tilde{N}} \times \mathbb{R} \text{ such that } A_{\tilde{h}} u_{\tilde{h}} - \lambda_{\tilde{h}} f(u_{\tilde{h}}) - g(u_{\tilde{h}}) = 0,$$

where $\tilde{N}$ is the number of interior points on the coarse space, $A_{\tilde{h}} \in \mathbb{R}^{\tilde{N}\times\tilde{N}}$ is the coefficient matrix corresponding to the Laplacian operator $-\Delta$, and $f(u_{\tilde{h}}) = [f((u_{\tilde{h}})_1), \ldots, f((u_{\tilde{h}})_{\tilde{N}})]^T$, $g(u_{\tilde{h}}) = [g((u_{\tilde{h}})_1), \ldots, g((u_{\tilde{h}})_{\tilde{N}})]^T \in \mathbb{R}^{\tilde{N}}$.

If we consider the linear approximation of the mapping $F(u, \lambda_{\tilde{h}})$ at $u_{\tilde{h}}$, the approximate solution $u_{\tilde{h}}$ can be corrected by solving the following equation

$$\begin{aligned} -\Delta e - \lambda_{\tilde{h}} f'(u_{\tilde{h}})e - g'(u_{\tilde{h}})e = \Delta u_{\tilde{h}} + \lambda_{\tilde{h}} f(u_{\tilde{h}}) + g(u_{\tilde{h}}) \qquad & \text{in } \Omega, \\ e = 0 \qquad & \text{on } \partial\Omega, \end{aligned} \tag{3}$$

on the fine grid with uniform meshsize h ($0 < h < \tilde{h} < 1$) on the x- and y-axis. The centered difference analogue of Equation (3) is of the form

Find $e_h \in \mathbb{R}^N$ such that $(A_h - \lambda_{\tilde{h}} C_h - D_h)e_h = -A_h(I_{\tilde{h}}^h u_{\tilde{h}}) + \lambda_{\tilde{h}} f(I_{\tilde{h}}^h u_{\tilde{h}}) + g(I_{\tilde{h}}^h u_{\tilde{h}})$, where N is the number of interior points on the fine space, $I_{\tilde{h}}^h : \mathbb{R}^{\tilde{N}} \to \mathbb{R}^N$ is the interpolation operator from $\mathbb{R}^{\tilde{N}}$ to $\mathbb{R}^N$, and $C_h = \text{diag}(f'((I_{\tilde{h}}^h u_{\tilde{h}})_1), \ldots, f'((I_{\tilde{h}}^h u_{\tilde{h}})_N))$, $D_h = \text{diag}(g'((I_{\tilde{h}}^h u_{\tilde{h}})_1), \ldots, g'((I_{\tilde{h}}^h u_{\tilde{h}})_N)) \in \mathbb{R}^{N\times N}$.

By considering the quadratic approximation of $F(u, \lambda_{\tilde{h}})$ at $u_{\tilde{h}}$, the further coarse grid correction is obtained by solving the following equation

$$\begin{aligned} -\Delta\tilde{e} - \lambda_{\tilde{h}} f'(u_{\tilde{h}})\tilde{e} - g'(u_{\tilde{h}})\tilde{e} = \tfrac{1}{2}(\lambda_{\tilde{h}} f''(u_{\tilde{h}})e_{\tilde{h}}^2 + g''(u_{\tilde{h}})e_{\tilde{h}}^2) \qquad & \text{in } \Omega, \\ \tilde{e} = 0 \qquad & \text{on } \partial\Omega, \end{aligned} \tag{4}$$

on the coarse grid. The centered difference counterpart of Equation (4) is of the form Find $e_{\tilde{h}} \in \mathbb{R}^{\tilde{N}}$ such that

$$(A_{\tilde{h}} - \lambda_{\tilde{h}} C_{\tilde{h}} - D_{\tilde{h}})e_{\tilde{h}} = \tfrac{1}{2}(\lambda_{\tilde{h}} E_{\tilde{h}}^2 f''(u_{\tilde{h}}) + E_{\tilde{h}}^2 g''(u_{\tilde{h}})),$$

where

$$C_{\tilde{h}} = \operatorname{diag}(f'((u_{\tilde{h}})_1), \dots, f'((u_{\tilde{h}})_{\tilde{N}})),\ D_{\tilde{h}} = \operatorname{diag}(g'((u_{\tilde{h}})_1), \\ \cdots, g'((u_{\tilde{h}})_{\tilde{N}})) \in \mathbb{R}^{\tilde{N}\times\tilde{N}},$$

and

$$E_{\tilde{h}} = \operatorname{diag}((I_h^{\tilde{h}} e_h)_1, \dots, (I_h^{\tilde{h}} e_h)_{\tilde{N}}) \in \mathbb{R}^{\tilde{N}\times\tilde{N}}$$

with $I_h^{\tilde{h}} : \mathbb{R}^N \to \mathbb{R}^{\tilde{N}}$ is the restriction operator, and

$$f''(u_{\tilde{h}}) = [f''((u_{\tilde{h}})_1), \dots, f''((u_{\tilde{h}})_{\tilde{N}})]^T, g''(u_{\tilde{h}}) = [g''((u_{\tilde{h}})_1), \\ \dots, g''((u_{\tilde{h}})_{\tilde{N}})]^T \in \mathbb{R}^{\tilde{N}}.$$

In order to obtain an accurate parameter λ_h on the fine grid, we set $u_h = I_{\tilde{h}}^h u_{\tilde{h}} + e_h + I_{\tilde{h}}^h e_{\tilde{h}}$ and compute the Rayleigh quotient

$$\lambda_h = \frac{u_h^T A_h u_h - u_h^T g(u_h)}{u_h^T f(u_h)},$$

where

$$f(u_h) = [f((u_h)_1), \dots, f((u_h)_N)]^T, \quad g(u_h) = [g((u_h)_1), \\ \dots, g((u_h)_N)]^T \in \mathbb{R}^N.$$

Therefore, we have the following algorithm.

Algorithm 2.1. A two-grid centered difference continuation algorithm for Equation (2).

1. Use the predictor-corrector continuation algorithm to find the approximate solution $(u_{\tilde{h}}, \lambda_{\tilde{h}})$ of Equation (2) on the coarse grid: Solve the parameter dependent nonlinear system

$$A_{\tilde{h}} u_{\tilde{h}} - \lambda_{\tilde{h}} f(u_{\tilde{h}}) - g(u_{\tilde{h}}) = 0.$$

2. For each continuation step, correct the approximate solution $(u_{\tilde{h}}, \lambda_{\tilde{h}})$ by the following steps:
   (i) Make a correction on the fine grid: Solve the linear system

$$(A_h - \lambda_{\tilde{h}} C_h - D_h)e_h = -A_h(I_{\tilde{h}}^h u_{\tilde{h}}) + \lambda_{\tilde{h}} f(I_{\tilde{h}}^h u_{\tilde{h}}) + g(I_{\tilde{h}}^h u_{\tilde{h}}).$$

(ii) Make a further correction on the coarse grid: Solve the linear system

$$(A_{\tilde{h}} - \lambda_{\tilde{h}} C_{\tilde{h}} - D_{\tilde{h}}) e_{\tilde{h}} = \frac{1}{2}(\lambda_{\tilde{h}} E_{\tilde{h}}^2 f''(u_{\tilde{h}}) + E_{\tilde{h}}^2 g''(u_{\tilde{h}})).$$

(iii) Set $u_h = I_{\tilde{h}}^h u_{\tilde{h}} + e_h + I_{\tilde{h}}^h e_{\tilde{h}}$.

(iv) Compute $\lambda_h = \dfrac{u_h^T A_h u_h - u_h^T g(u_h)}{u_h^T f(u_h)}$.

(v) If necessary, use (u_h, λ_h) as the predicted point, perform Newton's method on the fine grid space.

3. NUMERICAL RESULTS

We report some numerical results concerning the implementations of Algorithm 2.1. All of our computations were executed on a Pentium 4 computer using FORTRAN 95 with double precision arithmetic.

Example 1. We consider the following semilinear elliptic eigenvalue problem

$$\begin{aligned} -\Delta u - \lambda \sin u + u^4 &= 0 \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{5}$$

This example was numerically tested using two-grid centered difference discretization scheme. We discretized Equation (5) with coarse meshsize $\tilde{h} = 1/16$ and fine meshsize $h = 1/128$ and traced the solution branch of Equation (5) bifurcating at $(0, \lambda_{1,1})$. The discretization matrices are of order 161×161 and 12033×12033 on the coarse grid and the fine grid, respectively. The sample numerical result is shown in Table 1.

Table 1. Sample result for the solution curve of Equation (5) branching at $(0, \lambda_{1,1})$, $\tilde{h} = 1/16$, h = $1/128$, $\varepsilon = 5 \times 10^{-7}$, $tol = 5 \times 10^{-10}$, using Algorithm 2.1.

NCS	$\lambda_{\tilde{h}}$	$\|\|u_{\tilde{h}}\|\|_\infty$	λ_h	$\|\|u_h\|\|_\infty$
4	38.4795133	0.101983986D + 00	38.4902556	0.172834086D + 00
7	38.6611478	0.164473214D + 00	38.5932909	0.219566775D + 00
9	38.7407056	0.199730890D + 00	38.6577884	0.249217223D + 00
10	38.7792939	0.217406954D + 00	38.7018398	0.269068719D + 00
20	39.2130875	0.361247801D + 00	39.1839267	0.425219130D + 00
30	39.7540783	0.542339282D + 00	39.7561347	0.598334571D + 00
40	41.1645776	0.850916446D + 00	41.1716333	0.893770029D + 00
50	42.8752374	0.111421014D + 01	42.8832040	0.115216692D + 01
60	44.7653129	0.134224760D + 01	44.7737203	0.137838364D + 01
70	46.7730794	0.154296072D + 01	46.7817107	0.157864863D + 01
85	53.2183767	0.203164726D + 01	53.2264487	0.206916166D + 01
100	60.0268568	0.241186860D + 01	60.0317721	0.245309943D + 01

REFERENCES

1. C.C. Paige and M.A. Saunders (1975), Solution of sparse indefinite systems of linear equations. *SIAM Journal on Numerical Analysis*, 12, pp. 617–629.
2. F.A. Dul (1998), MINRES and MINERR are better than SYMMLQ in eigenpair computations. *SIAM Journal on Scientific Computing*, 19, pp. 1767–1782.
3. S.-L. Chang, C.-S. Chien and B.-W. Jeng (2004), Multigrid and two-grid schemes with Lanczos type solvers for the numerical solutions of eigenvalue problems, *preprint.*
4. C.-S. Chien and B.-W. Jeng (2006), A two-grid discretization schemes for semilinear elliptic eigenvalue problems, *SIAM Journal on Scientific Computing*, 27, pp. 1287–1304.

VIBRATION OF A BEAM WITH A BREATHING CRACK SUBJECT TO MOVING MASS

S.S. Law and X.Q. Zhu
Department of Civil and Structural Engineering, Hong Kong Polytechnic University, Kowloon, Hong Kong, People's Republic of China

Abstract The dynamic behaviour of a beam with a breathing crack subject to moving loads is studied. The two segments of the beam, separated by the crack, are related to one another by time varying connection matrices representing the interaction forces. Two different sets of admissible functions which satisfy the respective geometric boundary conditions are assumed for these fictitious sub-beams. Numerical procedure is presented for the dynamic response of the simple beam subject to moving loads.

Keywords: moving loads, breathing crack, dynamic response, damage detection.

1. INTRODUCTION

The analysis of a continuous elastic system subject to moving loads has been a subject of interest in many research areas. Lee and Ng [1] used the assumed mode method to analyze the dynamic response of a beam with a single-sided crack subject to a moving load on the top. The beam is modelled as two segments separated by the crack. Two different sets of admissible functions satisfying the respective geometric boundary conditions are then assumed for these two fictitious sub-beams. The rotational discontinuity at the crack is modelled by a torsional spring with an equivalent spring constant for the crack. The equality of transverse deflection at the crack is enforced by a linear spring of very large stiffness. Parhi and Behera [2] utilized a local stiffness matrix to model the crack section. Bilello and Bergman [3] modelled the damage in an Euler–Bernoulli beam as rotational spring whose compliance is evaluated using linear elastic fracture mechanics. All the above studies are related to simple open crack model.

G. R. Liu et al. (eds.), Computational Methods, 1963–1968.

In reality, partial crack closure often occurs due to (1) roughness interference, (2) wedging by corrosion or wear debris, and (3) elastic constraint in the wake of the plastic zone. The closure effects on the vibration response of a fatigue cracked steel Tee–beam were investigated experimentally by Zhang and Testa [4] Law and Zhu [3] used a damage function to model the crack zone in the reinforced concrete beam. In this paper, the dynamic behaviour of a Bernoulli–Euler beam with a breathing crack subject to a moving mass is studied. The two segments of the beam, separated by the crack, are related to one another by time varying connection matrices representing the interaction forces. Two different sets of admissible functions which satisfy the respective geometric boundary conditions are assumed for these fictitious sub-beams. Numerical procedure is presented for the dynamic response of a simple beam subject to moving loads.

2. DAMAGED BEAM ELEMENT WITH BREATHING CRACK

2.1 Rotational Spring Model

The rotational spring model is widely used to study cracked beams, in which the effect of structural damage is modeled through a local compliance [5]. There are many existing models to describe the deterioration in the structures, especially for inelastic structures. According to Richard and Abbott's hysteretic model [6], the moment-rotation relation at the damage cross-section of a beam can be written as,

$$M = \left[k_p + (k_0 - k_p)/\left(1 + \left|(k_0 - k_p)|\theta|\right|/M_0\right)^{1/n}\right]|\theta| \tag{1}$$

and the corresponding tangent stiffness is

$$k_{rd} = dM/d\phi = \left[k_p + (k_0 - k_p)/\left(1 + \left|(k_0 - k_p)|\theta|/M_0\right|^n\right)^{(n+1)/n}\right] \tag{2}$$

where k_0 is the initial rotational stiffness when the crack is closing and k_p is the stiffness when the crack full opens, M_0 is a reference moment and n is a parameter defining the curvature of the curve.

2.2 Cracked Beam Element

Figure 1 shows a prismatic isotropic beam element of length l with a breathing crack which is located at $l_c (l_c \neq 0)$ from left end. The cracked beam element is modeled as two segment beam connected by a virtual spring. Assuming the

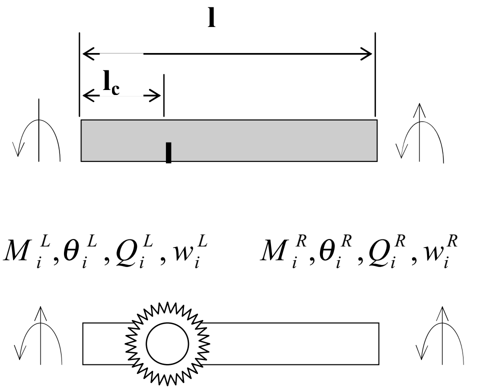


Figure 1. Cracked beam element.

rigidity of the undamaged beam is *EI*, the stiffness matrices link up the force and deformation at the ends of two segment beams are

$$\begin{Bmatrix} Q_i \\ M_i \\ Q_d^L \\ M_d^L \end{Bmatrix} = \frac{EI}{l_c^3} \begin{bmatrix} 12 & 6l_c & -12 & 6l_c \\ 6l_c & 4l_c^2 & -6l_c & 2l_c^2 \\ -12 & -6l_c & 12 & -6l_c \\ 6l_c & 2l_c^2 & -6l_c & 4l_c^2 \end{bmatrix} \begin{Bmatrix} w_i \\ \theta_i \\ w_d^L \\ \theta_d^L \end{Bmatrix}$$

$$\begin{Bmatrix} Q_d^R \\ M_d^R \\ Q_j \\ M_j \end{Bmatrix} = \frac{El}{(l-l_c)^3} \begin{bmatrix} 12 & 6(l-l_c) & -12 & 6(l-l_c) \\ 6(l-l_c) & 4(l-l_c)^2 & -6(l-l_c) & 2(l-l_c)^2 \\ -12 & -6(l-l_c) & 12 & -6(l-l_c) \\ 6(l-l_c) & 2(l-l_c)^2 & -6(l-l_c) & 4(l-l_c)^2 \end{bmatrix} \begin{bmatrix} w_d^R \\ \theta_d^R \\ w_j \\ \theta_j \end{bmatrix} \tag{3}$$

where $w_i, w_j, \theta_i, \theta_j$ are the displacements and rotations of two end nodes of the elements, Q_i, Q_j, M_i, M_j are the corresponding lateral shear forces and moments, $w_d^L, \theta_d^L, w_d^R, \theta_d^R$ are the two end displacements and rotations of the spring, Q_d^L, M_d^L, Q_d^R, M_d^R are the corresponding shear forces and moments.

According to the compatibility and equilibrium relations at the crack, the cracked element matrix can be given

$$\mathbf{K}_d = \mathbf{K}_1 + \mathbf{K}_2\mathbf{K}_3^{-1}\mathbf{K}_4 \tag{4}$$

where

$$\mathbf{K}_1 = \frac{EI}{l^3} \begin{bmatrix} 12/\delta^3 & 6l/\delta^2 & 0 & 0 \\ 6l/\delta^2 & 4l^2/\delta & 0 & 0 \\ 0 & 0 & 12/(1-\delta)^3 & -6l/(1-\delta)^2 \\ 0 & 0 & -6l/(1-\delta)^2 & 4l^2/(1-\delta) \end{bmatrix},$$

$$\mathbf{K}_2 = \frac{EI}{l^3}\begin{bmatrix} -12/\delta^3 & 6l/\delta^2 & 0 & 0 \\ -6l/\delta^2 & 2l^2/\delta & 0 & 0 \\ 0 & 0 & -12/(1-\delta)^3 & -6l/(1-\delta)^2 \\ 0 & 0 & 6l/(1-\delta)^2 & 2l^2/(1-\delta) \end{bmatrix}$$

$$\mathbf{K}_3 = \begin{bmatrix} -12/\delta^3 & 6l/\delta^2 & 0 & 0 \\ 6l/\delta^2 & S-4l^2/\delta & 0 & 0 \\ 0 & 0 & 12/(1-\delta)^3 & -6l/(1-\delta)^2 \\ 0 & -S & -6l/(1-\delta)^2 & S-4l^2/(1-\delta) \end{bmatrix},$$

$$\mathbf{K}_4 = \begin{bmatrix} -12/\delta^3 & -6l/\delta^2 & 0 & 0 \\ 6l/\delta^2 & 2l^2/\delta & 0 & 0 \\ 0 & 0 & -12/(1-\delta)^3 & 6l/(1-\delta)^2 \\ 0 & 0 & -6l/(1-\delta)^2 & 2l^2/(1-\delta) \end{bmatrix}$$

$$\delta = l_c/l,\ S = K_{rd}l^3/EI$$

3. EQUATION OF MOTION

Figure 2 shows the simply-supported beam under a moving mass m with constant speed v. The crack is located at l_d from left end. The supporting beam structure is discretized into $N-1$ beam element where N is the number of nodal points. With the assumption of Rayleigh damping, the equation of motion for the bridge can be written as

$$\mathbf{M}_b\ddot{\mathbf{R}} + \mathbf{C}_b\dot{\mathbf{R}} + \mathbf{K}_b\mathbf{R} = \mathbf{H}_m P \tag{5}$$

where $\mathbf{M}_b$, $\mathbf{K}_b$, $\mathbf{C}_b$ are the mass, stiffness and damping matrices of the simply-supported beam. $\ddot{\mathbf{R}}$, $\dot{\mathbf{R}}$, $\mathbf{R}$ are the nodal acceleration, velocity and displacement vectors of the beam respectively, and $\boldsymbol{H}_m P$ is the equivalent nodal load vector from the moving mass.

$$P = m\{g - d^2(w(x(t), t))/dt^2\} \tag{6}$$

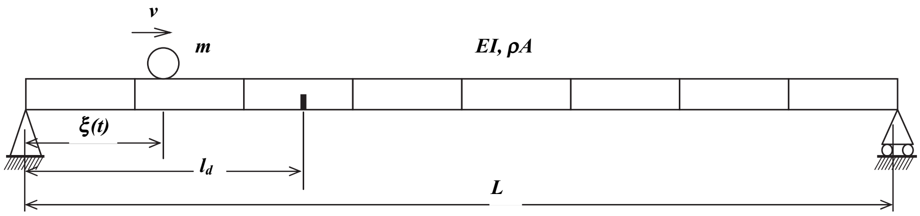


Figure 2. The damaged beam under a moving mass.

where m is the moving mass. $\mathbf{H}_m = \{0 \quad \cdots \quad H(\xi(t))_j^T \quad 0 \quad \cdots \quad 0\}^T$ with $(j-1)l \leq \xi(t) \leq jl$. $H(x)$ is the shape function of the beam element. The deflection of the beam at position x and time t can then be expressed as

$$w(x,t) = \mathbf{H}(x)\mathbf{R} \tag{7}$$

where $\mathbf{H}(x) = \{0 \quad \cdots \quad H(x)_j^T \quad 0 \quad \cdots \quad 0\}$ with $(j-1)l \leq x \leq jl$. Substitute Equation (7) into Equations (5) and (6), the equation of motion of the system is as follow

$$\mathbf{M}(t)\ddot{\mathbf{R}} + \mathbf{C}(t)\dot{\mathbf{R}} + \mathbf{K}(t)\mathbf{R} = mg\mathbf{H}_m \tag{8}$$

where

$$\begin{aligned}&\mathbf{M}(t) = \mathbf{M}_b + m\mathbf{H}_m\mathbf{H}(x), \quad \mathbf{C}(t) = \mathbf{C}_b + 2mv\mathbf{H}_m\mathbf{H}'(x),\\ &\mathbf{K}(t) = \mathbf{K}_b + mv^2\mathbf{H}_m\mathbf{H}''(x)\ .\end{aligned}$$

$\mathbf{H}'(x), \mathbf{H}''(x)$ are the first and second deviation of $\boldsymbol{H}(x)$. Newton–Raphson iteration algorithm is used to solve Equation (8). Then, dynamic responses of the beam subjected to a moving mass can be obtained from Equation (7).

4. CONCLUSION

A cracked beam element has been developed to analyze the dynamic behaviour of a beam with a breathing crack subjected to a moving mass. It is useful to study the nonlinear behaviour of the damage bridge structures subjected to moving vehicular loads.

ACKNOWLEDGEMENT

The work described in this paper was supported by a grant from the Hong Kong Polytechnic University Research Funding Project No.G-YW98.

REFERENCES

1. H.P. Lee and T.Y. Ng (1994), Dynamic response of a cracked beam subject to a moving load. *Acta Mechanica*, 106, pp. 221–230.
2. D.R. Parhi and A.K. Behera (1997), Dynamic deflection of a cracked beam with moving mass. In: *Proceedings of the Institute of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 211, pp. 77–87.

3. S.S. Law and X.Q. Zhu (2004), Dynamic behaviour of damaged concrete bridge structures under moving vehicular loads. *Engineering Structures*, 26, 9, pp. 1279–1293.
4. W.Z. Zhang and R.B. Testa (1999), Closure effects on fatigue crack detection. *Journal of Engineering Mechanics ASCE*, 125, 10, pp. 1125–1132.
5. C. Bilello and L.A. Bergman (2004), Vibration of damaged beams under a moving mass: theory and experimental validation. *Journal of Sound and Vibration*, 274, 3–5, pp. 567–582.
6. R.M. Richard and B.J. Abbott (1975), Versatile elastic-plastic stress-strain formula. *Journal of Engineering Mechanics Division ASCE*, 101, 4, pp. 511–515.

VALVE GEAR VIBRATIONAL ANALYSIS AND DEVELOPMENT OF NEW CAM DESIGN

J. Avsec and M. Oblak
University of Maribor, Faculty of Mechanical Engineering, Smetanova 17, 2000 Maribor, P.O. Box 224, Slovenia

Abstract This paper describes a valve gear vibrational analysis. In order to calculate the kinematic and dynamic values and to assess the minimum oil film thickness in the valve gear, the vibrational mathematical multidegree of freedom model of valve gear was used. In addition, the comparison of the results between the polysine cam and the new MULTICAM cam design was made. By means of the new cam design the Hertz pressures were reduced at the point of contact between the cam and the cam follower and the lubrication properties at the top of the cam improved.

Keywords: vibrational analysis, valve gear system, Runge–Kutta, method.

1. INTRODUCTION

The main task of the valve gear is the exhaust and inlet valve control. In internal combustion engines, the valve gear system has an important influence upon the power output and torque as well as on the exhaust gas emission. The forces required for the control are very high due to high accelerations occurring in the valve gear. Apart from the forces, wear and damage of the valve gear are caused also by the reactive reagents and high temperatures primarily on the exhaust side of the thermally loaded valve gear parts. The valve gear of truck and bus engines has experienced a fast development in the past years. In Otto engines for cars the electro-hydraulic valve gear with variable valve gear design has been established in the series production. It has an advantage over the conventional system primarily at the idling at lower engine loading. In racing car engines, the pneumatic valve gear is used already due to a very fast response and the correspondingly very short valve opening and closing times.

G. R. Liu et al. (eds.), Computational Methods, 1969–1974.

In truck and bus diesel engines only conventional valve gear designs are used in the series production.

2. EXISTING VALVE GEAR

In our case, the engine durability test was carried out in TAM BF 6L 515C diesel engine, incorporating a conventional valve gear (Figure 1). After the 3000-hour engine test the manifestations of an intensive wear appeared on the exhaust side of the valve gear. The wear of the exhaust cam was especially intensive right under the top. This type of wear is called the surface wear and it occurs at low peripheral speeds, when the lubrication properties are weakened and high friction occurs between the cam and the follower when the side pressure is high. Both the inlet and exhaust cams in the test engine were designed in accordance with the polysine curve theory [1]. This type of cam is designed so as to yield a continuous acceleration curve.

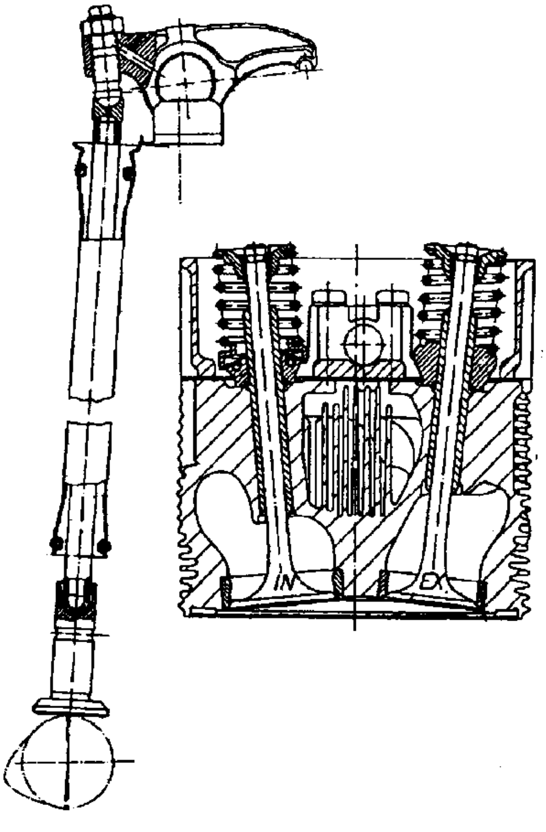


Figure 1. Valve gear system.

3. NEW MULTICAM CAM DESIGN

By using the dynamic model of the valve gear we analysed the causes of excessive wear. Since the cam (NI1), designed in accordance with the polysine curve, offered too few possibilities for an optimum cam profile, we wanted to manufacture a completely new type of cam with more possibilities for an optimum adjustment. At the same time, we wished to summarise some findings of authors on the dependence of the cam design on the fuel consumption and valve gear noise. Thus, on this basis the new MULTICAM cam (NI2) was created. Contrary to the conventional theory of polysine cam the motion in MULTICAM cam can be written by means of seven curves. Equations and boundary conditions for MULTICAM cam is possibly to find in the literature of Avsec et al. [2].

4. MATHEMATICAL VALVE GEAR DYNAMIC MODEL

The real valve gear system is unsuitable for the formulation of an equation of movement, mainly because it is described by means of partial differential equations. Consequently, the real system is replaced by an equivalent three-mass system (shown in Figure 2), which may be described through three ordinary second order differential equations. In this reaction, every element of the system is represented by two concentrated masses, connected by a weightless spring, having the stiffness of this element. In addition, rotational movement of a rocker arm is replaced by translatory movement. Portions of masses of two adjacent elements are integrated into a single mass (e.g., a portion of the mass of a push rod and a portion of the mass of a rocker arm form a single mass in the equivalent system), resulting in a system of three concentrated masses, interconnected with springs. Reduced masses and reduced stiffnesses of the equivalent system are determined on the basis of equality of kinetic and potential energy of the real and equivalent system. By introducing these variables into the system of dynamics equations a final form of differential equations is obtained, describing the movement of the equivalent model of the valve gear system:

$$\begin{bmatrix} m_1+m_0 & 0 & 0 \\ m_2 & -m_2 & 0 \\ m_3^* & -m_3^* & -m_3^* \end{bmatrix} \begin{Bmatrix} \ddot{z}_1 \\ \ddot{z}_2 \\ \ddot{z}_3 \end{Bmatrix} + \begin{bmatrix} C_{bv} & C_S & 0 \\ 0 & -C_S & C_k \\ C_0^* & -C_0^* & -(C_0^*+C_k) \end{bmatrix} \begin{Bmatrix} z_1 \\ z_2 \\ z_3 \end{Bmatrix} = \begin{Bmatrix} m_0\ddot{h}_p + C_{bv}h_p \\ 0 \\ F^* \end{Bmatrix} \tag{1}$$

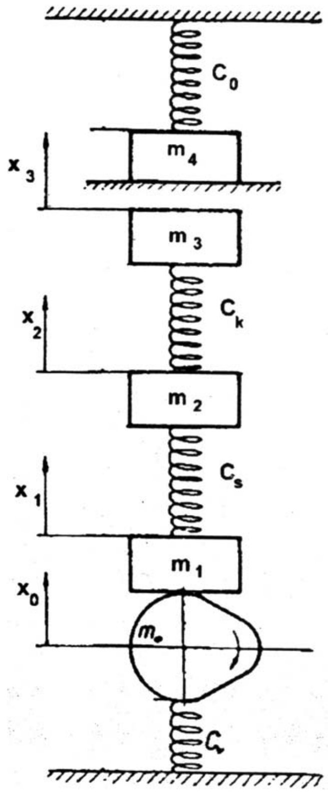


Figure 2. The mathematical model.

We have solved Equation (1) with the help of Runge–Kutta 4 numerical method. By means of the dynamic analysis we then carried out the review of values having a direct impact upon the valve gear wear: Control of Hertz pressures; Control of minimum oil film thickness according to elastohydrodynamic theory; Control of tangential stresses occurring some tenths of millimetres in the depth of the cam and being proportional to the normal tensions. Tables 1 and 2 show the values of the highest Hertz pressure and the minimum oil film thickness occurring in the inlet and outlet cam at the idling (500 rpm of crankshaft) and at

Table 1. Computation of force F_1 between cam follower and the push rod and minimum oil film thickness H_{min} at the top of cam.

	Inlet system	Exhaust system
F_1	1521 N	1109 N
H_{min}	0.142 μm	0.13 μm

Table 2. Computation of normal tensions (Hertz pressures) at the top of the cam.

Crankshaft velocity (rpm)	Exhaust cam (N/mm^2)	Inlet cam (N/mm^2)
500	722	590
2800	560	480

rpm exceeded by a 30% (2800 rpm of crankshaft), which can appear by vehicle moving downhill. In order to calculate the Hertz pressures we used the theory of rolling contacts [1]:

$$p_H(\mathrm{N/mm^2}) = 0.418\sqrt{\frac{F_1(\mathrm{N})E(\mathrm{N/mm^2})}{b_N(\mathrm{mm})}\left(\frac{1}{R_p(\mathrm{mm})}+\frac{1}{\mathrm{RON}(\mathrm{mm})}\right)} \quad (2)$$

According to the sources from the References [1, 2] the permissible Hertzian pressure loads are approximately 600 N/mm^2. As evident from Table 2, the loads are too high mainly on the exhaust cam. The minimum oil film thickness H_{min} was calculated using the elastohydrodynamic theory [2]:

$$H_{min}(\mathrm{m}) = 1.6 \cdot \mathrm{RON}(\mathrm{m})^{0.43}\eta(\mathrm{Pas})^{0.2}\,|v_{HD}(\mathrm{m/s})|^{0.7}\,(F_1(\mathrm{N}))^{-0.13}E(\mathrm{N/m^2})^{0.03} \quad (3)$$

The assessment of the minimum oil film thickness at the top of the exhaust cam (Table 1) does not provide any favourable results. As the largest loads appear at the top of the cam, where the highest wear was measured, it is necessary to reduce the normal tensions and improve the lubrication properties. Table 2 shows that the tensions at the top of the exhaust cam must primarily be lowered (SI1). This was achieved in two ways: The design exhaust cam allows an increase in the base circle diameter from 41.6 mm to 45.6 mm. By means of a new base circle diameter we calculated the new optimum polysine cam design (NI1).

5. RESULTS AND DISCUSSION

The comparison of the cam profiles is given in Figure 3. Figure 4 show the analytical calculated cam lift for the new, NI2 cam. This paper describes our theoretical and practical contribution to the optimization of valve gear design, including the ramp period. For these purpose we developed a new cam, called MULTICAM, which consists from seven different curves. To estimate the kinematics, dynamical and lubrication properties with the new cam profile the vibrational model of valve gear was used. At the new cam profile the Hertz

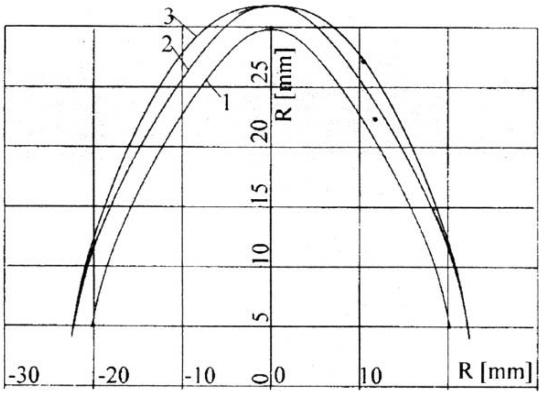


Figure 3. Cam profiles (1-SI1, 2-NI1, 3-NI2).

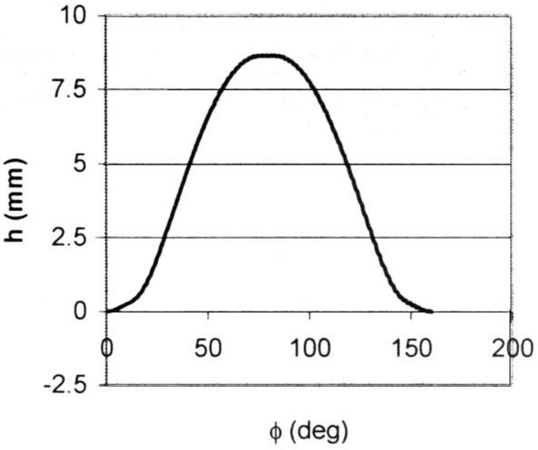


Figure 4. Cam lift for NI2 model obtained by vibrational analysis.

pressure on the contact between cam and cam follower was reduced and also oil film thickness was improved. The calculation of other important values, as for example the ratio between inertia forces and spring forces, shows favourable results too [2].

REFERENCES

1. V. Küntscher (1987), "Kraftfahrzeugmotoren", VEB Verlag Technik, Berlin.
2. J. Avsec, M. Marcic and M. Oblak (2002), Valve gear refinement, Journal of Mechanical Design, 124.

NATURAL FREQUENCY OF STEPPED BEAM HAVING MULTIPLE OPEN CRACKS BY TRANSFER MATRIX METHOD

Q. Chen, S.C. Fan and D.Y. Zheng
School of Civil and Environmental Engineering, Nanyang Technological University,
Nanyang Avenue, Singapore 639798

Abstract In this study, the transfer matrix method is used for the dynamic analysis of a stepped beam with arbitrary multiple transverse open cracks. The reduction in bending stiffness due to the presence of transverse open cracks or abrupt changes of cross-section is modelled by kind of massless rotational springs. The advantages are demonstrated through examples. Firstly, it yields purely analytical solutions which are more accurate than the numerical ones. Secondly, the size of the resulting eigen-matrix is small. For beams with all sorts of boundary restraint conditions, having many abrupt changes of cross-sections and/or with arbitrary multiple open cracks, the size of the eigen-matrix is still 4 by 4 (or less). Numerical examples are presented to validate the accuracy and efficiency of the present formulation.

Keywords: transfer matrix, stepped beam, multiple cracks, massless rotational spring.

1. INTRODUCTION

Knowing the dynamic behaviour of a stepped beam with cracks is of significant importance in engineering. In the past two decades, the vibration problems of an uncracked stepped beam or a cracked uniform beam have attracted the attentions of many researchers. A wealth of literature is available. However, less attention has been paid to the vibration problem of a stepped beam having cracks. Dimarogonas [1] presented a state-of-the-art review on methods available for solving the cracked beam problems. Recently, Li [2] presented solutions for a cracked multiple-stepped beam using analytical functions of vibration mode shape and rotational spring at the location with abrupt change of cross-sections. On the other hand, Khiem and Lien [3] used the transfer matrix

G. R. Liu et al. (eds.), Computational Methods, 1975–1979.

method to obtain natural frequencies of a uniform beam having multiple cracks. In this paper, the transfer matrix method is employed in conjunction with the implantation of artificial rotational springs to investigate the natural frequencies of a stepped beam having multiple open cracks. Essentially, the merits of the transfer matrix method are inherited, and the simplicity of modelling cracks by massless spring [4] is preserved.

2. THEORY AND FORMULATION

Firstly, consider a uniform beam segment j, of which the left end is at $x = x^{+}_{j-1}$ and the right end is at $x = x^{-}_{j}$ (where the superscript '$+$' denotes the right-hand side of the left-end junction; whereas the superscript '$-$' denotes the left-hand side of the right-end junction) having cross-sectional area A, moment of inertia I, mass density ρ and Young's modulus E. The general solution $\varphi(x)$ for the lateral free-vibration equation $\varphi''(x) - \lambda^4\varphi(x) = 0$ is:

$$\varphi(x) = \sum_{j=1}^{4} C_j K_j(\lambda^{-}x) \tag{1}$$

where $\lambda^4 = \rho A\omega^2/EI; \bar{x} = x^{-}_{j} - x^{+}_{j-1}; K_j$'s are functions: $K_1(\lambda\bar{x}) = (\cosh\lambda\bar{x} + \cos\lambda\bar{x})/2$, $K_2(\lambda\bar{x}) = (\cosh\lambda\bar{x} - \cos\lambda\bar{x})/2$, $K_3(\lambda\bar{x}) = (\sinh\lambda\bar{x} + \sin\lambda\bar{x})/2$, $K_4(\lambda\bar{x}) = (\sinh\lambda\bar{x} + \sin\lambda\bar{x})/2$; and C_j's are the coefficients to be determined. Write the state conditions at the left end ($x = x^{+}_{j-1}$) and the right end ($x = x^{-}_{j}$) as follows:

$$\begin{aligned}
\mathbf{Z}^{+}_{j-1} &= \{Z^{+}_{j-1,1}, Z^{+}_{j-1,2}, Z^{+}_{j-1,3}, Z^{+}_{j-1,4}\}^T \\
&= \{\varphi(x^{+}_{j-1}); \varphi'(x^{+}_{j-1}); -EI\varphi'''(x^{+}_{j-1}); El\varphi''(x^{+}_{j-1})\}^T; \\
\mathbf{Z}^{-}_{j} &= \{Z^{-}_{j1}, Z^{-}_{j2}, Z^{-}_{j3}, Z^{-}_{j4}\}^T = \{\varphi(x^{-}_{j}); \varphi'(x^{-}_{j}); EI\varphi'''(x^{-}_{j}); -EI\varphi''(x^{-}_{j})\}^T
\end{aligned} \tag{2}$$

Substituting the state conditions at the left end ($x = x^{+}_{j-1}$) into Equation (1) yields $C_1 = Z^{+}_{j-1,1}$, $C_2 = Z^{+}_{j-1,2}/\lambda$, $C_3 = Z^{+}_{j-1,4}/EI\lambda^2$, $C_4 = Z^{+}_{j-1,3}/EI\lambda^3$. Subsequently, the right-end ($x = x^{-}_{j}$) state condition can be expressed in terms of the left-end state conditions through Equation (1). In matrix form,

$$\mathbf{Z}^{-}_{j} = \mathbf{T}_j\mathbf{Z}^{+}_{j-1} \tag{3}$$

$$\mathbf{T}_j(\lambda, \bar{x}) = \begin{bmatrix}
K_1(\lambda, \bar{x}) & \lambda^{-1}K_2(\lambda\bar{x}) & K_4(\lambda\bar{x})/EI\lambda^3 & -K_3(\lambda\bar{x})/EI\lambda^2 \\
\lambda K_4(\lambda\bar{x}) & K_1(\lambda\bar{x}) & K_3(\lambda\bar{x})/EI & -K_2(\lambda\bar{x})/EI\lambda \\
-\lambda^3 EIK_2(\lambda\bar{x}) & -\lambda^2 EIK_3(\lambda\bar{x}) & -K_1(\lambda\bar{x}) & \lambda K_4(\lambda\bar{x}) \\
\lambda^2 EIK_3(\lambda\bar{x}) & \lambda EIK_4(\lambda\bar{x}) & \lambda^{-1}K_2(\lambda\bar{x}) & -K_1(\lambda\bar{x})
\end{bmatrix} \tag{4}$$

Secondly, consider the continuity conditions at the junctions between segments. Regardless of the presence of abrupt change of cross section at a junction, the following continuity conditions at an arbitrary junction (at x_j) must be satisfied; i.e.,

$$\varphi(x_j^-) = \varphi(x_j^+), \quad \varphi'(x_j^-) = \varphi'(x_j^+), \quad E_j I_j \varphi''(x_j^-) = E_{j+1} I_{j+1} \varphi''(x_j^+),$$
$$E_j I_j \varphi'''(x_j^-) = E_{j+1} I_{j+1} \varphi'''(x_j^+) \tag{5}$$

By using the same notation defined above, Equation (5) can be re-written in matrix form; i.e.,

$$\mathbf{Z}_j^+ = \mathbf{J}_j(0)\mathbf{Z}_j^- \qquad \text{where} \quad \mathbf{J}_j(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{6}$$

Thirdly, consider the continuity and discontinuity conditions at the junctions where an open crack is present between two segments. An artificial massless rotational spring having a flexibility coefficient β_k is introduced to account for the discontinuity of the first derivative. In this case, Equations (5) and (6) become

$$\varphi(x_j^-) = \varphi(x_j^+), \quad \varphi'(x_j^-) + \beta_k \varphi''(x_j^-) = \varphi'(x_j^+), \quad \varphi''(x_j^-) = \varphi''(x_j^+),$$
$$\varphi'''(x_j^-) = \varphi'''(x_j^+) \tag{7}$$

$$\mathbf{Z}_j^+ = \mathbf{J}_j(\beta_k)\mathbf{Z}_j^- \qquad \text{where} \quad \mathbf{J}_j(\beta_k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \beta_k \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{8}$$

Note that by setting $\beta_k = 0$, Equation (8) is reduced to the special case as in Equation (6). The flexibility coefficient β_k (for the kth open crack) can be expressed as a function of normalized crack depth ξ_k [4] where $\xi_k = \alpha_k / h_k(\alpha_k$ is the depth of the open crack and h_k is depth of the intact section without crack); i.e.,

$$\beta_k = 5.346\, h_k f(\xi_k)/EI \tag{9a}$$

$$f(\xi) = 1.8624\,\xi^2 - 3.95\,\xi^3 + 16.375\,\xi^4 - 37.226\,\xi^5 + 76.81\,\xi^6$$
$$-126.9\,\xi^7 + 172\,\xi^8 - 143.97\,\xi^9 + 66.56\,\xi^{10} \tag{9b}$$

Fourthly, a chain relationship can now be established. Starting from the extreme right-end segment (n) of a beam, we can relate the right-end state conditions to those of the left-end through Equation (3); i.e., $\mathbf{Z}_n^- = \mathbf{T}_n \mathbf{Z}_{n-1}^+$. Then, combining with Equation (6) or (8) leads to $\mathbf{Z}_n^- = \mathbf{T}_n \mathbf{Z}_{n-1}^+ = (\mathbf{T}_n \mathbf{J}_{n-1})\mathbf{Z}_{n-1}^-$,

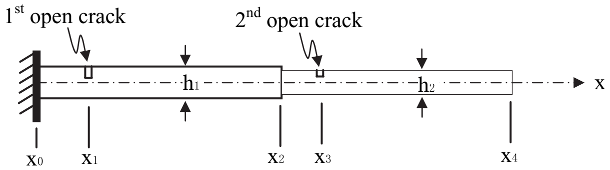


Figure 1. A cantilever beam with two steps and two open cracks.

which is the typical chain relationship between the right-end state conditions of two consecutive segments. Chaining up all segments yields

$$\begin{aligned}\mathbf{Z}_n^- &= (\mathbf{T}_n\mathbf{J}_{n-1})\,\mathbf{Z}_{n-1}^- = (\mathbf{T}_n\mathbf{J}_{n-1})(\mathbf{T}_{n-1}\mathbf{J}_{n-2})\mathbf{Z}_{n-2}^- = \cdots \\ &= (\mathbf{T}_n\mathbf{J}_{n-1})(\mathbf{T}_{n-1}\mathbf{J}_{n-2})\ldots(\mathbf{T}_2\mathbf{J}_1)\mathbf{Z}_1^- \end{aligned} \tag{10}$$

Substituting $\mathbf{Z}_1^- = \mathbf{T}_1\mathbf{Z}_0^+$ into Equation (10) leads to

$$\mathbf{Z}_n^- = [(\mathbf{T}_n\mathbf{J}_{n-1})(\mathbf{T}_{n-1}\mathbf{J}_{n-2})\mathbf{Z}_{n-2}^- \ldots (\mathbf{T}_2\mathbf{J}_1)\mathbf{T}_1]\mathbf{Z}_0^+ \tag{11}$$

Letting $\mathbf{Q} = (\mathbf{T}_n\mathbf{J}_{n-1})(\mathbf{T}_{n-1}\mathbf{J}_{n-2})\ldots(\mathbf{T}_2\mathbf{J}_1)\mathbf{T}_1$, we have $\mathbf{Z}_n^- = \mathbf{Q}\mathbf{Z}_0^+$, in which **Q** relates the state conditions of the extreme right-end to those of the extreme left-end. Note that **Q** is 4 × 4 matrix having its 16 elements q_{mn}, which are functions of the frequency ω only. By imposing the boundary constraints at the two ends, **Q** will be often reduced to a 2 × 2 eigen-matrix and subsequently to a single frequency equation. For example,

$$q_{12}q_{34} - q_{32}q_{14} = 0 \quad \text{(for both ends simple supported)} \tag{12}$$

$$q_{13}q_{24} - q_{23}q_{14} = 0 \quad \text{(for both ends fixed)} \tag{13}$$

$$q_{33}q_{44} - q_{43}q_{34} = 0 \quad \text{(for cantilevered beam left-end fixed)} \tag{14}$$

Example. To validate the present formulation, a two-stepped cantelever beam having two one-sided open cracks is analysed (see Figure 1). The data are Section 1 (width $b = 0.3\,m$, depth $h_1 = 0.7\,m$), Section 2 (same b, but $h_2 = 0.5\,m$); 4 segments ($x_0 = 0, x_1$ *varies*, $x_2 = 5.5\,m, x_3 = 6.5\,m, x_4 = 11\,m$); two cracks (depth $a_1 = 0.3\,m$ at x_1; depth $a_2 = 0.2\,m$ at x_3; $E = 210\,GPa$; $\rho = 7800\,kg/m^3$. Good accord with finite-element results [5] is observed (see Figure 2).

3. CONCLUSIONS

Modelling of open crack by rotational spring in conjunction with application of the transfer matrix method was presented. In solving the free vibration frequency of a multiple stepped beam having arbitrary number of open cracks, it always leads to a simple eigen-matrix of size only 4 × 4 (or less for common

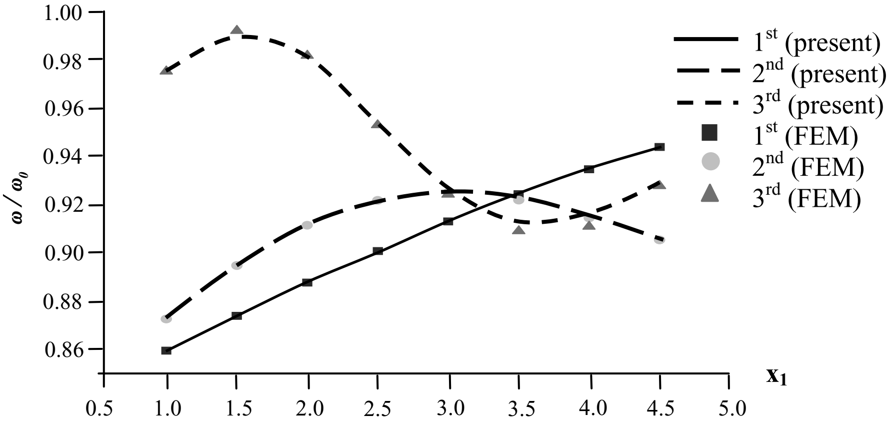


Figure 2. Frequency reduction (ω/ω_0) for the first 3 modes against varies x_1.

boundary conditions). The example showed that the present formulation is simple, versatile, effective and accurate.

REFERENCES

1. A.D. Dimarogonas (1996), Vibration of cracked structures: a state of the art review. *Enngineering Fracture Mechanics*, **55**, pp. 831–857.
2. Q.S. Li (2001), Vibratory characteristics of multi-step beams with an arbitrary number of cracks and concentrated masses. *Applied Acoustics*, **62**, pp. 691–706.
3. N.T. Khiem and T.V. Lien (2001), A simplified method for natural frequency analysis of a multiple cracked beam. *Journal of Sound and Vibration*, **245**, **4**, pp. 737–751.
4. T.G. Chondros, A.D. Dimarognas and J. Yao (1998), A continuous cracked beam vibration theory. *Journal of Sound and Vibration*, **215**, pp. 17–34.
5. D.Y. Zheng and N.J. Kessissoglou (2004), Free vibration analysis of a cracked beam by finite element method. *Journal of Sound and Vibration*, **273**, pp. 457–475.

FUNDAMENTAL MATRIX ESTIMATION BASED ON A GENERALIZED EIGENVALUE PROBLEM

H. X. Zhong, Y.P. Feng and Y.J. Pang
College of Computer Science and Technology, Jilin University, Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, Changchun, 130012, P.R. China

Abstract A new method for estimating the fundamental matrix is proposed. Using eigenvectors corresponding to the two smallest eigenvalues obtained by the orthogonal least-squares technique, we construct a 3×3 generalized eigenvalue problem. Its solution gives not only the fundamental matrix but also the corresponding epipoles. The new method performs well as compared with several existing linear methods.

Keywords: fundamental matrix, epipolar geometry, orthogonal least-squares, generalized eigenvalue problem.

1. INTRODUCTION

Two images of a single rigid scene are related by the epipolar geometry, which can be described by a matrix called fundamental matrix. It is a 3×3 matrix of rank 2 [1]. It is independent of scene structure, and can be computed from correspondences of image points alone. Many methods have been proposed for estimating fundamental matrix. They can be divided into three classes [2] linear, iteration and robust. Linear methods are foundations of other methods.

In this paper, we propose a new linear method for estimating the fundamental matrix by constructing a 3×3 generalized eigenvalue problem. The performance of the new method is compared with other linear techniques.

G. R. Liu et al. (eds.), Computational Methods, 1981–1985.

2. DESCRIPTION OF THE METHOD

Assume that two image points of a same 3D space point in image I and I' are $\mathbf{m} = (x, y, 1)^T$ and $\mathbf{m}' = (x', y', 1)^T$ respectively, then the epipolar geometry relation:

$$\mathbf{m}^T\mathbf{F}\mathbf{m}' = 0 \tag{1}$$

is satisfied, where the 3×3 singular matrix $\mathbf{F}$ is the fundamental matrix. For the epipoles e′ and e, we have

$$\mathbf{F}\mathbf{e}' = 0, \quad \mathbf{F}^T\mathbf{e} = 0 \tag{2}$$

For a set of correspondences $\{(\mathbf{m}_i, \mathbf{m}_i')|i = 1, 2, \ldots, n\}$ with $\mathbf{m}_i = (x_i, y_i, 1)^T$, $\mathbf{m}_i' = (x_i', y_i', 1)^T$ and $n \geq 7$, Equation (1) can be writing as:

$$\mathbf{U}_n\mathbf{f} = 0 \tag{3}$$

where

$$\mathbf{f} = (F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33})^T \tag{4}$$

$$\mathbf{U}_n = \begin{bmatrix} x_1x' & x_1y_1' & x_1 & y_1x_1' & y_1y_1' & y_1 & x_1' & y_1' & 1 \\ x_2x_2' & x_2y_2' & x_2 & y_2x_2' & y_2y_2' & y_2 & x_2' & y_2' & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_nx_n' & x_ny_n' & x_n & y_nx_n' & y_ny_n' & y_n & x_n' & y_n' & 1 \end{bmatrix}_{n\times 9} \tag{5}$$

It is important to note that a fundamental matrix has only 7 degrees of freedom. There are only 7 independent parameters among the 9 elements [3]. In Equation (3), only 7 parameters need to be solved and they can be expressed by the remaining two parameters. A least-squares method can be used to solve Equation (3):

$$\min_{\mathbf{f}\in R^9}(\mathbf{U}_n\mathbf{f}, \mathbf{U}_n\mathbf{f}) \tag{6}$$

where (·,·) denotes the Euclid inner product. The vector $\mathbf{f}$ is only defined up to an unknown scale factor. What we need is the non-trivial solution $\mathbf{f} \neq \mathbf{0}$. We thus impose the constraint on the fundamental matrix: $\|\mathbf{f}\| = 1$ where $\|\mathbf{f}\|$ denotes the Euclid norm of $\mathbf{f}$. Then Equation (6) becomes a classical constrained minimization one:

$$\min_{\|\mathbf{f}\|=1}(\mathbf{U}_n\mathbf{f}, \mathbf{U}_n\mathbf{f}) = \min_{\mathbf{f}\neq\mathbf{0}} \frac{(\mathbf{U}_n\mathbf{f}, \mathbf{U}_n\mathbf{f})}{(\mathbf{f}, \mathbf{f})} \tag{7}$$

We introduce a Lagrange function

$$L(\mathbf{f}, \lambda) = \|\mathbf{U}_n\mathbf{f}\|^2 + \omega\left(1 - \|\mathbf{f}\|^2\right) \tag{8}$$

where ω is the Lagrange multiplier. The solutions of Equation (7) result in the following eigenvalue problem:

$$\mathbf{U}_n^T \mathbf{U}_n \mathbf{f} = \omega \, \mathbf{f}, \quad \|\mathbf{f}\| = 1 \tag{9}$$

The nine eigenvalues of $\mathbf{U}_n^T \mathbf{U}_n$ are arranged in the increasing order: $0 \leq \omega_1 \leq \omega_2 \leq \ldots \leq \omega_9$, and the corresponding unit orthogonal eigenvectors are written as: $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_9$, respectively.

If the matrix $\mathbf{F}_1$ formed by $\mathbf{f}_1$ based on Equation (4) satisfies the rank-2 constraint, $\mathbf{F}_1$ is just the fundamental matrix. For $\mathbf{F}_1$, however, the rank-2 constraint will generally not be satisfied. We assume that $\mathbf{F}_1$ is not singular. To obtain the fundamental matrix with the rank-2 constraint, we approximately express the solution to Equation (3) in terms of

$$\mathbf{f} = \alpha \mathbf{f}_1 + \beta \mathbf{f}_2, \quad \alpha^2 + \beta^2 = 1 \tag{10}$$

Let $\mathbf{F}_1$ and $\mathbf{F}_2$ be 3×3 matrices formed by $\mathbf{f}_1$ and $\mathbf{f}_2$, respectively. The fundamental matrix $\mathbf{F}$ can be represented as $\mathbf{F} = \alpha \mathbf{F}_1 + \beta \mathbf{F}_2$. The singularity of $\mathbf{F}$ can be imposed by letting the following system of equations

$$(\alpha \mathbf{F}_1 + \beta \mathbf{F}_2) \mathbf{z} = 0 \tag{11}$$

has non-zero solution $\mathbf{z} \neq 0$. If $\beta = 0$, Equation (11) has trivial solution $\mathbf{z} = 0$ only, since $\mathbf{F}_1$ is non-singular. Therefore we have $\beta \neq 0$. Let $\lambda = -\alpha/\beta$, we get the following generalized eigenvalue problem

$$(\mathbf{F}_2 - \lambda \mathbf{F}_1) \mathbf{z} = 0, \quad \mathbf{z} \neq \mathbf{0}. \tag{12}$$

There is at least 1 real eigenvalue and at most 3 real ones for Equation (12). Let λ denote a real eigenvalue. Using Equation (10) and noticing $\lambda = -\alpha/\beta$ yields

$$\|\mathbf{U}_n \mathbf{f}\|_2{}^2 = \left(\frac{\lambda^2}{1 + \lambda^2} \right) \omega_1 + \left(\frac{1}{1 + \lambda^2} \right) \omega_2 \tag{13}$$

From Equation (13), we see that the second smallest eigenvalue is weighted by $1/(1 + \lambda^2)$ thus its contribution is reduced. For the case of Equation (12) having 3 real eigenvalues λ_1, λ_2 and λ_3, based on Equation (13), we choose $\lambda = \lambda_i$ where $|\lambda_i| = \max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\}$. Finally, the corresponding $\mathbf{F}$ and the epipole in I are given by:

$$\mathbf{F} = \frac{1}{\sqrt{1 + \lambda^2}} (\mathbf{F}_2 - \lambda \mathbf{F}_1), \quad \mathbf{e}' = \mathbf{z}_i \tag{14}$$

3. EXPERIMENTAL RESULTS

The method proposed in this paper is compared with other two different linear methods by using synthetic images with the Gaussian noise and outliers, and two real images shown in Figure 1. The methods include: Method 1-*the method proposed in this paper*, Method 2-*the orthogonal least-squares technique*, Method 3-*the modified orthogonal least-squares method*. We use 8 correspondences to calculate F. Data normalization technique proposed by Hartley and Zisserman [4] is used. The mean and standard deviation of the distance between points and epipolar lines computed for 100 trials are listed in Table 1. Method 2 obtains a rank-3 fundamental matrix. Method 3 yield rank-2 fundamental matrix, but distances become worse. For the proposed method, rank-2 constraint is satisfied, and mean and standard deviation of distance are smaller than other two methods. The results also indicate that the sensitivity to noise of the proposed method is weaker than other methods.

(1) (2)

Figure 1. Real images with correspondences: (1) Outdoor scene, (2) Indoor scene.

Table 1. Comparison of the mean and standard deviation of the distance (in pixels) between correspondences points and epipolar lines.

Image	Noise	Outlier (%)	Methods 1		Methods 2		Methods 3	
Synthetic image	0	0	0.00007	0.00009	0.0002	0.0003	0.0002	0.0002
	0.1	0	1.0614	1.1986	1.5623	2.092	1.605	1.9653
	0.5	0	3.4683	4.3818	3.9786	5.2487	4.1394	4.8859
	1.0	0	4.6128	5.0021	5.9686	7.7197	5.7432	5.9899
	0	12.5	4.0042	6.6476	5.5926	9.7352	6.1835	10.486
	0.1	12.5	4.1647	6.0372	5.6528	8.6008	5.8042	7.9764
	0.5	12.5	4.3803	5.961	6.9934	9.6698	7.1518	8.4868
	1.0	12.5	8.0714	9.2385	8.1241	10.592	8.4277	9.9636
Image 1			0.6133	0.5409	0.7914	0.8622	0.9997	0.9062
Image 2			2.1770	2.2602	2.4964	2.7708	2.4755	2.3879

4. CONCLUSIONS

In this paper, we present a new method for estimating the fundamental matrix. Using the two eigenvectors corresponding to the two smallest eigenvalues achieved by the orthogonal least-squares technique, we construct a 3 × 3 generalized eigenvalue problem. The solutions to the problem give not only the fundamental matrix but also the corresponding epipoles. The proposed approach is easy to be implemented, and the high quality results can be achieved.

REFERENCES

1. Q.-T. Luong and O.D. Faugeras (1996), The fundamental matrix: theory, algorithms, and stability analysis. *International Journal of Computer Vision*, 17,1, pp. 43–75.
2. X. Armangué and J. Salvi (2003), Overall view regarding fundamental matrix estimation. *Image and Vision Computing* 21, 2, pp. 205–220.
3. Z.Y. Zhang (1998), Determining the epipolar geometry and its uncertainty: a review. *International Journal of Computer Vision*, 27, 2, pp. 161–198.
4. R.I. Hartley and A. Zisserman (2000), Multiple View Geometry in Computer Vision, Cambridge University Press, London.

SKELETAL REDUCTION OF EIGEN-VALUE PROBLEMS OVER THIN SOLIDS

K. Suresh
Department of Mechanical Engineering, University of Wisconsin—Madison, 347 Mechanical Engineering Building, 1513 University Avenue, Madison, WI 53706

Abstract Extant dimensional reduction methods for thin solids are largely restricted to 'regular' solids, such as plates and shells that permit a mid-surface plus thickness representation. In this paper, we propose a skeletal reduction that is mathematically well-defined for all solids. The proposed method is illustrated and compared with the mid-surface based reduction using the Helmholtz eigen-value problem as a vehicle.

Keywords: dimensional reduction, medial axis transform, eigen-value, plates, shells.

1. INTRODUCTION

Consider the eigen-value problem:

$$\text{Find } \{u, \lambda\},\ u \in H_0^1(\Omega),\ \lambda \in \mathbb{C} \text{ such that}$$

$$\int_\Omega \nabla u . \nabla v d\Omega = \lambda \int_\Omega u v d\Omega,\ \forall v \in H_0^1(\Omega) \tag{1}$$

In particular, we focus here on solids Ω that are geometrically thin in that one or more spatial dimensions of Ω is much smaller than the third; examples of such solids include plates and shells. For thin solids, a fully discretized finite element approximation of (1) is not desirable for reasons stated, for example, in [1]; instead, a dimensional reduction is often recommended. Extant dimensional reduction methods hinge on a mid-surface representation of Ω, wherein Ω is expressed as a lower-dimensional manifold (a.k.a. mid-surface) ω plus a relatively small, but possibly varying thickness $H(\omega)$, i.e.:

$$\Omega = \{p + \hat{n} H(p)\varepsilon | p \in \omega(s,t), \hat{n} : \text{ normal}(p), -1 \leq \varepsilon \leq 1\}. \tag{2}$$

G. R. Liu et al. (eds.), Computational Methods, 1987–1991.

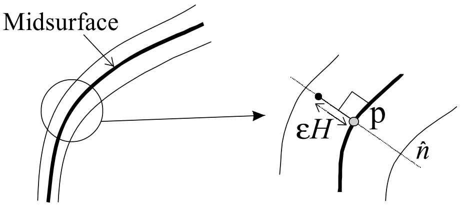


Figure 1. Example to illustrate (2).

Figure 1 illustrates the above definition. The mid-surface representation (2) leads to a convenient sub-space of solutions, typically consisting of low order polynomials in the thickness parameter ε [2]. A dimensionally reduced model of (1) can be obtained by seeking solutions in such sub-spaces. Dimensional reduction methods have been successfully applied in the analysis of thin solids such as plates and shells [3, 4].

Now consider the L-bracket in Figure 2 that is thin but exhibits an abrupt change in geometry. Once again, a fully discretized solution is not desirable since $H \ll L$. In order to apply extant dimensional reduction methods, one must first find a suitable quasi-disjoint decomposition $\Omega = \bigcup \Omega_i$ such that each Ω_i permits the representation (2). More formally, we seek a decomposition satisfying the following properties:

(1) $\Omega = \bigcup_{i=1}^{N} \Omega_i$ ... else the decomposition does not cover the original set.
(2) $\Omega_i \cap \Omega_{j \neq i} = \phi$ ... else the decomposition leads to ambiguity.
(3) $\Omega_i = \{p + \hat{n} H(p)\varepsilon | p \in \omega_i(s,t), \hat{n} : normal(p), -1 \leq \varepsilon \leq 1\}$ ... else the objective of dimensional reduction is not met.
(4) $\omega = \cup \omega_i$ must be connected ... else a lower-dimensional problem cannot be posed over ω. (3)

Unfortunately, for most solids that exhibit a sudden change in geometry (including the one in Figure 2), it is not possible to find such a decomposition. Further, violating one or more conditions from (3) results in both lack of

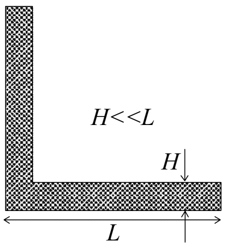


Figure 2. An L-bracket.

automation [5] and irrecoverable loss in accuracy, as illustrated in the numerical section below.

2. PROPOSED METHOD

It is objective of this paper to introduce a new method of dimensional reduction that can be applied to all thin polyhedral solids (non-polyhedral solids require a small modification not discussed here). The proposed method is based on the *skeletal representation* of Ω [6, 7] which guarantees that it is always possible to find a *skeletal decomposition* with the following properties:

(1) $\Omega = \bigcup_{i=1}^{N}(\Omega_i^+ \cup \Omega_i^-)$
(2) $\Omega_i^\pm \cap \Omega_{j \neq i}^\pm = \phi$
(3) Each $\Omega_i^\pm$ permits the skeletal representation: $\Omega_i^\pm = \left\{ p + \hat{k}_i^\pm R(p)\varepsilon | p \in m_i(s,t), \hat{k}_i^\pm : \text{projectors}(p), 0 \leq \varepsilon \leq 1 \right\}$
(4) $\mathfrak{M} = \cup m_i$ is connected (4)

Observe the similarities and differences between (3) and (4). The manifold $\mathfrak{M}$ in (4) is called the medial axis of Ω, while $R(p)$ is the radius function. Note that the medial axis $\mathfrak{M}$, in general, will not coincide with the mid-surface ω The vectors $\hat{k}_i^\pm$ in (4) are defined to be the two projectors from a point on the medial axis to the nearest boundary points; Figure 3 illustrates this definition. The mathematical properties of skeletal representations are now well understood [7], while its computation has been addressed by Srinivasan [8] and Etzion [9]; Figure 4 illustrates the skeletal decomposition of the solid in Figure 2, where the darker shade sub-domains correspond to Ω_i^+, while lighter ones correspond to Ω_i^-.

The main contribution of this paper is to propose a (Galerkin) sub-space of solutions that are explicit functions of the thickness parameter ε in (4). In this paper, we shall restrict ourselves to polynomial sub-spaces of ε. Note that such sub-spaces may be expanded to include, for example, singularity functions at re-entrant corners. Once a sub-space is selected, then a dimensionally reduced eigen-value problem over $\mathfrak{M}$ may be obtained via standard symbolic integration [5]. Since skeletal reduction rests on sub-space projections, it inherits the

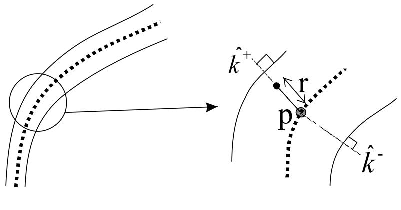


Figure 3. Example to illustrate (4).

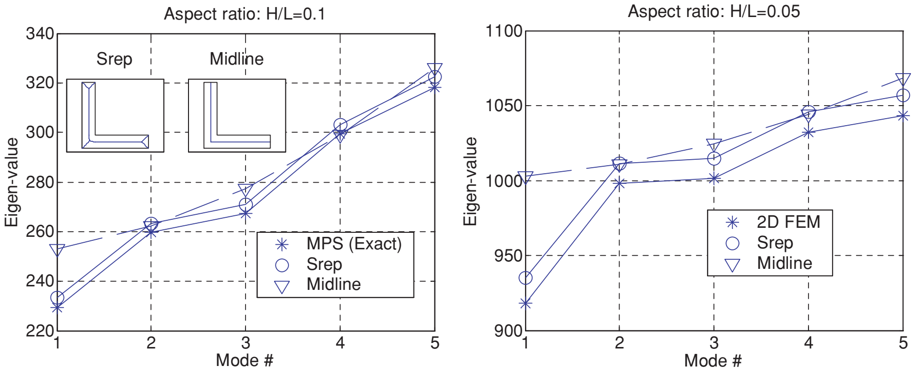


Figure 4. Skeletal decomposition of L-bracket.

standard properties of a Galerkin approximation [1]. These and other theoretical properties of the proposed method will be discussed in a forthcoming paper.

3. NUMERICAL EXPERIMENT

For lack of space, we consider only two numerical experiments involving the Helmholtz eigen-value problem (1) when Ω is the 2-D L-bracket illustrated in Figure 2. The experiments involve comparing the eigen-values λ_{mid} obtained via the mid-element based reduction (3), when conditions (3)–(1) and (3)–(2) are not satisfied, with the eigen-values λ_{skel} obtained via the skeletal reduction (4), when all four conditions were satisfied. In the first experiment below with $H/L = 0.1$, the solution λ_{mps} obtained via the method of particular solutions [10] is used as a reference. In the second experiment ($H/L = 0.05$), the solution λ_{fem} is obtained via a fully discretized finite element solution, and is used as a reference. (We are unable to compare the eigen-vectors since the eigen-vectors u_{mid} are not well-defined, when some of the conditions in (3) are violated; the eigen-vectors u_{skel} are however well-defined.)

Observe in Figure 5 that the first (critical) eigen-value λ_{mid} departs considerably from the expected value while the eigen-values λ_{skel} are considerably

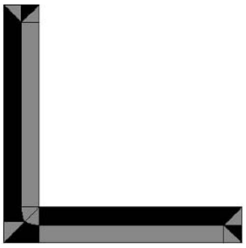


Figure 5. Comparison of eigen-values for the L-bracket for two aspect ratios.

more accurate (with a loss in accuracy inherent to dimensional reduction). The computational costs for mid-element and skeletal reduction are almost identical (and significantly lower than either of the two 2-D methods). Further, skeletal reduction can be easily automated making it a practical alternate to mid-element based reduction.

REFERENCES

1. M. Ainsworth and M. Arnold (2001), Computable error bounds for some simple dimensionally reduced models on thin domains. *IMA Journal of Numerical Analysis*, 21, pp. 81–105.
2. M. Vogelius and I. Babuska (1981), On a dimensional reduction method. I. The optimal selection of basis function. *Mathematics of Computation*, 37, 155, pp. 31–46.
3. P.G. Ciarlet (1990), *Plates and Junctions in Elastic Multistructures*. Springer-Verlag, New York.
4. V.V. Vasiliev (2000), Modern conceptions of plate theory. *Composite Structures*, 48, pp. 39–48.
5. K. Suresh (2003), Generalization of the mid-element dimensional reduction. *Journal of Computing and Information Science in Engineering*, 3, pp. 308–314.
6. H. Blum and R.N. Nagel (1978), Shape description using weighted symmetric axis features. *Pattern Recognition*, 10, pp. 167–180.
7. E.C. Sherbrooke and N.M. Patrikalakis (1996), Differential and topological properties of medial axis transforms. *Graphical Models and Image Processing*, 58, 6, pp. 574–592.
8. V. Srinivasan and L.R. Nackman (1987), Voronoi diagram of multiply connected polygonal domains I: Algorithm. *IBM Journal of Research*, 31, 3, pp. 1–31.
9. M. Etzion and A. Rappoport (2002), Computing voronoi skeletons of a 3-D polyhedron by space subdivision. *Computational Geometry*, 21, pp. 87–120.
10. T. Betcke and L.N. Trefethen (2003), Reviving the method of particular solutions. *Technical Report NA-03/12*, Computing Laboratory, Oxford University, Oxford, UK.

EXACT SOLUTION FOR THE FREE VIBRATION OF A TAPERED BEAM WITH ELASTIC END ROTATIONAL RESTRAINTS

Reza Attarnejad,[1] Negin Manavi[2] and Alireza Farsad[3]
[1] *Civil Engineering Department, University of Tehran*
[2] *Student of M.S. Hydraulically structures, Civil Engineering Department, University of Tehran*
[3] *Graduated Student of M.S. Structural Engineering, Civil Engineering Department, University of Tehran*

Abstract An exact solution for the free vibration of a tapered beam with elastic end rotational restraints problem is obtained using the Euler–Bernoulli theory of beam vibrations by programming them in FORTRAN 90. Natural frequencies and mode shape details of an Euler–Bernoulli beam with ends on elastically support are derived. Free vibration of uniform and tapered beam with elastic restraint is studied using free vibration equation and the well-known Bessel functions. In this way exact results can be obtained for the entire cross sectional laws which allow one to reduce the equation of motion to a couple second-order Bessel equations. The resulting displacement is then expressed in terms of Bessel functions. Natural frequencies and mode shapes are determined using governed boundary conditions (spring-free) and obtaining the matrix of coefficient. By equaling the matrix determinant to zero and finding the roots actually natural frequencies are derived. The paper ends with numerical examples, which confirm the usefulness of the proposed method, and in good agreement with some previously known results from journal papers and finite element methods. The accuracy of the method considering simplicity of it is very precious.

Keywords: Euler–Bernoulli beam, free vibration, natural frequency, mode shape, dynamic response, tapered beam, exact solution, elastic end.

1. INTRODUCTION

The importance of this subject is obvious considering great use of exact solution for structural elements. For many structural elements in actual use

G. R. Liu et al. (eds.), Computational Methods, 1993–2003.

it proves to be convenient to study the dynamic behaviour, and the quantities that have an influence on the results by schematizing them: this often leads to the variable cross-section beam model. The results thus obtained supply, by simple analogy, useful indications regarding the influence of the different parameters for complex problem whose solution proves to be difficult to obtain. This simplification, therefore, allows optimization of the structures of mobile arms used in cybernetics, towers, solar panel frames, tall building etc.

Many researchers have studied free vibration characteristics of tapered beams with linear end rotational restraints. Dubil, Conway (1965) gave some preliminary results for more common boundary conditions. Goel (1976), Auciello (1995), Auciello and De Rosa (1996), Auciello and Ercolano (1997) proposed an interesting generalization to the non-classical boundary conditions. Liz Graciela Nallim, Ricardo Oscar Grossi (1999) proposed a general algorithm for the study of the dynamical behaviour of beams. Auciello (2000) gave a paper about free vibration of a restrained shear-deformable tapered beam with a tip mass at its free end. Naidu and Rao and Raju (2001) considered free vibration behavior of tapered beam with non-linear elastic end rotational restraints. Auciello (2001) worked on the transverse vibrations of non-uniform beams with axial loads and elastically restrained ends. Naguleswaran (2002) formulated vibration of a Euller–Bernoulli beam on elastic end supports and with up to three step changes in cross-section. Naguleswaran (2003) also formulated vibration and stability of an Euler–Bernoulli beam on elastic end supports and with up to three step changes in cross-section and in axial force. Naguleswaran (2004) also formulated vibration of an Euler–Bernoulli stepped beam carrying a non-symmetrical rigid body at the step. And many other researchers whose results use in this paper and their names are in references [1–16].

2. EQUATION OF MOTION

The equation of motion of an Euler–Bernoulli beam.

$$\frac{\partial^2}{\partial x^2}\left(EI(x)\frac{\partial^2 w}{\partial x^2}\right)+m(x)\frac{\partial^2 w}{\partial t^2}=P_z \tag{1a}$$

$$\theta(x)=-\frac{\partial w}{\partial x} \tag{1b}$$

$$M(x)=-EI(x)\frac{\partial^2 w}{\partial x^2} \tag{1c}$$

$$V(x)=-\frac{\partial}{\partial x}\left(EI(x)\frac{\partial^2 w}{\partial x^2}\right) \tag{1d}$$

Where

$EI(x) =$ bending stiffness of beam
$m(x) =$ mass per unit length of beam
$P_z =$ Vertical load per unit length
$w =$ Displacement of beam
$\theta =$ Rotation of beam
$M =$ Bending moment of beam
$V =$ Shear of beam
t, x express time and axis of beam.

Considering condition of problem, above equation can be solved and used in analysis of a beam.

The equation of motion of a tapered Euler–Bernoulli beam:

$$\frac{\partial^2}{\partial x^2}(EI)(x)\frac{\partial^2 V}{\partial x^2} + \rho A(x)\frac{\partial^2 V}{\partial t^2} = 0 \tag{2}$$

The moment of Inertia and cross-section area express as follow:

$$I(x) = I_o\xi^{N+2}$$
$$A(x) = A_0\xi^{N} \tag{3}$$
$$\xi = 1 + C\frac{x}{a} \tag{4}$$

It is assumed that:

$$V(\xi, t) = V(\xi)\sin(\omega t) \tag{5}$$

Substituting Equation (3, 4, 5) into Equation (2), equation can be obtained.

$$\xi^2\frac{d^4 V}{d\xi^4} + 2(N+2)\xi\frac{d^3 V}{d\xi^3} + (N+2)(N+1)\frac{d^2 V}{d\xi^2} - \frac{\lambda_g^4}{C^4}V = 0 \tag{6}$$

$$\lambda_g = L\sqrt[4]{\frac{\rho A_g\omega^2}{EI_g}} \tag{7}$$

Solution for Equation (6) is:

$$V(\varphi) = \frac{1}{\varphi^N}\{C_1 J_N(\varphi) + C_2 Y_N(\varphi) + C_2 Y_N(\varphi) + C_3 I_N(\varphi) + C_4 K_N(\varphi)\} \tag{8}$$

$$\varphi = \frac{2\lambda_g}{C}\sqrt{\xi} \tag{9}$$

And finally following equations express: φ rotation, M moment, and s shear in each node.

$$\theta(\varphi) = -\frac{\lambda_g}{\varphi^N}\{C_1 J_{N+1}(\varphi) + C_2 Y_{N+1}(\varphi) - C_3 I_{N+1}(\varphi) + C_4 K_{N+1}(\varphi)\} \quad (10)$$

$$M(\varphi) = \frac{EI_g \xi^{N+2}\lambda_g^2}{h^2\varphi^N}\{C_1 J_{N+2}(\varphi) + C_2 Y_{N+2}(\varphi) + C_3 I_{N+2}(\varphi) + C_4 K_{N+2}(\varphi)\} \quad (11)$$

$$S(\varphi) = \frac{EI_g \xi^{N+\frac{1}{2}}\lambda_g^3}{h^3\varphi^N}\{C_1 J_{N+1}(\varphi) + C_2 Y_{N+1}(\varphi) + C_3 I_{N+1}(\varphi) + C_4 K_{N+1}(\varphi)\} \quad (12)$$

K, I, Y, J are first-order and second-order Bessel functions and modified first-order and second-order Bessel functions [19].

3. BOUNDARY CONDITION

3.1 Pined-Pined

$$\begin{array}{ll} V(0) = 0 & V(L) = 0 \\ M(0) = 0 & M(L) = 0 \end{array} \quad (13)$$

$$A = \begin{bmatrix} J_N(\varphi_0) & Y_N(\varphi_0) & I_N(\varphi_0) & K_N(\varphi_0) \\ J_{N+2}(\varphi_0) & Y_{N+2}(\varphi_0) & I_{N+2}(\varphi_0) & K_{N+2}(\varphi) \\ J_N(\varphi_L) & Y_N(\varphi_L) & I_N(\varphi_L) & K_N(\varphi_L) \\ J_{N+2}(\varphi_L) & Y_{N+2}(\varphi_L) & I_{N+2}(\varphi_L) & K_{N+2}(\varphi_L) \end{bmatrix} \quad (14)$$

$$\varphi_0 = \frac{2\lambda_g}{C}, \varphi L = \varphi_0\sqrt{1+C} \quad (15)$$

3.2 Clamped-Clamped

$$\begin{array}{ll} V(0) = 0 & V(L) = 0 \\ \theta(0) = 0 & \theta(L) = 0 \end{array} \quad (16)$$

$$A = \begin{bmatrix} J_N(\varphi_0) & Y_N(\varphi_0) & I_N(\varphi_0) & K_N(\varphi_0) \\ J_{N+1}(\varphi_0) & Y_{N+1}(\varphi_0) & -I_{N+1}(\varphi_0) & K_{N+1}(\varphi) \\ J_N(\varphi_L) & Y_N(\varphi_L) & I_N(\varphi_L) & K_N(\varphi_L) \\ J_{N+1}(\varphi_L) & Y_{N+1}(\varphi_L) & -I_{N+1}(\varphi_L) & K_{N+1}(\varphi_L) \end{bmatrix} \quad (17)$$

3.3 Clamped-Free (Cantilever)

$$
\begin{array}{ll}
V(0)=0 & M(L)=0 \\
\theta(0)=0 & S(L)=0
\end{array}
\tag{18}
$$

$$
A=\begin{bmatrix}
J_N(\varphi_0) & Y_N(\varphi_0) & I_N(\varphi_0) & K_N(\varphi_0) \\
J_{N+1}(\varphi_0) & Y_{N+1}(\varphi_0) & -I_{N+1}(\varphi_0) & K_{N+1}(\varphi) \\
J_{N+2}(\varphi_L) & Y_{N+2}(\varphi_L) & I_{N+2}(\varphi_L) & K_{N+2}(\varphi_L) \\
J_{N+1}(\varphi_L) & Y_{N+1}(\varphi_L) & I_{N+1}(\varphi_L) & -K_{N+1}(\varphi_L)
\end{bmatrix}
\tag{19}
$$

3.4 Spring-Free

$$
\begin{array}{ll}
S(0)=K_u\times V(0) & S(L)=0 \\
M(0)=K_\theta\times\theta(0) & M(L)=0
\end{array}
\tag{20}
$$

$$
A=\left[\begin{array}{cc}
\dfrac{EI_g\xi^{N+1}\lambda_g^2}{L^2\varphi_0^N}J_{N+2}(\varphi_0)+\dfrac{K_\theta\lambda_g}{\varphi_0^N}J_{N+1}(\varphi_0) & \dfrac{EI_g\xi^{N+1}\lambda_g^2}{L^2\varphi_0^N}Y_{N+2}(\varphi_0)+\dfrac{K_\theta\lambda_g}{\varphi_0^N}Y_{N+1}(\varphi_0) \\
\dfrac{EI_g\xi^{N+\frac{1}{2}}\lambda_g^3}{L^3\varphi_0^N}J_{N+1}(\varphi_0)-\dfrac{K_U}{\varphi_0^N}J_N(\varphi_0) & \dfrac{EI_g\xi^{N+\frac{1}{2}}\lambda_g^3}{L^3\varphi_0^N}Y_{N+1}(\varphi_0)-\dfrac{K_U}{\varphi_0^N}Y_N(\varphi_0) \\
J_{N+2}(\varphi_L) & Y_{N+2}(\varphi_L) \\
J_{N+1}(\varphi_L) & Y_{N+1}(\varphi_L)
\end{array}\right.
$$

$$
\left.\begin{array}{cc}
\dfrac{EI_g\xi^{N+1}\lambda_g^2}{L^2\varphi_0^N}I_{N+2}(\varphi_0)-\dfrac{K_\theta\lambda_g}{\varphi_0^N}I_{N+1}(\varphi_0) & \dfrac{EI_g\xi^{N+1}\lambda_g^2}{L^2\varphi_0^N}K_{N+2}(\varphi_0)+\dfrac{K_\theta\lambda_g}{\varphi_0^N}K_{N+1}(\varphi_0) \\
\dfrac{EI_g\xi^{N+\frac{1}{2}}\lambda_g^3}{L^3\varphi_0^N}I_{N+1}(\varphi_0)-\dfrac{K_U}{\varphi_0^N}I_N(\varphi_0) & -\dfrac{EI_g\xi^{N+\frac{1}{2}}\lambda_g^3}{L^3\varphi_0^N}K_{N+1}(\varphi_0)-\dfrac{K_U}{\varphi_0^N}K_N(\varphi_0) \\
I_{N+2}(\varphi_L) & K_{N+2}(\varphi_L) \\
I_{N+1}(\varphi_L) & -K_{N+1}(\varphi_L)
\end{array}\right]
\tag{21}
$$

For the other boundary conditions considering above equations formulation can be derived.

When C has small value for example $C \leq 0.01$ the beam physically closes to uniform beam and *K, L* (modified first-and second-order Bessel function) overflows. Considering that when their parameter is large they incline to simple mathematical functions, so this problem is dissolvable. For frequency computing (determinant equal to zero) these two functions multiply by each other by determinant of a 4×4 or $4(n+1) \times 4(n+1)$ *matrix* and also it is like this for mode shapes computing.

When their parameter is large:

$$\begin{aligned}&\alpha \to \infty\\ &I_m(\alpha) \overset{\sim}{=} e^{\alpha} P_1(\alpha)\\ &K_n(\alpha) \overset{\sim}{=} e^{-\alpha} P_2(\alpha)\end{aligned} \tag{22}$$

P_1, P_2 is known and limited functions, which, defined considering conditions. So the answer of $I_m(\alpha) \times K_m(\alpha)$ will be limited.

If the parameters differ from each other when they both are large:

$$I_m(\alpha)K_n(\beta) = e^{\alpha-\beta} P_1(\alpha)P_2(\beta) \tag{23}$$

Therefore, if C has a small value or ω has large value, the value of α, β will be large, but they are a little bit different from each other so they never overflow Generally when the parameter is large:

$$\begin{aligned}P_1(\alpha) &= \frac{1}{\sqrt{2\pi\alpha}}\\ P_2(\alpha) &= \sqrt{\frac{\pi}{2\alpha}}\end{aligned}$$

4. MODE SHAPES AND FREQUENCIES OF BEAM

Each beam has 2 boundary conditions in each edge; therefore, there are 4 conditions (equations) in each beam. If conditions express as a matrix and the equation was related to free vibration so:

$$[A]_{4\times 4} \begin{Bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{Bmatrix} = \{0\} \tag{24}$$

Solving Equation (29) results: C_1, C_2, $C_3 C_4$ and natural frequencies of beam are values, which make the determinant of [A] zero.

$$A = \begin{bmatrix} m_1 & m_2 & m_3 & m_4 \\ n_1 & n_2 & n_3 & n_4 \\ p_1 & p_2 & p_3 & p_4 \\ q_1 & q_2 & q_3 & q_4 \end{bmatrix} \tag{25}$$

Forming equation for 3 first lines:

$$\begin{aligned} m_1C_1 + m_2C_2 + m_3C_3 + m_4C_4 &= 0 \\ n_1C_1 + n_2C_2 + n_3C_3 + n_4C_4 &= 0 \\ p_1C_1 + p_2C_2 + p_3C_3 + p_4C_4 &= 0 \end{aligned} \tag{26}$$

Dividing them by C_4:

$$C'_1 = \frac{C_1}{C_4}, \quad C'_2 = \frac{C_2}{C_4}, \quad C'_3 = \frac{C_3}{C_4} \tag{27}$$

$$\begin{aligned} m_1C'_1 + m_2C'_2 + m_3C'_3 + m_4 &= 0 \\ n_1C'_1 + n_2C'_2 + n_3C'_3 + n_4 &= 0 \\ p_1C'_1 + p_2C'_2 + p_3C'_3 + p_4 &= 0 \end{aligned} \tag{28}$$

$$\Rightarrow \begin{bmatrix} m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \\ p_1 & p_2 & p_3 \end{bmatrix} \begin{Bmatrix} C'_1 \\ C'_2 \\ C'_3 \end{Bmatrix} = - \begin{Bmatrix} m_4 \\ n_4 \\ p_4 \end{Bmatrix} \tag{29}$$

$$\Rightarrow \begin{Bmatrix} C'_1 \\ C'_2 \\ C'_3 \end{Bmatrix} = - \begin{bmatrix} m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \\ p_1 & p_2 & p_3 \end{bmatrix}^{-1} \begin{Bmatrix} m_4 \\ n_4 \\ p_4 \end{Bmatrix} \tag{30}$$

Coefficients of shape mode obtain from $C_1^{''}, C_2^{''}, C_3^{''}$.

$$U(\varphi) = \frac{1}{\varphi^n}\{C_1^{'}J_n + C_2^{''}Y_n + C_3^{''}I_n + K_n\}C_4 \tag{31}$$

Computing of derivatives considering constancy of unknowns is available easily.

If the beam is a stepped beam four boundary conditions are added for connectivity of displacement, rotation, moment, shear at stepped cross section. Equations of motion of beam express in the form of $n + 1$ differential equations, which have answers that are corrected for stepped cross section. Generally for n-step, matrix of coefficient will be $4(n + 1) \times 4(n + 1)$ and unknowns will be $4(n + 1)$.

If there is just one-step, matrix of coefficients will be 8×8 and unknown coefficients will be C_1 to C_8. If answers in area between $0, b$ express as U_1, θ_1, M_1, S_1 and in area between b, h express as U_2, θ_2, M_2, S_2, boundary conditions will be:

$$\begin{aligned} &V_1(0) = 0 && V_1(b) = V_2(b) \\ &\theta_1(0) = 0 && \theta_1(b) = \theta_2(b) \\ &M_2(h) = 0 && M_1(b) = M_2(b) \\ &S_2(h) = 0 && S_1(b) = S_2(b) \end{aligned} \tag{32}$$

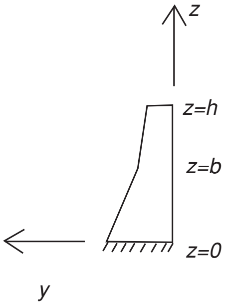


Figure 1. Beam with one-step.

5. NUMERICAL EXAMPLES

Written equations and relations were used for analysis of some examples and results were compared with previous researches and softwares. Here there are some tables and figures of compared results of present method and some journal papers, which show high accuracy of, proposed method.

$$\varphi_0 = \frac{2\lambda_g}{C}, \quad \varphi_L = \varphi_0\sqrt{1+C}, \quad \lambda_g = L_4\sqrt{\frac{\rho A_g \omega^2}{E I_g}}$$

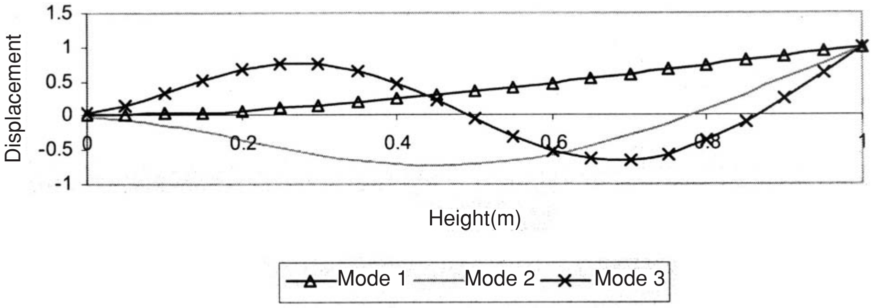


Figure 2. Comparison of mode shape 1, 2, 3 of spring-free beam with $K_\theta = K_U = 4.5 \times 10^8$.

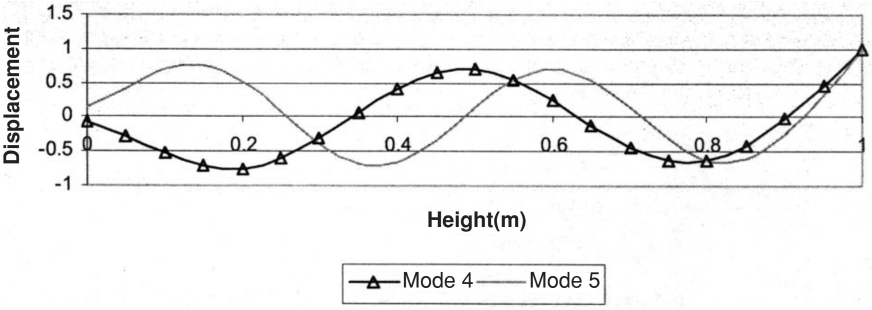


Figure 3. Comparison of mode shape 4, 5 of spring-free beam with $K_\theta = K_U = 4.5 \times 10^8$.

Table 1. Comparison of non-dimensional natural frequencies of present method and some journal papers (cantilever beam).

	λ_1	λ_2	λ_3	λ_4	λ_5
$\frac{h_0}{h_L} = 2$					
Naguleswaran (1994)	3.8238	18.317	47.265	–	–
Lee (1992)	3.8238	18.3172	47.2648	–	–
Ho & Chen (1998)	3.8238	18.3173	47.2648	–	–
Present Method	3.8238	18.3173	47.2648	90.4505	148.0016
$\frac{h_0}{h_L} = 5$					
Naguleswaran (1994)	4.2925	15.743	36.885	–	–
Present Method	4.2925	15.743	36.885	68.116	109.594

Table 2. Comparison of natural frequencies of present method and some journal papers (cantilever beam, $h_L = 15\,m,\ L = 180\,m,\ E = 3.5 \times 10^7 \frac{ton}{m^2}$, $m = 3.7\frac{ton}{m}$.

	ω_1	ω_2	ω_3	ω_4	ω_5
$\frac{h_0}{h_L} = 2$					
Naguleswaran (1994)	6.1124	29.279	75.554	–	–
Lee (1992)	6.1124	29.281	75.554	–	–
Ho & Chen (1998)	6.1124	29.282	75.554	–	–
Present Method	6.1124	29.2805	75.5538	144.587	236.584
$\frac{h_0}{h_L} = 5$					
Naguleswaran (1994)	6.8616	25.165	58.961	–	–
Present Method	6.8616	25.165	58.961	108.886	175.188

Table 3. Result of natural frequency of present method (Spring-free beam, $h_L = 15\,m, L = 180\,m, E = 3.5 \times 10^7 \frac{ton}{m^2}, m = 3.7\frac{ton}{m}$.

	$K_\theta = K_U$	ω_1	ω_2	ω_3	ω_4	ω_5
$\frac{h_0}{h_L} = 2$						
Present method	*1.0E + 20*	6.1124	29.2805	75.5538	144.587	236.584
Present method	*1.0E + 10*	6.1124	29.2803	75.5528	144.583	236.574
Present method	*1.0E + 07*	6.0756	29.073	74.4917	140.63	224.906
Present method	*1.0E + 06*	5.7608	27.2751	64.3021	108.626	173.757
Present method	*1.0E + 05*	3.7988	18.1595	40.3738	86.5519	158.183
$\frac{h_0}{h_L} = 5$						
Present method	*1.0E + 20*	6.8616	25.165	58.961	108.886	175.188
Present method	*1.0E + 10*	6.8616	25.165	58.960	108.884	175.183
Present method	*1.0E + 07*	6.8147	25.003	58.384	107.055	170.218
Present method	*1.0E + 06*	6.4137	25.633	52.963	89.068	134.570
Present method	*1.0E + 05*	4.0038	17.013	34.398	66.080	116.642

Table 4. Non-dimensional natural frequencies results of present method for different spring stiffness of beam.

	$h_0 = 15m, h_L = 13m$						
	K_θ	K_U	λ_1	λ_2	λ_3	λ_4	λ_5
Clamped-free (uniform)	Infinity	Infinity	5.6204	35.223	98.624	193.264	319.480
Clamped-free	Infinity	Infinity	5.7137	33.731	92.800	181.007	298.641
Spring-free	4.50E + 08	Infinity	5.7135	33.730	92.799	181.006	298.639
Spring-free	4.50E + 07	Infinity	5.7112	33.724	92.790	180.993	298.623
Spring-free	4.50E + 06	Infinity	5.6886	33.668	92.696	180.862	298.458
Spring-free	4.50E + 05	Infinity	5.4719	33.130	91.796	179.614	296.881
Spring-free	4.50E + 04	Infinity	3.9638	29.606	85.827	171.280	286.260
Spring-free	4.50E + 08	4.50E + 08	5.7134	33.725	92.761	180.860	298.238
Spring-free	4.50E + 07	4.50E + 07	5.7104	33.674	92.408	179.493	294.352
Spring-free	4.50E + 06	4.50E + 06	5.6811	33.162	88.621	163.193	249.633
Spring-free	4.50E + 05	4.50E + 05	5.4026	28.298	61.848	118.296	210.163
Spring-free	4.50E + 04	4.50E + 04	3.6543	14.543	44.853	107.962	201.431

Table 5. Non-dimensional natural frequencies of present method for different spring stiffness of beam.

	$h_0 = 15m, h_L = 13m$						
	K_θ	K_U	λ_1	λ_2	λ_3	λ_4	λ_5
Clamped-free	Infinity	Infinity	11.4361	41.942	98.268	181.476	291.980
Spring-free	4.50E + 08	Infinity	11.4285	41.925	98.240	181.438	291.930
Spring-free	4.50E + 07	Infinity	11.3603	41.773	97.994	181.095	291.491
Spring-free	4.50E + 06	Infinity	10.7125	40.381	95.763	178.011	287.549
Spring-free	4.50E + 05	Infinity	6.7589	33.819	85.845	164.683	270.777
Spring-free	4.50E + 04	Infinity	1.8644	29.322	79.378	156.189	260.220
Spring-free	4.50E + 08	4.50E + 08	11.4280	41.914	98.170	181.177	291.219
Spring-free	4.50E + 07	4.50E + 07	11.3556	41.664	97.278	178.333	283.431
Spring-free	4.50E + 06	4.50E + 06	10.6691	39.319	87.992	147.723	223.589
Spring-free	4.50E + 05	4.50E + 05	6.5964	28.185	56.999	109.847	194.123
Spring-free	4.50E + 04	4.50E + 04	1.8229	14.096	39.012	94.803	178.376

6. CONCLUSION

An exact solution for the free vibration of a tapered beam with elastic end rotational restraints problem is obtained using the Euler–Bernoulli theory of beam vibrations. Natural frequencies and mode shape details of an Euler–Bernoulli beam with ends on elastically support were derived. Numerical examples were presented which, confirmed the usefulness of proposed method. Comparison of results and those of various research works shows high accuracy of method and good performance of it was demonstrated.

REFERENCES

1. J.R. Banerjee and F.W. Williams (1985), 'Exact Bernoulli–Euler dynamic stiffness matrix for a range of tapered beams'. *International Journal for Numerical Methods in Engineering*, 21, pp. 2289–2302.
2. S.Y. Lee, S.M. Lin and C.T. Wu (2004), 'Free vibration of rotating non-uniform beam, an elastically restrained root and a tip mass'. *Journal of Sound and Vibration*, 273, pp. 477–492.
3. N.M. Auciello (2001), 'On the transverse vibrations of non-uniform beams with axial loads and elastically restrained ends'. *International Journal of Mechanical Sciences*, 43, pp. 193–208.
4. N.M. Auciello and M.A. De Rosa (2004), 'Two approaches to the dynamic analysis of foundation beams subjected to sub tangential forces'. *Computers and Structures*, 82, pp. 519–524.
5. Raju Nadiu (2001), 'Free vibration behavior of tapered beams with non-linear elastic end rotational restraints'. *Journal of Sound and Vibration*, 240, 1, pp. 195–202.
6. S.Y. Lee, H.Y. Ke and Y.H. Kuo (1992), 'Exact solutions for the analysis of general elastically restrained uniform beams'. *International Journal of Applied Mechanical*, 59 pp. 205–212.
7. S.H., Ho and C.K. Chen (1998), 'Analysis of general elastically end restrained non-uniform beams differential transform'. *Applied Mathematical Modeling*, 22, pp. 219–234.
8. S. Naguleswaran (1994), 'A direct solution for the vibration of Euler–Bernoulli wedge and cone beams'. *International Journal of Sound and Vibration*, 172, pp. 289–304.
9. S. Naguleswaran (2002), 'Natural frequencies, sensitivity and mode shape details of an Euler–Bernoulli beam with one-step change in cross section and with end on classical support'. *Journal of Sound and Vibration*, 252, 4, pp. 751–767.
10. S. Naguleswaran (2002), 'Vibration of an Euler–Bernoulli beam on elastic end supports and with up to three step changes in cross-section'. *International Journal of Mechanical Sciences*, 44, pp. 2541–2555.
11. S. Naguleswaran (2003), 'Vibration and stability of an Euler–Bernoulli beam with up to three-step changes in cross-section and in axial force'. *International Journal of Mechanical Sciences*, 45, pp. 1563–1579.
12. S. Naguleswaran (2004), 'Vibration of an Euler–Bernoulli stepped beam carrying a non-symmetrical rigid body at the step'. *Journal of Sound and Vibration*, 271, pp. 1121–1132.
13. Alireza Farsad (2003), 'Analysis of Variable Thickness Plates and Shells', M.S. Thesis, Faculty of Engineering, University of Tehran, Iran.