Ligang Yao · Shuncong Zhong
Hisao Kikuta · Jih-Gau Juang
Masakazu Anpo *Editors*

Advanced Mechanical Science and Technology for the Industrial Revolution 4.0

Springer

Advanced Mechanical Science and Technology for the Industrial Revolution 4.0

Ligang Yao · Shuncong Zhong
Hisao Kikuta · Jih-Gau Juang
Masakazu Anpo
Editors

Advanced Mechanical Science and Technology for the Industrial Revolution 4.0

Springer

Editors
Ligang Yao
Fuzhou University
Fuzhou, Fujian
P. R. China

Shuncong Zhong
Fuzhou University
Fuzhou, Fujian
P. R. China

and

Shanghai University
Shanghai
P. R. China

Hisao Kikuta
Osaka Prefecture University
Sakai, Osaka
Japan

Jih-Gau Juang
National Taiwan Ocean University
Keelung
Taiwan

Masakazu Anpo
State Key Laboratory of Photocatalysis on Energy and Environment, Fuzhou University
Fuzhou, Fujian
P. R. China

and

Osaka Prefecture University
Sakai, Osaka
Japan

ISBN 978-981-10-4108-2 ISBN 978-981-10-4109-9 (eBook)
https://doi.org/10.1007/978-981-10-4109-9

Library of Congress Control Number: 2017952926

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Foreword

The 1st Joint Symposium on Advanced Science and Technology for the Industrial Revolution 4.0 was held at Fuzhou University in China on December 1–3, 2016, under cooperation and collaboration among Fuzhou University (FZU), Osaka Prefecture University (OPU), and National Taiwan Ocean University (NTOU). In the scientific program, we had four excellent invited keynote lectures by Prof. Y.B. Kim (Korea), Prof. N. Sugimura (Japan), Prof. J.Y. Yen (Taiwan), and Prof. Y.L. Tu (Canada) as well as the 17 oral presentations and 23 poster presentations from these three universities. We had more than 250 participants, and many active discussions and exchange of ideas during the conference.

The idea of this joint symposium was proposed by Prof. Ligang Yao, the Dean of the School of Mechanical Engineering and Automation of FZU at the end of 2015. It was profoundly impressive to start the 1st joint symposium after various discussions and exchanges among the three universities since the beginning of 2016. At the opening ceremony, Prof. Zhigang Huang, Vice President of FZU, provided an impressive opening address on behalf of the President of FZU, Prof. Xianzhi Fu, to introduce some characteristic features of FZU and the School of Mechanical Engineering and Automation. This is followed by cochairs of the joint symposium, Prof. Hisao Kikuta, the Head of the Department of Mechanical Engineering of OPU and Prof. Jih-Gau Juang, Vice President of NTOU and the Department of Communication, Navigation and Control Engineering that provided warm welcome and informative speech on behalf of the participants.

These proceedings entitled "Advanced Mechanical Science and Technology for the Industrial Revolution 4.0" include 37 articles of which main contents were presented at the 1st joint symposium. The book covers most of important aspects related with the industrial revolution 4.0, with Robotics and Artificial Intelligence in Part I; Sensors and Measurements in Part II; Intelligent Systems in Part III, and Modern Design Methodologies in Part IV which are the most prospective themes in the field of mechanical engineering nowadays.

This book is well edited to include many contributors from China, Japan, Taiwan, Korea, and Canada and provides the current fashion in the research field of mechanical engineering in Asia on being focused on the industrial revolution 4.0 which is the most important topic nowadays in the mechanical engineering field.

Prof. J.S. Dai
April 2017 King's College London, London, UK

Contents

Part I
Robotics and Artificial Intelligence

Simulation and Formulation of Rehabilitation Strategies for a Novel Ankle Rehabilitation Robot

Zhiwei Liao, Zongxing Lu, Chen Peng and Ligang Yao

Abstract A novel ankle rehabilitation robot and its mechanism composition, were proposed. The 3D modeling of the proposed ankle rehabilitation robot and the 3D modeling of the complete human body were established. The integration of the two 3D models was further proceeded by importing the robot models into the AnyBody simulation software. A driving function was also designed on the basis of the sine function; the inverse kinematic analysis and muscle force analysis of the integrated model are carried out. The driving function design was executed to determine the relationships of the movement of the angles for the proposed ankle rehabilitation robot with the activity and passive traction of the muscles. A rehabilitation strategy was also developed according to different motion amplitudes, velocity, and different external environments in dorsal flexion motion. This research would provide a reference for further design of the structure and control of rehabilitation robots.

Keywords Ankle rehabilitation · Biomechanics · Simulation · Sine function · Rehabilitation strategies

1 Introduction

Foot supporting the human body to complete a series of daily activities is an important limb of the human body. The ankle is an essential joint in the foot, and it is the heaviest joint of the body [1]. However, aging deters the recovery ability of the ankle and thus causes joint pain. Consequently, ankle injury affects a person's daily life. Thus, ankle rehabilitation training should be recommended for patients with this injury. For instance, exercise therapy as a rehabilitation training method involves movement tasks to promote the repair of damaged parts in a reasonable working space by using auxiliary sports equipment. The ankle rehabilitation robot has been exten-

Z. Liao · Z. Lu · C. Peng · L. Yao (✉)
School of Mechanical Engineering and Automation, Fuzhou University,
No. 2, Xueyuan Road, Fuzhou 350116, Fujian, P. R. China
e-mail: ylgyao@fzu.edu.cn

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_1

sively investigated [2–8]. Warnica et al. investigated the effects of muscle activation and passive ankle stiffness on the amplitude and frequency of postural sway [9]. They proposed that the active contraction of ankle muscles may affect balance control in quiet standing, but its effects are different from the use of ankle foot orthosis. Naito et al. proposed a method to estimate muscle length parameters based on measured data [10], especially Hill muscle model of muscle and bone structures. An estimation method based on individual muscle skeletal model analysis is also utilized to quantify the muscle length parameters passive joint torque characteristics. Benjamin et al. examined the effects of mechanical assistance on the biomechanical properties of the related muscular system [11]. They focused on system dynamics in a range of biologically active muscle and bone spring stiffness and revealed that mechanical assistant causes a muscle tendon unit to decrease the tune, reduce the required output power, and improve the mechanical power output of muscle active components. Various human motions have been subjected to biomechanical analysis by using a typical musculoskeletal model, in which a Hill-type muscle model is commonly utilized to calculate muscle tension forces [12–14]. Tan et al. built a standard three-dimensional human skeletal muscle model system and its corresponding mechanical model [15]. They established a transformation method between a standard skeletal muscle system and an individual skeletal muscle model to perform the kinematics and dynamics analyses of the typical movement of the human body and to quantify muscle activation associated with participation in exercise. They also developed the software that can predict muscular force of the human body and perform dynamic calculations. Zhang et al. investigated the effect of local muscle fatigue on the body sensation of the ankle joint by analyzing the changes in the position of the ankle joint in the sagittal plane and the changes in the muscle force [16]. They also provided theoretical and experimental bases for the prevention and treatment of ankle joint injury and rehabilitation training. By Detecting the dynamic balance ability of some college students, simultaneously collecting surface EMG signals and evaluating the ankle muscle strength training, Zhao et al. determined the effects of changes in ankle-joint-related muscle activity and muscle strength on homeostasis to understand the mechanism and relationship between ankle joint muscle activity and muscle strength on dynamic balance ability [17].

These researches provided the fundamental basis for the developing of the ankle rehabilitation robot. However, the ankle biomechanics and rehabilitation strategies have not yet to be developed to establish a complete rehabilitation system and this paper focuses on the simulation and formulation of the rehabilitation strategies for a proposed ankle rehabilitation robot.

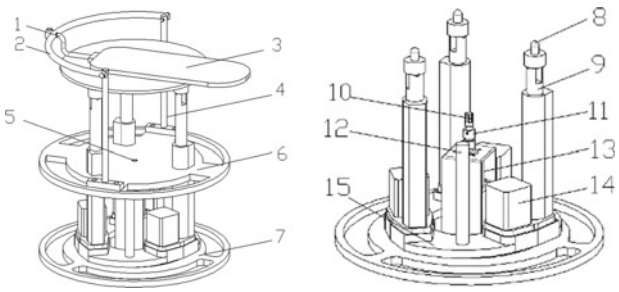
2 Novel Ankle Rehabilitation Robot

The proposed ankle rehabilitation robot is composed of link, rocker, moving platform, two bearing rods, spline groove, rotating platform, base platform, three hemisphere sleeves, three push rods, spline shaft, coupler, three studdles, a motor reducer, and four stepper motors. As shown in Fig. 1.

The proposed robot composed of series and parallel parts, simple structure, small size, light weight, easy to carry and easy to move. It can help patients with ankle plantar or dorsal flexion motion, inversion or eversion motion, and abduction or adduction motion perform ankle rehabilitation exercise. The series part is consisting of a motor reducer, a spline shaft, a rotating platform, bearing rods, a rocker, a link, and a moving platform. A parallel part is consisting of a base platform, three stepper motors, two bearing rods, a rocker, a link and a moving platform. The stepper motors are controlled by inputting the commands into the computer and controlling its motion through the control platform. The three motors which on the support rotating platform push the push rod up and down to drive the moving platform to realize the planter, dorsal flexion motion and inversion, eversion motion, the another one which is fixed on the base platform drives the rotation platform to realize the abduction and adduction motion. The parallel part contains the three motors, so it controls the patients' plantar flexion or dorsal flexion motion and inversion or eversion motion, whereas the series part controls the patients' abduction and adduction motion.

3 Simulation and Rehabilitation Strategy Formulation

The simulation and rehabilitation strategy formulation are the fundamentals for developing the ankle rehabilitation robot and the 3D modeling. In this section, driving function is designed to simulate the rehabilitation motion of ankle and the rehabilitation strategies are formulated. Firstly, the 3D model of human body has been built by using the AnyBody simulation software. The process of modeling mainly includes inputting standard human model to the software, changing model's parameters and retaining the required part of the simulation. Then, the pose of the



1. Link 2. Rocker 3. Moving platform 4. Bearing rod 5. Spline groove 6. Rotating Platform 7. Base platform 8. Hemisphere sleeve 9. Push rod 10. Spline shaft 11. Coupler 12. Studdle 13. Motor reducer 14. Stepper motor 15. Support rotation platform

Fig. 1 Diagram of the ankle rehabilitation robot

model has been adjusted and the ankle rehabilitation robot has been inputted to the software. At this basis, adding constraints make the whole mechanism has only 3 DOFs. Then, by taking Fourier function as driving function, the planter and dorsal flexion motion have been investigated. By inputting the driving function to the AnyBody simulation software, the related muscle' activity and the passive traction of the muscle from different angles have been analyzed. At this basis, the rehabilitation strategy has been formulated. The specific process is as follows (Fig. 2).

3.1 Modeling

A human body model was obtained from the AnyBody model repository (AnyBody Managed Model Repository, AMMR). To increase the running speed, this paper retained the trunk, including the head, neck, pelvis and the right leg, and removed the upper limbs and the left leg, namely setting "TRUNK = 1", "RIGHT_ARM = 0",

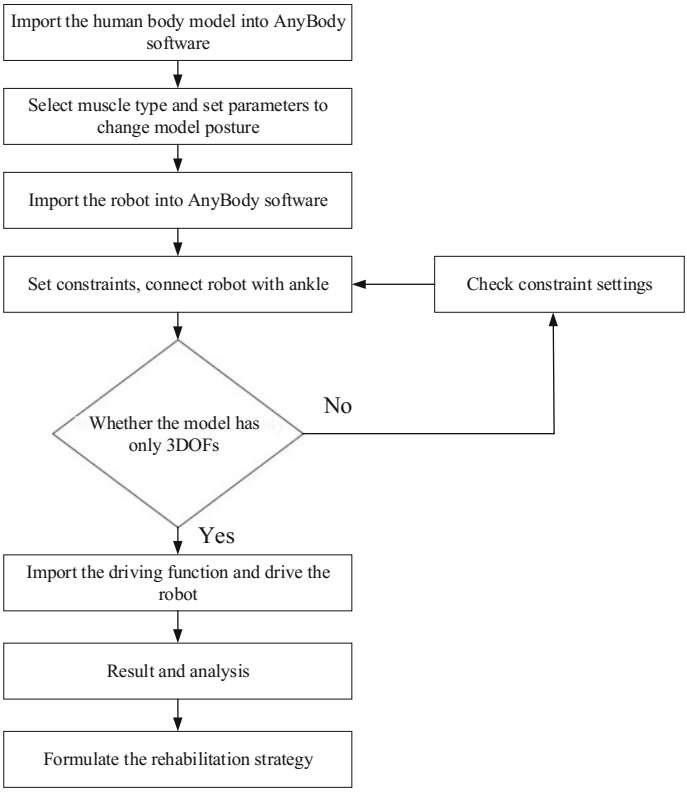


Fig. 2 Process of modeling, simulation and rehabilitation strategy

"LEFT_ARM = 0", "RIGHT_LEG = 1" and "LEFT_LEG = 0". Muscle type selects "AnyMuscleModel3E". The muscle types are the three element muscle models of Hill. The Hill-type muscle model is more similar to the characteristic of real human body. The model is built on the basis of European standards, which recommends a default height of 180 cm and weight of 75 kg. In the actual simulation experiment, we can change the parameters of the human body model depending on different experimental objects and situations (Fig. 3).

The model of the human body connected to the rehabilitation robot was assigned 3 degrees of freedom to simulate the three kinds of motion of the ankle rehabilitation robot. Thus, we added 16 constraints. The following procedures were employed:

(1) Using a fixed joint "AnyStdJoint" to attach the pelvis to a point "Hpoint" in the global coordinate system. The coordinates of the point are set to the default position (0.046, 1.16, 0). In this way, the 6 DOFs are fixed.

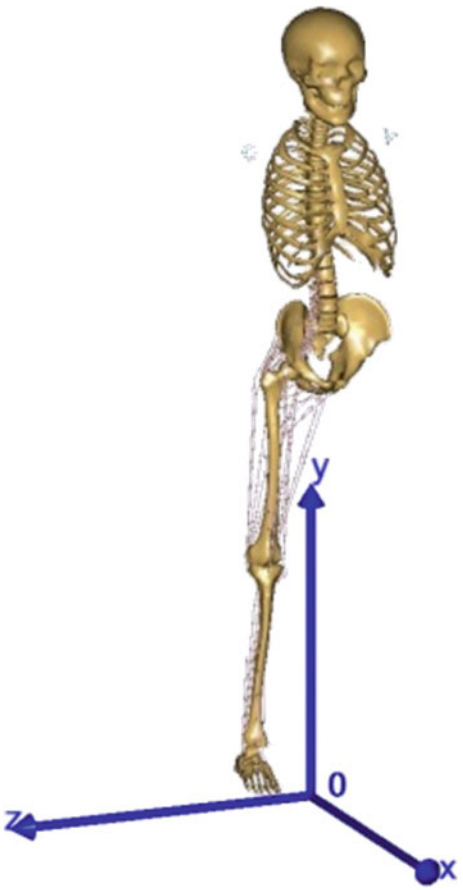


Fig. 3 Human body modeling

(2) Using actuator "AnyKinEqSimpleDriver" to fix the three rotation directions of the trunk (Lateral Bending, Rotation and Extension) and the initial position where the velocity is set to 0. The setting "DriverPos = {0, 0, 0} and the, DriverVel = {0, 0, 0}". In this manner, the 3 DOFs are fixed.
(3) Similar with step (2) and by using actuator "AnyKinEqSimpleDriver" to fix the neck rotation; the initial position and the velocity is set to 0. The setting "DriverPos = {0} and the, DriverVel = {0}". In this manner, 1 DOF are fixed.
(4) Using "AnyStdJoint" to fix the rehabilitation robot pedal which constraints human foot, given that the pedal can drive the ankle to rotate. In this manner, the 6 DOFs are fixed.

After the process was completed, the robot left 3 degrees of freedom to achieve 3 kinds of motion (Fig. 4).

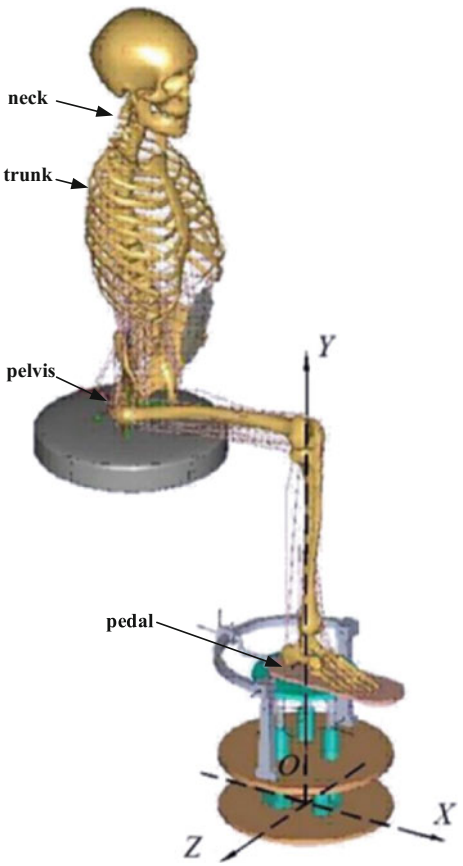


Fig. 4 Ankle rehabilitation robots connected with the human body

3.2 Driven Function Formulation

Because the process of analyzing the three types of motion are similar, the following will focus on the plantar and dorsal flexion motions.

The changing tendency of the ankle joint angle is similar to the sine function. AnyBody provides a drive "AnyKinEqFourierDriver", which can perform a reciprocating stroke. This drive is based on Fourier expansion as the drive function, which can be further expressed as follows.

$$\theta(t) = \sum_{i=1}^{n} A_i \sin(\omega_i t + B_i), i = 1, 2, 3, \ldots, n \tag{1}$$

Where A and B are Fourier coefficients, ω_i is angular velocity, $\omega_i = (i-1)2\pi f$, and f is frequency.

If n = 2,

$$\theta(t) = A_1 \sin B_1 + A_2 \sin(2\pi f + B_2) \tag{2}$$

This function is a sine function of time variation.

According to the direction of motion and the working space of the ankle plantar and dorsal flexion motion, we designed a drive function based on sine function, and the maximum angle of the plantar and dorsal flexion motion are limited in -60° and 30° respectively, where $\theta_1 \in [-\frac{\pi}{3}, \frac{\pi}{6}]$ and T is 1 s:

$$\theta_1(t) = \frac{\pi}{4}\sin(2\pi t + 0.108\pi) - \frac{\pi}{12} \tag{3}$$

Comparing the Fourier expansion (2), we can obtain the Fourier coefficient $A_1 = -\frac{\pi}{12}, A_2 = \frac{\pi}{4}, B_1 = \frac{\pi}{2}, B_2 = 0.108\pi$.

3.3 Formulation of Rehabilitation Strategy

This paper considers the early stage of rehabilitation. In this stage, a patient's ankle moves with the help of an ankle rehabilitation robot. The patient's ankle joint surrounding the ligament and muscle force should be controlled to avoid the recurrence of patient's ankle injury. The Fourier coefficient A_2 and movement period T are the variables. The maximum muscle activity and passive traction F_P are output parameters. The maximum muscle activity is the percentage of the muscle strength, which performs the highest muscle activity. Relaxed muscles correspond to the highest percentage of 100% when the actual value is smaller than the degree of activity.

In summary, a low muscle activity and F_P is beneficial for patients in this rehabilitation stage.

The muscle of the posterior group of the leg is the main part of the force when the ankle performs plantar and dorsal flexion motion. Thus, we mainly consider the maximum muscle activity and the passive traction on the muscle of the posterior group of the leg.

At $n = 2$ and $A_1 = A_2 = A_3 = 0$, we can obtain the following Fourier expansion when the ankle performs dorsal flexion motion (4).

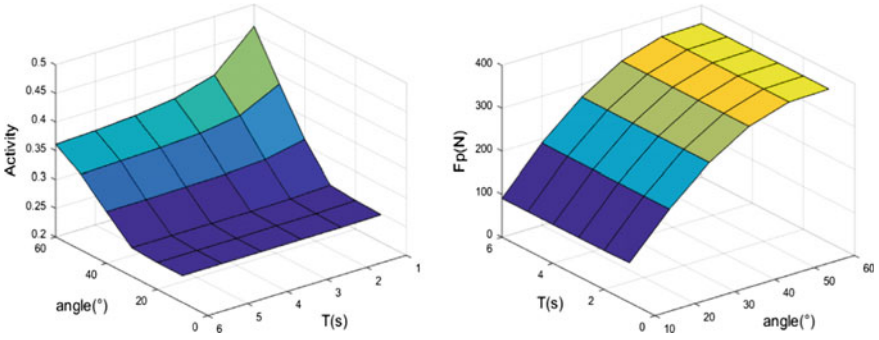
$$\theta_c(t) = A_2 \sin\frac{2\pi t}{T} \tag{4}$$

Importing the drive function into the hinge Z and supposing that the variation range of A_2 is $[10°, 60°]$[17], we determine the interval as 10°, the variation range of T as $[1s, 6s]$, and the interval as 1 s. The maximum muscle activity and F_P are shown in Fig. 5a, b respectively.

In Fig. 5a, the maximum muscle activity is maintained at approximately 25% when dorsal flexion motion occurs from 10° to 30°. The maximum muscle activity improves significantly if the maximum margin continuously increases. The activity is retained at 35–45% when dorsal flexion motion reaches 60°. The muscle activity also increases when the period is shortened. In Fig. 5b, passive traction is also enhanced as the maximum angle is improved. The maximum passive traction gradually slows down and maintains approximately 350 N when dorsal flexion motion reaches 60°. Changes in period slightly affect passive traction.

On the basis of these findings, we assume that the variation range of A_2 is $[25°, 40°]$, the interval is 1°, and the period is constant. The maximum muscle activity is shown in Fig. 6.

In Fig. 6, the maximum muscle activity remains unchanged when the maximum rotation angle is less than 29°. The maximum muscle activity increases as the



(a) Maximum muscle activity (b) Passive traction F_P

Fig. 5 Dorsal flexion motion 10° ∼ 60° maximum muscle activity and passive traction F_P

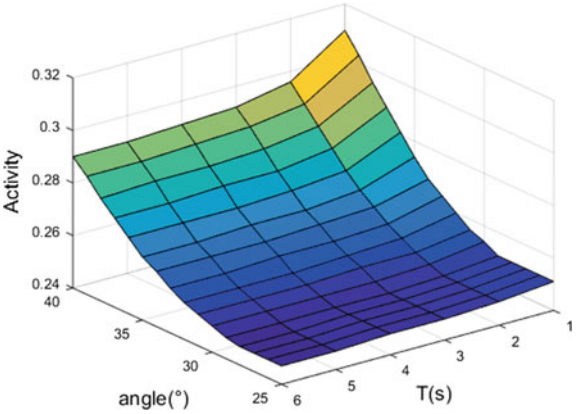


Fig. 6 Dorsal flexion motion 25° ∼ 40° maximum muscle activity

rotation angle increases when the maximum rotation angle is greater than 29°. The activity improves to 29–31% when the rotation angle reaches 40°. The muscle activity also increases when the period is shortened. Thus, the maximum rotation angle should be less than 29° in the early stage of rehabilitation to avoid the sudden increase in muscle activity.

4 Conclusions

In the present work, a novel ankle rehabilitation robot was proposed. The proposed robot could help patients with ankle plantar, dorsal flexion motion, inversion, eversion motion, and abduction, adduction motion perform ankle rehabilitation exercise. Using AnyBody, we simulated a human ankle motion on the ankle rehabilitation robot and designed a drive function based on ankle plantar, dorsal flexion motion. The maximum muscle activity and passive traction at different ankle joint angles were obtained. On the basis of the obtained results, the research found changes in period slightly affect passive traction, but affect muscle activity, the muscle activity increases when the period is shortened. So we discussed the rehabilitation strategy of dorsal flexion motion. The maximum rotation angle should be less than 29° in the early stage of rehabilitation. This strategy effectively avoided the possibility of patients' injury during ankle rehabilitation.

References

1. J. Wang, Rehabilitation of ankle joint and foot. Chinese Rehabilitation Theory and Practice. **14**(12), 1197–1198 (2008)
2. M. Girone, G. Burdea, M. Bouzit, The Rutgers Ankle orthopedic rehabilitation interface. Proc. ASME Dyn. Syst. Control Div **67**(1), 305–312 (1999)

3. J. Dai, T. Zhao, C. Nester, Sprained ankle physiotherapy based mechanism synthesis and stiffness analysis of a robotic rehabilitation device. Autonomous Robots. **16**(2), 207–218 (2004)
4. H. Takemura, T. Onodera, D. Ming et al., Design and Control of a Wearable Stewart Platform-Type Ankle-Foot Assistive Device. Int. J. Adv. Rob. Syst. **9**(5), 1–7 (2012)
5. A. Agrawal, V. Sangwan, S.K. Banala et al., Design of a novel two degree-of-freedom ankle-foot orthosis. J. Mech. Des. **129**(11), 1137–1143 (2007)
6. J. Yoon, J. Ryu, K.B. Lim, Reconfigurable ankle rehabilitation robot for various exercises. J. Robotic Syst. **22**(S1), S15–S33 (2006)
7. G. Liu, J. Gao, H. Yue et al., Design and kinematics analysis of parallel robots for ankle rehabilitation Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on. IEEE: 253–258 (2006)
8. Y. Li, L. Yao, Mathematical Modeling and Kinematics Analysis for a Novel Ankle Rehabilitation Robot. Advances in Reconfigurable Mechanisms and Robots II. Springer International Publishing, 10.1007/978-3-319-23327-7_49 (2016).
9. W. Meagan, W. Tyler, P. Stephen et al., The influence of ankle muscle activation on postural sway during quiet stance. Gait & Posture. **39**(4), 1115–1121 (2014)
10. N. Hisashi, A. Yasushi, M. Ayu et al., Identification of Individual Muscle Length Parameters from Measurements of Passive Joint Moment Around the Ankle Joint. J. Biomech. Sci. Eng. **7**(2), 168–176 (2012)
11. D. Benjamin, S. Gregory, Influence Robertson of parallel spring-loaded exoskeleton on ankle muscle-tendon dynamics during simulated human hopping. 33rd Annual International Conference of the IEEE EMBS, Boston, Massachusetts, USA, August 30–September 3: 583–586 (2011)
12. K. Hase, N. Yamazaki, Development of three dimensional whole-body musculo-skeletal model for various motion analyses. JSME International Journal Series C, Dynamics, Control, Robotics, Design and Manufacturing. **41**(1), 25–32 (1997)
13. S.L. Delp, J.P. Loan, A computational framework for simulating and analyzing human and animal movement. Comput. Sci. Eng. **2**(5), 46–55 (2000)
14. Y. IHisashi, Identification of Individual Muscle Length Parameters from Measurements of Passive Joint Moment Around the Ankle Joint. JBSE. **7**(2): 168–176 (2012)
15. G. Tang, Simulation and analysis of human body typical motion. (Shanghai, college of mechanical and power engineering, Shanghai Jiao Tong University. 2011) TP391.9
16. Q. Zhang et al. The effect of local muscle fatigue on the body sensation of the ankle joint. **31** (3): 68–74 (2011)
17. T. Zhao, H. Yu, J. Dai, Ankle rehabilitation robot based on 3-RSS/S parallel mechanism. Journal of Yanshan University. **29**(6), 471–475 (2005)

Internet of Things Technology Applies to Two Wheeled Guard Robot with Visual Ability

Chih-Hui Chiu and Yu-shiou Huang

Abstract In this study, a two wheeled guard robot (TWGR) system with visual ability is realized. The hardware of the TWGR includes a chassis carrying dc motors for wheels, the input/output (I/O) board connects the TWGR system to a personal computer (PC) and Personal Digital Assistant (PDA), driver circuits for the motors, the necessary sensors and filter circuit needed to measure system states. Moreover, camera system and image transmission module are also connected to the chassis. Moreover, based on Internet of Things (IoT) technology, all system state and real-time image from the visual system can be transmitted to the cloud server through the Internet. Administrator is able to access the cloud server to monitor each TWGR immediately. Finally, TWGRs can be controlled remotely to deal with any emergency situation.

Keywords Inverted pendulum · Fuzzy control · System integration · Internet of things (IoT)

1 Introduction

In recent years, the rapid development of semiconductor industry speeds up the robot development. To satisfy the needs of human, the various functions of the robot is be designed. Many mobile robot studies have received many attentions. The ball robot has been constructed in 1994 [1]. The similarity of this wheel-based mobile robot likes unicycle-type mobile robot [2], two-wheeled inverted pendulum

C.-H. Chiu (✉) · Y. Huang
Department of Communications, Navigation and Control Engineering, National Taiwan Ocean University, No.2, Beining Rd, Jhongjheng District, Keelung City 20224, Taiwan, R.O.C
e-mail: chchiu@ntou.edu.tw

Y. Huang
e-mail: 10567020@ntou.edu.tw

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_2

[3], mobile wheeled inverted pendulum [4], and intelligent omnidirectional mobile robots. [5] …etc.

Recently, technology of the Internet of things (IoT) has been developed rapidly [6–8]. In general, connect all the things to the Internet, it is called IoT. IoT has a wide range of applications, from wearable devices, smart cars, smart home, intelligent transportation, smart city, to the industrial networking and so on. IoT can make the real world digitization, so it has a very wide range of uses. Through the Internet of things, machines, equipment or personnel can be managed by a remote control center. Moreover, it can also be used to the family device control, car remote control, as well as the search location for stolen items, etc.

In this study, a two wheeled guard robot (TWGR) system with visual ability is designed. A TWGR can be modelled as an inverted pendulum on two coaxial wheels that have independent drives. This system is composed of a chassis carrying a DC motor coupled to a gearbox for each wheel. The image transmission system with the camera is also added. Clearly, a transmission loss exists with this configuration. Obviously, the mathematical model for a TWGR system is complex and inexact. From a control perspective, the conventional control technologies always require a good understanding of the plant. Because an exact dynamic TWGR model is difficult to obtain, controlling such a system using conventional control technologies is also difficult. Therefore, this study applies a Mamdani-type fuzzy control strategy for angle and position control of a TWGR system. Moreover, based on above IoT technology, the TWGR system can send the images back to the control center or other devices for watching wirelessly. By this way, through the Internet of things, the TWGA can be managed by a remote control center.

2 Design Method

A TWGR can be considered an inverted pendulum and is a physically unstable system. This system mainly consists of a platform, pose sensors, and a controller. The platform includes the robot body, the left and right wheels with encoders, and the electronic circuits. The TWGR in this work consists of a base that carries a DC motor, which is connected to a reduction gear for each wheel. The control algorithms are executed using a microprocessor, W78E058B40DL, which is produced by Winbond company. W78E058B40DL is an 8-bit microcontroller, which has a programmable, flash-erasable read-only memory to refresh the firmware. W78E058B40DL has 32 kilobytes of main read-only memory, 4 kilobytes of auxiliary read-only memory, 512 bytes of random access memory, four bi-directional and bit-addressable input/output (I/O) ports, one 4-bit multipurpose programmable port, three timers/counters, and one full-duplex serial port. Moreover, several self-created circuits are used in the TWGR system. The I/O circuits connect the TWGR to the microcontroller, the driver circuits for the motors, and the sensors and filter circuits that are required to measure the rider states.

In the following, the fuzzy controller of the TWGR is described. First, The TWGR can be simply considered as an inverted pendulum on a two coaxial wheels (Fig. 1). By moving the two coaxial wheels back and forth, the body of the two-wheel robot will keep balance. Therefore, we assume two dc motors will receive the same force command in balance control.

The following is a review of fuzzy inference systems. Generically, a fuzzy model is a collection of fuzzy rules $Rule^j, j = 1, \ldots, m$, where m is the number of the fuzzy rules. Each rule is expressed in the following form:

$$Rule^j : \text{If } y_1 \text{ is } \mu^j_{A_1}, \ldots, y_n \text{ is } \mu^j_{A_n}, \text{ then } u_j \text{ is } \mu_{B_j}$$

where u_j is the output variable, y_i, $i = 1, \ldots, n$ are the input variables, n is the number of inputs, $\mu_{A_1}, \ldots, \mu_{A_n}$ are standard fuzzy sets, and $\mu_{B_1}, \ldots, \mu_{B_m}$ are fuzzy singletons.

In this study, the center-of-gravity method and the *product-sum* operation were used for defuzzification. Specifically, the output, U_{Fuzzy}, of the fuzzy system can be obtained as

$$U_{Fuzzy} = \frac{\sum_{j=1}^{m} w_j \prod_{i=1}^{n} \mu^j_{A_i}(y_i)}{\sum_{j=1}^{m} \prod_{i=1}^{n} \mu^j_{A_i}(y_i)} \tag{1}$$

where w_j is the point where $\mu_{B_j} = 1$.

3 Results

This work uses a fuzzy theory to design a vision based security robot. Through the Internet of things, the TWGR can be controlled by a remote control center. Moreover, it can also send remote image back by the camera on the TWGR. In this

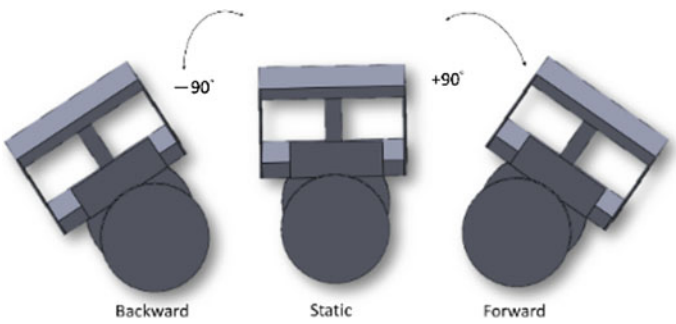


Fig. 1 The scope of the body

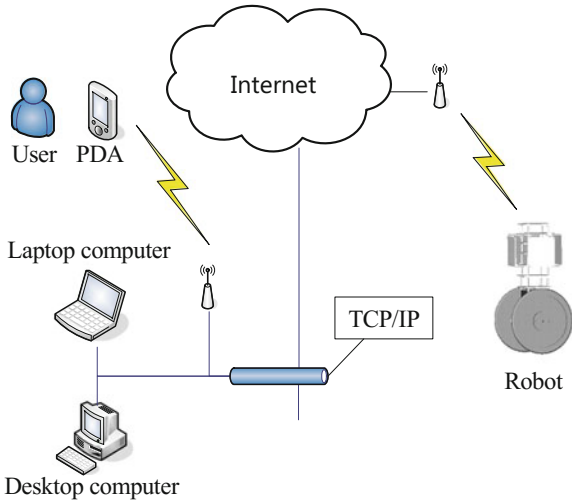


Fig. 2 User schematic

study, the notebook and the PDA are used to monitor the robot through the Internet network. According to IoT technology, it can connect all TWGRs to the Internet (shown in Figs. 2 and 3). Then, the control center can control several robots at the same time. It also can use PDA or other terminals to monitor multiple images from TWGRs. Figure 4 is the experimental system. The TWGR states can use a PDA to monitor. This experimental result is shown in Fig. 5. Figure 6 shows the remote image send by the TWGR.

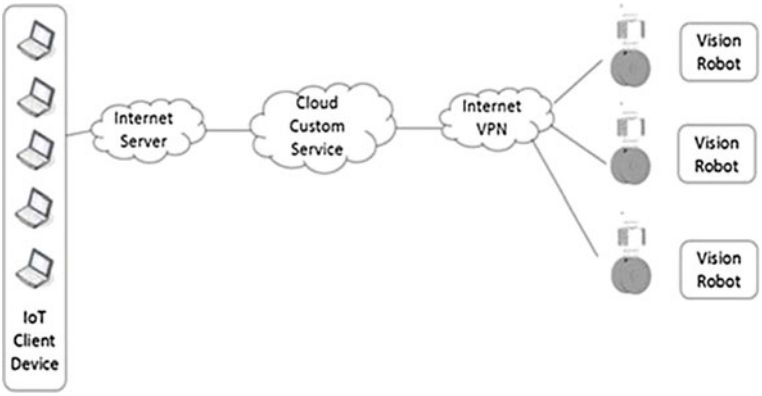
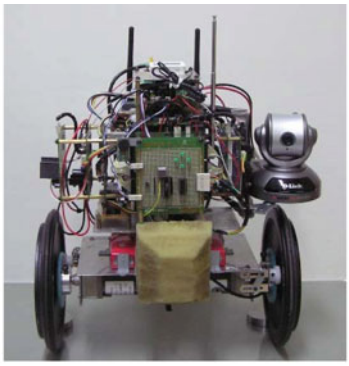


Fig. 3 Iot System block diagram



(a) TWGR

(b) PDA and control center

Fig. 4 TWGR experimental system

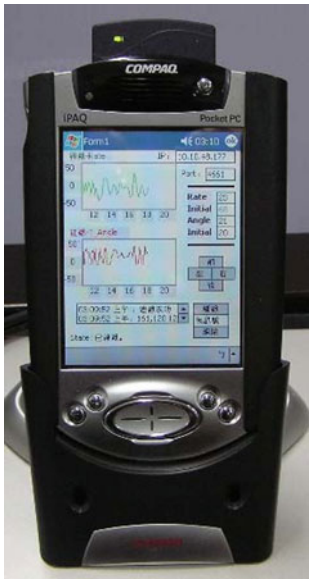


Fig. 5 PDA control center

Fig. 6 Remote image monitor interface

4 Conclusions

In this study, the model-free fuzzy control strategy for a TWGR with visual ability has been proposed. Based on basic fuzzy theory, all the simple structure controllers are designed. Finally, 8-bit microprocessors are used to implement this fuzzy controller in this robot. Based on Internet of Things (IoT) technology, all system state and real-time remote image from the visual system can be transmitted to the cloud server through the Internet. Administrator is able to access the cloud server to monitor each TWGR immediately. Experimental results have shown that the proposed control strategy has very good efficiency.

References

1. Y.S. Ha, S. Yuta, Trajectory tracking control for navigation of self-contained mobile inverse pendulum. *In Proc. IEEE/RSJ Int'l. Conf. on Intelligent Robots and Systems*, p. 1875–1882, 1994
2. E.H. Guechi, J. Lauber, M. Dambrine, M. Defoort, Output feedback controller design of a unicycle-type mobile robot with delayed measurements. IET Control Theory Appl. **6**(5), 726–733 (2012)
3. C.H. Chiu, The design and implementation of a wheeled inverted pendulum using an adaptive output Recurrent cerebellar model articulation controller. IEEE Trans. Ind. Electron. **57**(5), 1814–1822 (2010)
4. J. Huang, F. Ding, T. Fukuda, T. Matsuno, Modeling and velocity control for a novel narrow vehicle based on mobile wheeled inverted pendulum. IEEE Trans. Control Syst. Technol. **21** (5), 1607–1617 (2013)
5. H.C. Huang, SoPC-based parallel ACO algorithm and its application to optimal motion controller design for intelligent omnidirectional mobile robots. IEEE Trans. Ind. Inform. **9**(4), 1828–1835 (2013)

6. M.W. Condry, C.B. Nelson, Using smart edge IoT devices for safer, rapid response with industry IoT control operations. Proc. IEEE **104**(5), 938–946 (2016)
7. J. Wan, S. Tang, Z. Shu, Di Li, S. Wang, M. Imran, A.V. Vasilakos, Software-defined industrial internet of things in the context of industry 4.0. IEEE Sens. J. **16**(20), 7373–7380 (Oct. 15, 2016)
8. T. Bangemann, M. Riedl, M. Thron, C. Diedrich, Integration of classical components into industrial cyber-physical systems. Proc. IEEE **104**(5), 947–959 (2016)

Implementation of a Biomimetic Flapping-Wing Robot Based on Internet of Things Technology

Chih-Hui Chiu and Chieh-Min Lin

Abstract The purpose of this study is to complete the design, production and implementation of the machine animal of bionic bird. Objectives of this paper was on design and manufacture of small-scale flight mechanism with a pair of wings. The implementation of mechanical structure, system integration of the bionic robot bird platform and the design of controller has completed. It has flight, remote control, attitude feedback ability. The point of this paper, is to construct the image transmission system and integrates with the bionic robot bird platform. Based on the system platform we established above, adding the camera lens, image transmission and other modules to the fuselage. And the interface of the original system is integrated and compatible. Dynamic sensors are used to measure flight information of the bionic robot bird. Moreover, according to image transmission unit, the bionic robot bird has the ability of aerial photography. Finally, the remote images from the flying robot bird are provided to other platform by using the network of the ground control center, in purpose to integrate the function of the Internet of Things (IoT) .

Keywords Biomimetic robot control · System integration · First person view (FPV) camera · Unmanned aerial vehicle (UAV) · Internet of things (IoT)

C.-H. Chiu (✉)
Department of Communications, Navigation and Control Engineering,
National Taiwan Ocean University, No.2, Beining Rd, Jhongjheng District,
Keelung City 20224, Taiwan, R.O.C
e-mail: chchiu@ntou.edu.tw

C.-M. Lin
Department of Electrical Engineering, Yuan Ze University, Chung Li 32003,
Taoyuan City, Taiwan, R.O.C
e-mail: m0832456@gmail.com

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_3

1 Introduction

Flying in the sky is always human's dream. In past decades, there are a lot of researches and theories has mentioned about flying [1–16]. In the early stages, some researches were focus on bird's wing and try to flying in the sky by using flapping wing actions. However, these researches were failed to accomplish desire purpose. Recently, with technological advancement and semiconductor innovation, human's flying dream has come true.

In recent years, technology of the Internet of things (IoT) has been developed rapidly [17–20]. In general, connect all the things to the Internet, it is called IoT. IoT has a wide range of applications, from wearable devices, smart cars, smart home, intelligent transportation, smart city, to the industrial networking and so on. IoT can make the real world digitization, so it has a very wide range of uses. Through the Internet of things, machines, equipment or personnel can be managed by a remote control center. Moreover, it can also be used to the family device control, car remote control, as well as the search location for stolen items, etc.

Biomimetic flapping-wing robot is a new type of micro aerial vehicle (MAV). The target of this study is to develop a mimicking bird, which has a fuselage with wings. In this paper, the fuselage with wings is built firstly. Then, controller design and system integration are also completed. It has flight, remote control, and attitude feedback ability. This study focuses on the integration and implementation of the bionic bird system with the image platform to achieve the objectives of the IoT.

The hardware of the system consists of the following modules: body platform, sensor and filter circuit, controller circuit, remote control circuit, power circuit, image transmission module.

The body platform includes a brushless DC motor, two servo motors and a carbon fiber fuselage. The bionic bird is basically constructed on two wings connected to the reduction gear mechanism and brushless DC motor. By flapping the two wings up and down, the bionic bird is capable to keep flying. Sensors in this system include multi-axis inclinometer, gyro and a GPS. The bird's posture can be measured by those sensors. Filter circuit is used to reduce the noise combined with sensors' signals. Controller circuit includes system on-a-programmable-chip (SoPC) board and radio frequency (RF) receiving circuit. It is the centre of the whole system. SoPC board is utilized for the controller center. The control command is a pulse width modulation (PWM) signal sent to brushless DC motor and servo motors. The brushless speed controller and servo drive receives the PWM command and sends the required force to drive the motors. RF receiving circuit is used as the interface between user and the bionic bird.

This work mainly focused on image capture and system integration. Using the first person view (FPV) camera for aerial photographing, then send the images back to the ground station. The previous established system platform using dynamic sensor to measure system state of the bionic bird, such as tilt angle, angular velocity, GPS position and so on, through digital signal processing, an accurate calculation of the bird's flight information can be obtained.

Finally, images form the bionic bird will return to the ground station and provide to the network live platform. Through the Internet, the proposed bionic bird has ability to connect many device such as mobile phones or computers. Users can watch real time images remotely.

2 Methods and Results

In this study, an ornithopter bird shown in Fig. 1 is presented, which is made by carbon fiber fuselage and nylon fabric wing. Several electric modules are equipped. The control algorithms are executed by a microprocessor. Inertial sensors are used to measure bird posture. Wireless transceiver is the communication channel between remote controller and bird. The bird's location is got by an onboard GPS sensor. A 1000 mAh Lithium-Polymer battery provides power for the whole system, the DC brushless motor features on high efficiency, high power, high torque characteristics which drives the wing flapping action, two servo motors that has the advantage of small size, lightweight, high-speed, high torque control the rudder and elevator. The size of the bird fuselage is 140-cm long, 80-cm wide. The hardware photo of the bird including the image transmission module which consist of a main

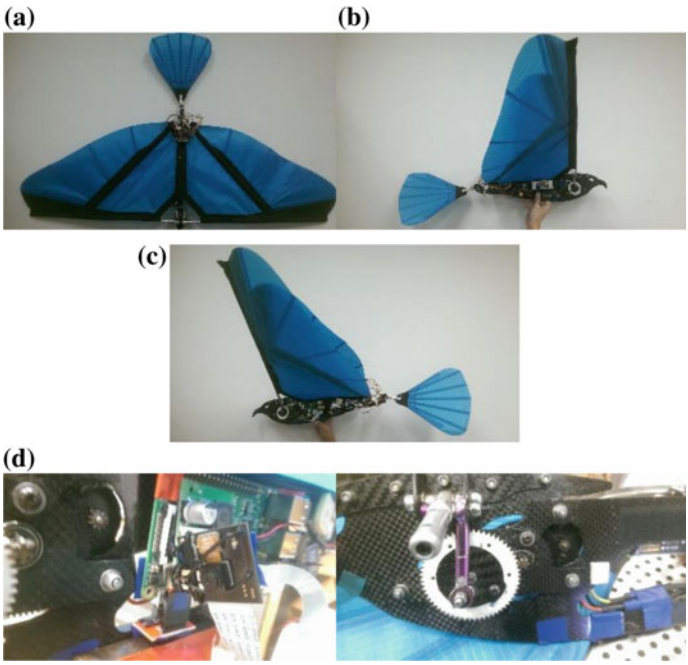


Fig. 1 The biomimetic bird. **a** Top view of system, **b** Right view of system, **c** Left view of system **d** FPV camera installation position

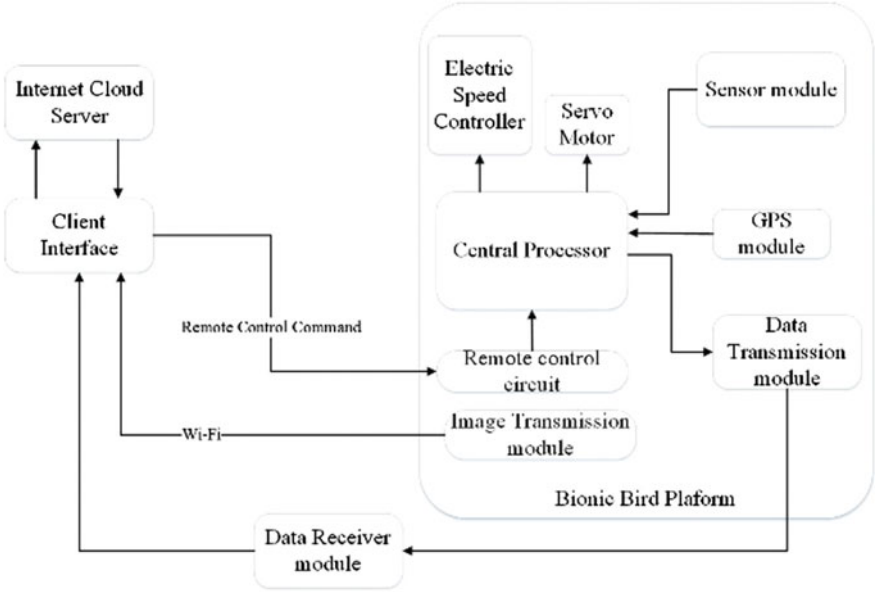


Fig. 2 Block diagram of proposed bionic robot bird

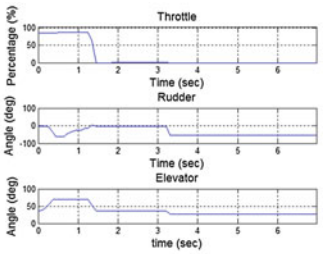
board and a FPV camera was shown in Fig. 1. Block diagram of proposed system is shown in Fig. 2.

The biomimetic bird real world control performance is shown in the following Figs. 3 and 4.

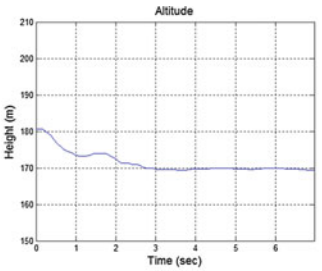
(a) Short distance (b) Middle distance

(c) Long distance

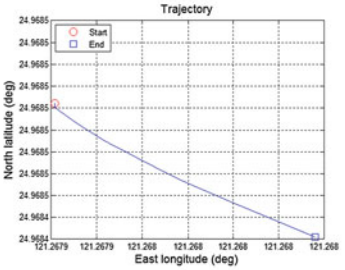
Fig. 3 Bird's flying perspective



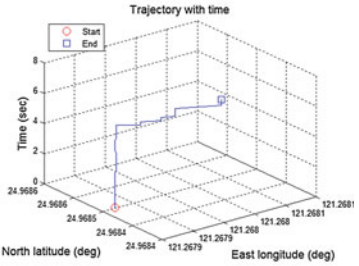
(a) throttle and system angle



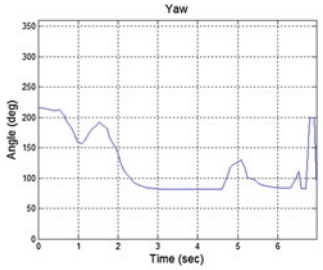
(b) altitude of system



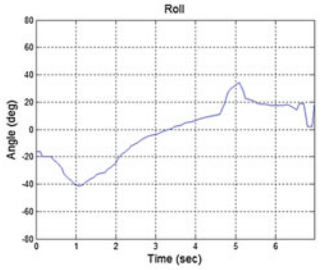
(c) 2D trajectory



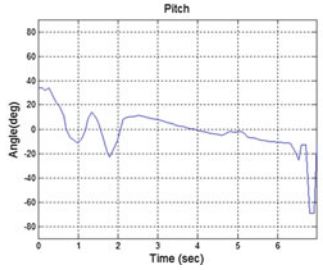
(d) 2D trajectory with time



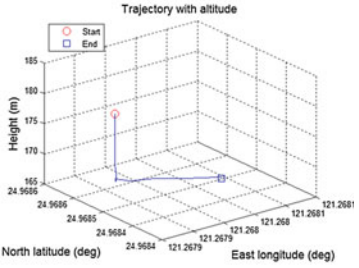
(e) yaw angle



(f) roll angle



(g) Pitch angle



(h) 3D trajtory

Fig. 4 System performence

The application of IoT technology on the proposed bionic bird is shown in Fig. 5. The addition of the image transmission system in this paper is a kind of embedded IoT interface. With the use of the wireless network function and network streaming software, all users can watch the flight process by computers or mobile devices instantly and remotely, as the schematic shown in Fig. 6. According to above presented function, the implementation of IoT in this system has been essentially completed.

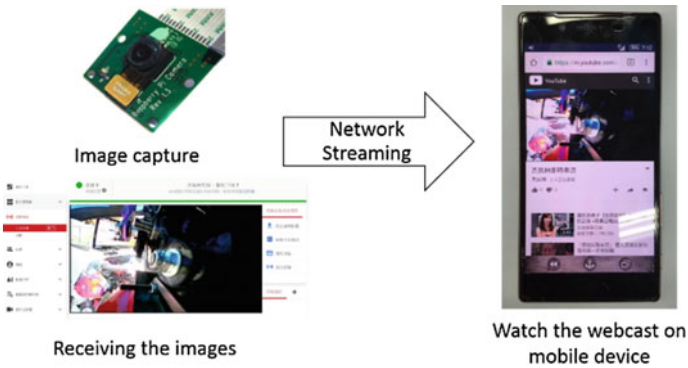


Fig. 5 Network streaming process

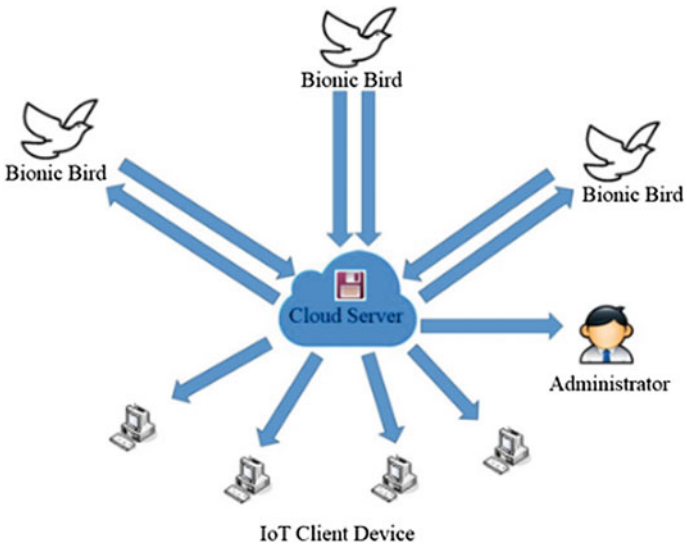


Fig. 6 Schematic of IoT function of proposed system

3 Conclusions

This paper has completed the bionic bird platform, including controller, electric circuit design, flight parameter feedback, aerial photographing and implementation of IoT, the set target has been reached. The proposed platform successfully joined the aerial photographing function. Images from the bionic bird will return to the ground station. In addition, the returned images are provided to the network live platform through the Internet, users can access the images instantly and remotely. IoT technology not only makes the proposed system become a powerful intelligent device, but also creates new possibilities of the UAV technology. And finally, effectively develop many applications in the people's livelihood, military, and industry.

References

1. J.D. DeLaurier, J.M. Harris, A study of mechanical flapping-wing flight. Aeronautical J. R. Aeronautical Soc., 277–286 (October, 1993)
2. F.E. Mehler, The Structural testing and Modification of A Full-Scale Ornithopter's Wing Spars, Master Thesis, University of Toronto Institute for Aerospace Studies, 1997
3. World First Ornithopter Takes Flight, Online Available: http://www.bbc.co.uk/news/world-us-canada-11410410
4. R. Madangopal, Z.A. Khan, S.K. Agrawal, Energetics-based design of small flapping-wing micro air vehicles. IEEE/ASME Trans. Mechatron. **11**(4), 433–438 (2006)
5. R. Żbikowski, C. Galiński, C.B. Pedersen, Four-bar linkage mechanism for insectlike flapping wings in hover: concept and an outline of its realization. ASME J. Mech. Des. **127**, 817–824 (2005)
6. S.H. McIntosh, S.K. Agrawal, Z. Khan, Design of a mechanism for biaxial rotation of a wing for a hovering vehicle. IEEE/ASME Trans. Mechatron. **11**(2), 145–153 (2006)
7. Y.F. Zhang, B.F. Song, Y.Z. Zhang, Development of Flapping Wing Micro Air Vehicle, 26th international congress of the aeronautical sciences, 2008
8. G.C.H.E. de Croon, K.M.E. de Clercq, R. Ruijsink, R.R. Remes, C. de Wagter, Design, aerodynamics, and visionbased control of The DelFly. Int. J. Micro Air Vehicles **1**(2), 71–98 (2009)
9. L.J. Yang, C.K. Hsua, J.Y. Hoa, C.K. Feng, Flapping wings with pvdf sensors to modify the aerodynamic forces of a micro aerial vehicle. Sens. Actuators A: Phys. **139**, 95–103 (2007)
10. J.H. Park, K.J. Yoon, Designing a biomimetic ornithopter capable of sustained and controlled flight. J. Bionic Eng. **5**, 39–47 (2008)
11. Aerodynamic Lightweight Design with Active Torsion, Online Available: http://www.festo.com/net/SupportPortal/Downloads/46270/Brosch_SmartBird_en_8s_RZ_110311_lo.pdf, 2011
12. C.S. Lin, C.B. Hwu, W.B. Young, The thrust and lift of an ornithopter's membrane wings with simple flapping motion. Aerosp. Sci. Technol. **10**, 111–119 (2006)
13. H. Duan, Q LI, Model and Control of Flight Attitude For Flapping Wing Micro Aerial Vehicle, 8th IEEE international conference on control and automation, 2010
14. S.S. Baek, F.L. Garcia Bermudez, and R.S. Fearing, Flight control for target seeking by 13 gram ornithopter, 8th IEEE/RSJ international conference on intelligent robots and systems, 2011
15. A.A. Paranjape, S.J. Chung, J. Kim, Novel dihedral-based control of flapping-wing aircraft with application to perching. IEEE Trans. Rob. **29**(5), 1071–1084 (2013)

16. J.S. Lee, J.K. Kim, Jae-hung Han, Charles P. Ellington, Periodic Tail Motion Linked to Wing Motion Affects the Longitudinal Stability of Ornithopter Flight, J. Bionic Eng. **9**(1), 18–28 (2012), ISSN 1672-6529
17. J. Weinman, The Economics and Strategy of Manufacturing and the Cloud, IEEE Cloud Computing **3**(4), 6–11, July-Aug. 2016
18. J. Wan, S. Tang, Z. Shu, Di Li, S. Wang, M. Imran, A. V. Vasilakos, Software-Defined Industrial Internet of Things in the Context of Industry 4.0, IEEE Sens. J. **16**(20), 7373–7380, Oct. 15, 2016
19. T. Pfeiffer, J. Hellmers, E.M. Schön, J. Thomaschewski, Empowering user interfaces for industrie 4.0. Proc. IEEE **104**(5), 986–996 (2016)
20. M.W. Condry, C.B. Nelson, Using smart edge IoT devices for safer, rapid response with industry IoT control operations. Proc. IEEE **104**(5), 938–946 (2016)

Application of a Multirotor Unmanned Aerial Vehicle to Automatic Inspection of Object Surface

Yi-Sheng Lee, Shu-Ya Tong and Jih-Gau Juang

Abstract This research has two main parts that are the use of different sensors to make quadcopter as a detective aircraft and the use of C++ and MATLAB on quadcopter software simulation. In order to maintain altitude holding in hostile environment, air pressure altimeter and optical flow are utilized. Ultrasonic sensors are used in autonomous obstacle avoidance. In addition, Pixy (CMUcam5) camera is used to do image processed and image tracking. The data caught by ultrasonic sensor and Pixy camera are processed in Arduino board. In software simulation, we use C++ programming to achieve Gyroplane pitch and roll angles of the rule table, combine images of feedback data tracking, and balance flight movement independently through their rules tables. MATLAB software simulations for PID and fuzzy control of a flying robot to meet independent verification tracking needs are given.

Keywords Quadcopter · Image tracking · Obstacle avoidance

1 Introduction

In recent years, the quadcopters are widely used in aerial photography, shipping items, even disaster relief. Furthermore, some people use them as fishing tools. The progress of quadcopter development is growing fast in these few years. Thus dynamic quadcopter program and control design experiments are more and more popular in nowadays. The article references [1] and [2] conduct the design of systems' PD controller and state feedback control law to derivate designed PD controller. In [3, 4], Simulink of MATLAB is applied to simulate and adjust PD gains, flight direction, and tracking objects. After software simulations, this study will then test hardware APM2.6 fly control board and integrate RC Radio, 3DR number biography, Pixy CMUcam5, and ultrasonic measurement components. The

Y.-S. Lee · S.-Y. Tong · J.-G. Juang (✉)
Department of Communications, Navigation and Control Engineering,
National Taiwan Ocean University, Keelung, Taiwan
e-mail: jgjuang@ntou.edu.tw

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_4

Gyroplane pitch and the roll angle of the rule table can be adjusted by setting to manual or automatic mode through RC Radio [5, 6].

2 System Description

A quadcopter, which is driven by four motors, is a multi-axis aircraft. Different from a fixed-wing aircraft, the quadcopter takes off with the rotation of the propeller. The four rotors of the quadcopter are in the same size and distribute symmetrical position. By adjusting the relative speed between different rotors, we can adjust torque and control stable flight, rotation, and turning of the quadcopter. This is varying from a helicopter which has only two rotors. The major one is in charge of driving thrust, and the minor is responsible for the function of conteracting the torque produced by the main rotor, and controlling rotation of the helicopter. Most quadcopters sold in the market drive propellers with cross steel body, which is the main structure, and four SDMODEL motors. The four rotors provide power source with quadcopter flight required thrust and change of flight attitude. Quadcopter's flight condition is controlled by the voltage's increase and decrease which are insured by the four motors to reach flight state. Quadcopters have four propellers. The steering of two of them are forward, the others are opposite, as shown in Fig. 1 [7].

3 Control Scheme

Flowchart of tracking target is shown in Fig. 2 [7]. In this section, the main system structure of entire quadcopter track is introduced, which is the combination of hardware and systems and their respective relationships. The four-rotor equipped with a RC Radio, Ardupilot (APM2.6) flight control board, Arduino Mega 2560

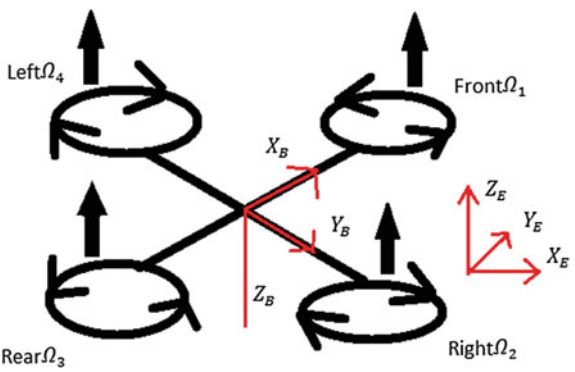


Fig. 1 Sketch map of quadcopter's rotors

board, four sets of motors, propellers, Electronic Speed Controller (ESC), and a plurality of sensors to provide stable customize and track target to quadcopter.

(a) **Quadcopter Hardware Control**

In the part of hardware control, in addition to Ardupilot (APM2.6) flight control board, install an Arduino Mega 2560 board to do the expansion. Firstly, connect the control section of RC Receiver to Ardunio Mega 2560, as in Fig. 3.

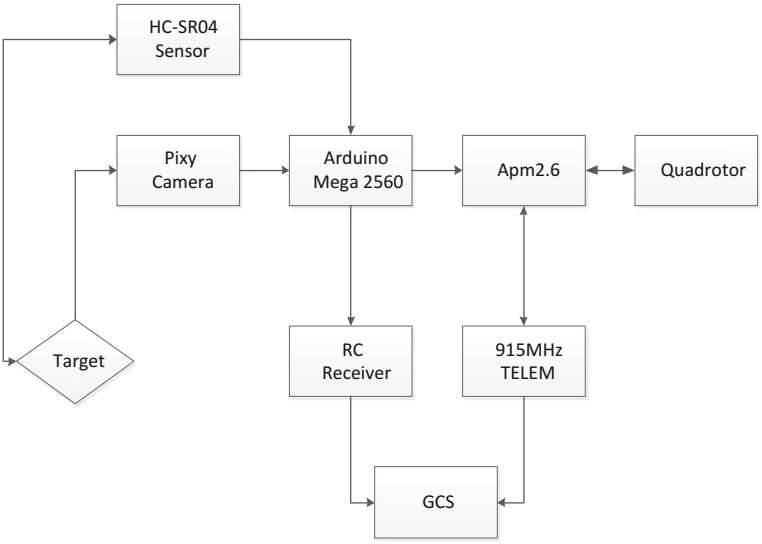


Fig. 2 Flowchart of tracking target

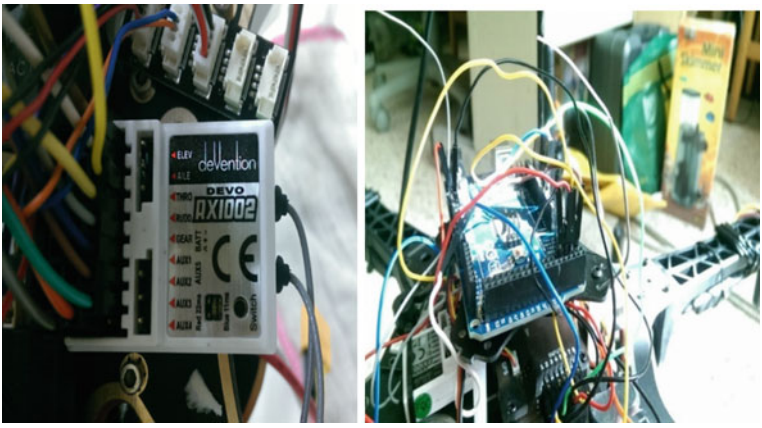


Fig. 3 RC receiver and Arduino Mega controller

(b) **Obstacle Avoidance Controller**

When it comes to obstacle avoidance, ultrasound sensors (Fig. 4) are installed at the bottom of each four faces of quadcopter to accomplish obstacle avoidance and altitude hold. Detection distance of the ultrasonic sensor is approximately 2 cm to 3 m, burst frequency is 40 kHz, current is 30 mA, and the voltage is 5 V.

(c) **Hover Point Control**

As to hover point, inertia might happen during flight. Optical flow sensor (Optical Flow v10) is used to obtain x and y values to keep at the hover point, as shown in Fig. 5.

(d) **Image Tracking Control**

In tracking control, the Pixy Camera is used. It sends block information to Arduino at 1 Mbits/second which means Pixy can send more than 6000 detected

Fig. 4 HC-SR04 ultrasonic sensor

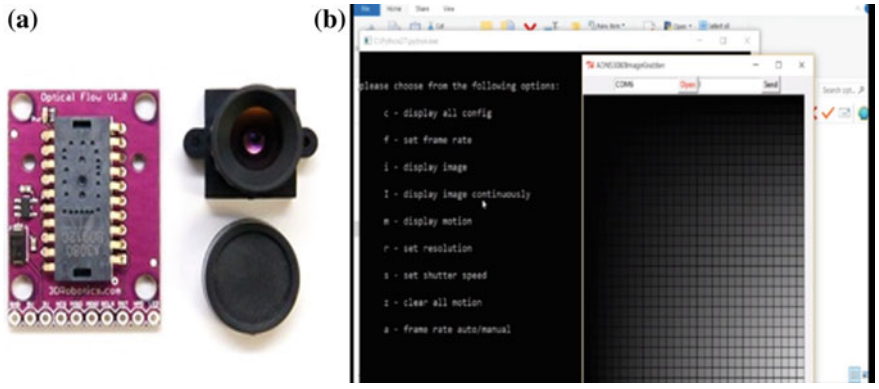


Fig. 5 **a** Optical flow sensor (Optical Flow v10), **b** interface and image of Optical Flow sensor installation

objects per second or 135 detected objects per frame (Pixy can process 50 frames per second.). Pixy uses the algorithm based on color filtering (hue-based color filtering algorithm) to identify objects. It can also detect various colors, and can read seven different colors' directions and area at the same time, as illustrated in Fig. 6.

4 Experiment Result

(a) Quadcopter Aircraft Tracking Simulation

In quadcopter's simulation, the dynamic equation and transfer function of each system are constructed to Simulink by the MATLAB package software. The simulator uses the parameters and controller gain values from Reference [2] and [8], as illustrated in Fig. 7.

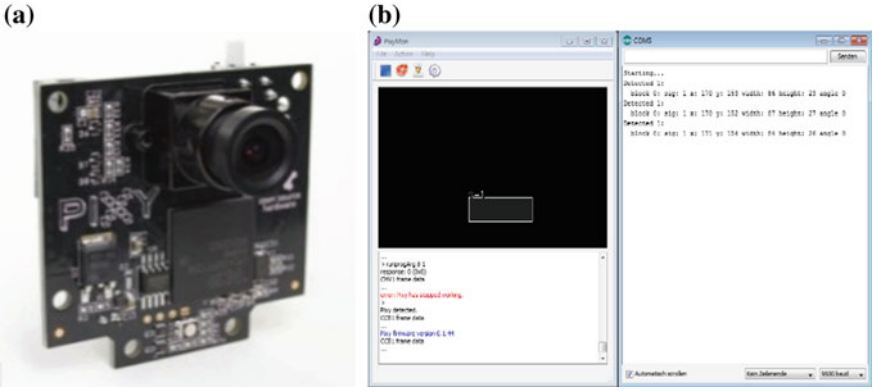


Fig. 6 **a** Pixy camera, **b** the interface of reading the size and direction

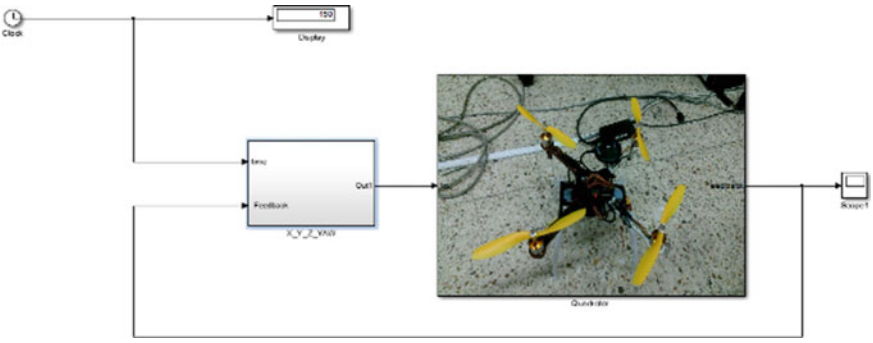


Fig. 7 Quadcopter system simulation

There are two major system blocks, which are the quadrotor and the guide law systems. The former is responsible for simulating the dynamic response of quadcopter. The latter gives the dynamic system commands such as *x _cmd*, *y _cmd, z _ cmd,* and *yaw _ cmd*, etc., with the location of the feedback after calculate. The design task is to move height, X-axis, and Y-axis sequentially. The height is set to 10 meters, X-axis displacement is 10 meters, Y-axis displacement is 10 meters. It flies on a circle, as shown in Fig. 8.

Figure 9 is the instruction maps for X, Y, Z, and the actual flight profile of the aircraft. Purple stands for the reference, and yellow represents the actual flight curve. There is a lag time that waits until the aircraft rise up in the Z-axis, then begin circling.

(b) **Indoor and Outdoor Actual Test**

In actual test of target tracking and obstacle avoidance, a target is tracked autonomously, then start cruising flight and obstacle avoidance capabilities, as shown in Fig. 10.

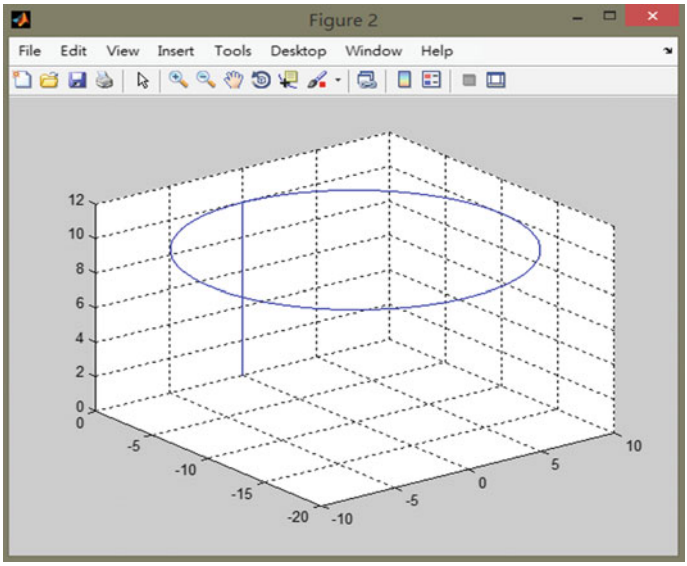


Fig. 8 Simulated mission trajectory

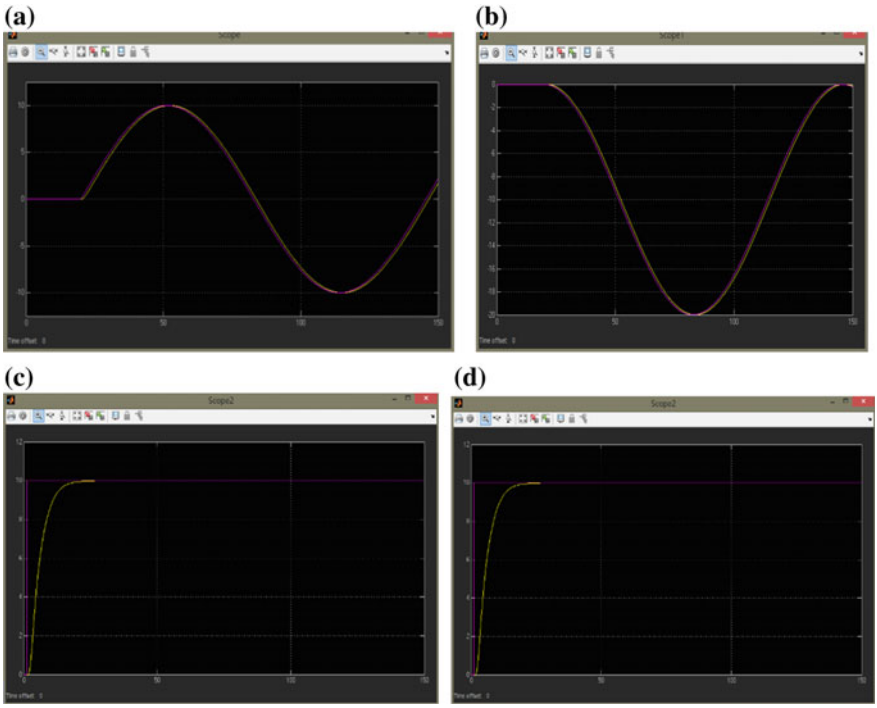


Fig. 9 **a** X-axis direction command and response, **b** Y-axis direction command and response, **c** Z-axis direction command and response, **d** motor speed

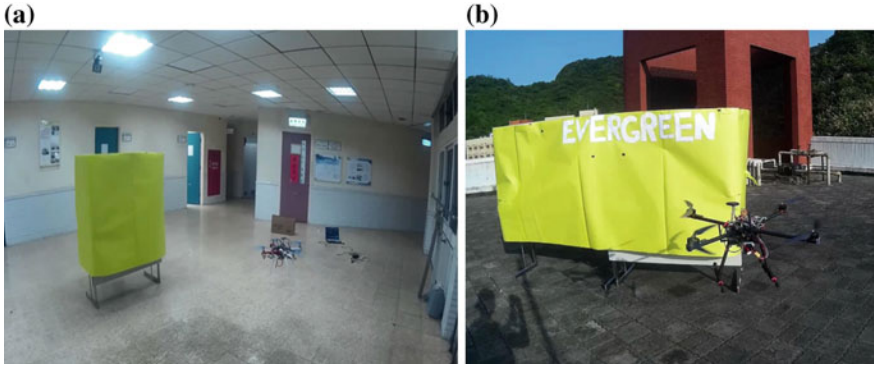


Fig. 10 **a** The actual test flight in a circle around the target (indoor), **b** the actual test flight in a circle around the target (outdoor)

5 Conclusions

Software simulations show that the fuzzy PD control truly improves the efficacy over the conventional PD controller. This study uses seven and three fuzzy sets in simulations. Initially, it had been supposed that the settling time of output response using three fuzzy sets would be much shorter. In the hardware implementation, a ship model is set up as the object. This model is used to conduct real tracking flight and object detection. The quadcopter had flied around the target in reality, and catch the predefined target color as hovering point. The real flight was integrated by multiple sensors. Input the values of every aspect of sensors in Arduino mega 2560 and conduct signal integration in program. When it comes to the part of program, use the ultrasonic sensor to conduct obstacle avoidance as the first priority in flight control. The left and right direction is controlled by the value of color catching and the size of the object, and it is the second priority in flight control. Pixy image recognition is used to grab the target then control the direction. Predefined target color can be caught in outdoor and indoor environments. But in the part of catching color, manual adjustment of the predefined color is needed.

References

1. B.C. Kuo, *Automatic Control Systems*, 9th edn. (Wiley, USA, 2010)
2. C.L. Lee, Displacement Controller Design of Quadrotor, Master Thesis, Department of Aerospace Engineering, NCKU, 2014
3. Y.J. Li, Study of Quadrotor Formation Flight, Master Thesis, Department of Aerospace Engineering, TKU, 2014
4. A.L. Salih, M. Moghavvemi, H.A.F. Mohamed, K.S. Gaeid, Flight PID Controller Design for a UAV Quadrotor. Scientific Research and Essays (2010)
5. R. Dasgupta, R. Mukherjee, A Novel Approach of Sensor Data Retrieving using a Quadcopter in Wireless Sensor Network Forming Concentric Circular Topology, Dr. Amitava Gupta Innovation Lab (Feb 17–19, 2015)
6. C.C. An, Developed a Wireless Remote Control and Multi-rotor Aerial Vehicle Systems Integration of Video Surveillance, Master Thesis, Southern Taiwan University of Science and Technology, 2014
7. Y.S. Lee, J.G. Juang, Sensor Integration and Image Tracking for Quadcopter Application, Proc. of National Symposium on Systems and Science Engineering, July 8, 2016
8. J.P. How, B. Bethke, A. Frank, D. Dale, J. Vian, Real-Time Indoor Autonomous Vehicle Test Environment, IEEE Control System Magazine, 2008

Task Execution Based-on Human-Robot Communication and Pointing Gestures

Peiqing Yan, Bingwei He, Liwei Zhang and Jianwei Zhang

Abstract Service robots have already acquired the skills to execute explicit tasks. But service robots cannot deal with the situations where ambiguities exist. When the robot encounters ambiguities such as multiple objects belonging to a same category in the surrounding environment meeting needs of the user, the robot should provide options for the user to choose by communication. In order to address this problem, we fuse verbal language and pointing gesture information to enable a robot to execute a vague task, such as "Give me the book". In our work, we propose a progressive approach that allows robots to analyze the meaning of task assigned by the user and acquire direction information from human gestures. The conducted experiments show that a humanoid robot NAO can execute vague task by applying the proposed approach.

Keywords Human-Robot communication · Pointing gesture · Verbal language command

1 Introduction

We aim to enable the robot to get adequate task information via verbal language communication rather than a robot that just can execute easy and default behavior (e.g. [1–8]). One of the intuitive interacting ways between humans and robots is implemented via verbal language. In daily life, we often need help from someone to take something over to us. The partner may have no idea about the location of the target object and we will offer information through verbal language and pointing gestures. The partner obtains spatial information via pointing gestures and gets task

P. Yan · B. He · L. Zhang (✉)
School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou, People's Republic of China
e-mail: lw.zhang@fzu.edu.cn

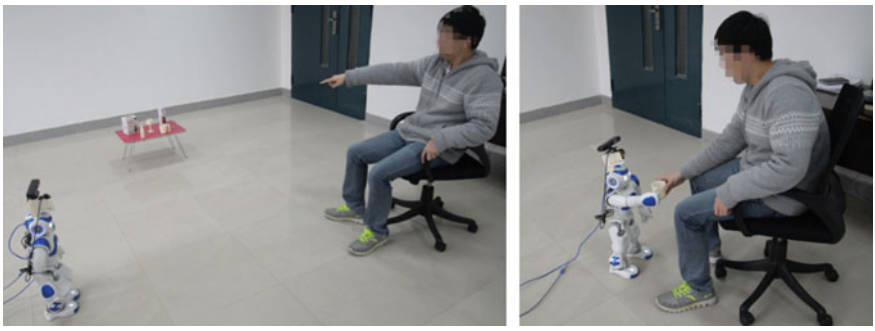
J. Zhang
TAMS, Department of Informatics, University of Hamburg, Hamburg, Germany

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_5

information via verbal information. The partner will first go to the pointed to area to seek for the target. After arriving at the area, there exist several objects that accord with the target. The partner will ask questions to confirm which object is the target one. Inspired by this point, our approach is capable of letting the robot resolve ambiguities by asking questions.

For a robot with common sense, a complete task should at least contain verbal information about what and where the target is, which action to be executed. Whether the target object is associated with other relevant objects need to be found out by perceiving the environment. For instance, a person wants to drink some coffee but he is busy with working. Then it can be convenient if a robot is able to help him bring a cup of coffee. However, the user does not habitually have all information included in just one sentence, like "Please bring me a small cup of coffee on the desk over there in this room". Instead, we may just say "bring me some coffee" to the helper, which merely contains information about an action which is "bring" and an object which is "coffee". But these two elements are insufficient for the helper to perform the task. In order to accomplish the task, the robot needs to know the location of the coffee which can be acquired via the user's gesture. After arriving at the target area, the robot may encounter with another question that there are a various kinds of cups, so the robot needs to confirm which cup satisfies with the need of the user.

To accomplish these vague tasks, robots' capability to analyze tasks information is the first and a challenging part of our work. We propose an approach that the verbal signals of tasks are first translated into texts which is then segmented using word segment technology. With these segmented phrases, the robot tries to search essential elements that composite a complete task. A task to be executed should at least contain information about "where", "what", and "action". If the segmented phrases lack any essential elements, the robot tries to inquire the missing elements from the user. (Fig. 1)



(a) Guiding the robot (b) Receiving the target object

Fig. 1 The robot accomplishes the task assigned by the user

2 Approach

An approach combining visual perception, verbal communication, task planning and action execution ensures robots are capable of obtaining information in a progressive process. The process of a robot obtaining task information based on the proposed approach is shown in Fig. 2. We focus on endowing robots the capability to interact with humans to accomplish assigned tasks based on the environment around the robots. The task analysis begins when detected verbal signals is considered as helping signals. When helping signals are detected, the robot will first find out the user's location and then face to the user. The robot can translate every word said by the users into texts, which, however, cannot be understood by the robot. We thus propose an approach that makes the robot understand the user's intention and judge if the user has provided sufficient information. In order to help the robot to understand the user, we define three task categories which are shown in Table 1.

2.1 Task Analysis

We first parse verbal commands into a structured form [2]. Verbal commands are then inferred to low-level instructions, such as "Bring me the book of John on the desk". A verb clause shown in Fig. 2 is the representation of the verbal command.

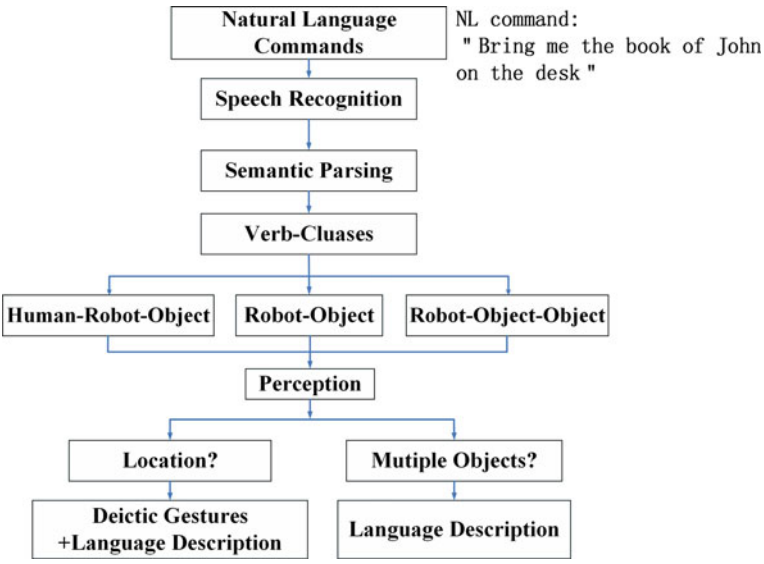


Fig. 2 The process of the proposed approach

Table 1 Three Task Categories and low-level Instructions

Tasks categories	Low-level instructions
Robot-target-user	***Seek*** *(target);* ***Grasp*** *(target);* ***Seek*** *(user);* ***Goto*** *(user);* ***Seek*** *(user);* ***Deliver*** *(target, user)*
Robot-target-object	***Seek*** *(target);* ***Grasp*** *(target);* ***Seek*** *(recipe);* ***Goto*** *(recipe);* ***Seek*** *(recipe);* ***Action*** *(recipe);*
Robot-target	***Seek*** *(target);* ***Action*** *(target);*

Each verb clause represents a single verbal command. The universal representation of a verb clause VC is shown as follow,

$$VC = (v, [[location]\,[modi]\,target], r, [recipe]) \tag{1}$$

containing the action v, an object [*target*] to be manipulated, a recipient subject or object [*recipe*], and preposition relationship r (such as "to" and "in") between [*target*] and [*recipe*]. For instance, the following verbal command can be parsed into a verb clause.

To parse verbal commands into verb-clauses $\{VC\}$, the Tensorflow-based dependency parsing library, Syntaxnet, is applied. The verb is assigned to v. The *dobj* is assigned to target and *iobj* is assigned to *recipe*. Object noun, personal pronoun and name can be assigned to *recipe*. The modifier of *dobj* is assigned to *modi*. For instance, "*the blue book of John*" has an adjective *"blue"* and a preposition *"of John"*. The relationship between the *target* and *recipe* is assigned to relationship r.

In this paper, we define three task categories which are shown in Table 1. The complex tasks are constituted of the three task categories. The verbal commands are parsed into verb clauses, and then grounded to low-level robotic instructions. The grounded robotic instructions are shown in Table 1. We define the executable actions containing *Action* = {*locate, grasp, release, seek, goto, push, pull, stop, place, turn, pour, deliver}*.

Information containing location, target objects, and actions are indispensable for the robot to execute the task. In the process of executing the task, there are ambiguities coming from the environment such as the description about the target is not unique. For instance, the task "Give me the blue book of John on the desk" can provide enough information to confirm the target object. However, the user often assigns tasks from his own perspective, such as "Bring me the book of John". The robot does not have prior knowledge about which book belonging to John. The robot thus needs to ask questions to acquire more properties about the book. The inquiry form and reply from the user is shown in Table 2. The sample task mentioned above is a Robot-Target-User type task and can be grounded to low-level

instructions {*Seek(book); Goto(book); Grasp(book); Seek(user); Goto (user); Deliver(book, user)*}. Supplementary information need to be inquired to let the robot execute the task. As shown in Table 2, when the robot has no idea about the location of the target, the robot will inquire and wait for verbal description about the location and the pointing gesture. As for the extra properties about the target, we consider three types of properties including color, volume, and relative location to identify the unique target object. These three types of properties are assigned to [*modi*] to update the instructions. The Volume contains biggest, largest, smallest and tallest. The Color contains green, pink, black, etc. The Relative location contains farthest, left, right, nearest.

Table 2 Human-Robot Communication

Task information (Where, What, Action)	Supplementary information	Inquiry form	Reply from the user
(unknown, unknown, unknown)	all	Sorry, I couldn't understand that, please tell me how I can help you again!	Assign the task again
(unknown, [target],[v])	where	Where is [targets]?	[location]
([location], [target], unknown)	action	Which action did you want me to execute?	Action
([location], [target], [v])	confirmation	Is the [targets] on the [location] over there? (with pointing gestures)	Yes
([location], [targets], [v])	what	There are multiple [targets], which is the target?	Color/Volume/ Relative location

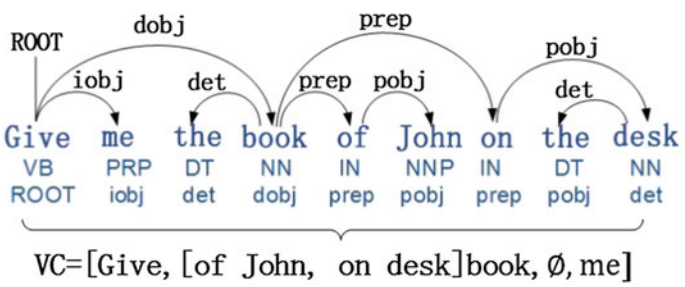


Fig. 3 A verbal command is parsed into a verb clause

2.2 Deictic Gesture

Gesture is a significant assistant interaction way in daily life. The pointing gesture combined with natural language command can give the robot direction and verbal task information to locate the target object in the indoor environment. The skeleton data [9] is retrieved using the ASUS Xtion to form a direction vector which is shown in Fig. 3. The robot can reduce the search area along the direction. The vector goes through the 3D head coordinates and the 3D hand coordinates (e.g. [10–12]). If the user does not provide pointing gesture for the robot and the robot does not know the location of the target object, supplementary information about the direction need to be inquired. The robot thus needs to ask the user where to find the target. Under the request of the robot, the user gives the robot an approximate direction to know where to find the target.

The robot in our experiments is mounted with a RGB-D camera and a laser sensor and all the sensor data are transformed to the robot's torso coordinate. A parameterized vector can be represented as follows:

$$\overrightarrow{L} = T_{camera \rightarrow torso}(\overrightarrow{P}_{head} + t(\overrightarrow{P}_{head} - \overrightarrow{P}_{hand})), \tag{2}$$

where $\overrightarrow{L}$ is the direction vector, $t \in R^{+}$ is the parameter of the direction vector, $T_{camera \rightarrow torso}$ is the transformation matrix between the RGB-D camera frame and the torso frame of the NAO robot. $\overrightarrow{L}_{head}$ is the head coordinate and $\overrightarrow{L}_{hand}$ is the hand coordinate. The target object is static in the environment, so we need to establish a global fixed coordinate. We use the state of the art mapping toolkit RTAB-Map [13] to reconstruct the environment map and the map frame is the global fixed coordinate. After reconstructing the environment, the robot can seek for the static target object along the direction provided by the user. Because the pointing gesture and the skeleton data are inaccurate, the direction vector will not hit the target object. To resolve this problem, we define a cone as the searching space which is shown as follows:

$$(\overrightarrow{P} - TT_{camera \rightarrow torso}\overrightarrow{P}_{head}) \cdot (T\overrightarrow{L}) \geq \cos\theta, \tag{3}$$

where $\overrightarrow{P}$ is all the points inside the searching cone, θ is vertex angle of the searching cone *and* T is the 6D pose of the torso in the global fix frame. The point cloud in the map is structured in an octree and the searching algorithm extracts a set of points that lie in the cone. The user in Fig. 4 used pointing gesture to provide direction information for the robot. The space in the cone is the target searching space. The cone reduces the searching space in the whole environment.

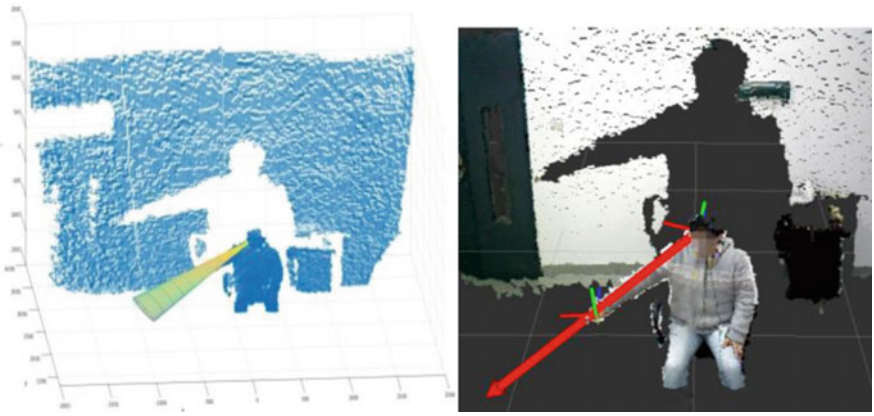


Fig. 4 The searching cone and the pointing vector

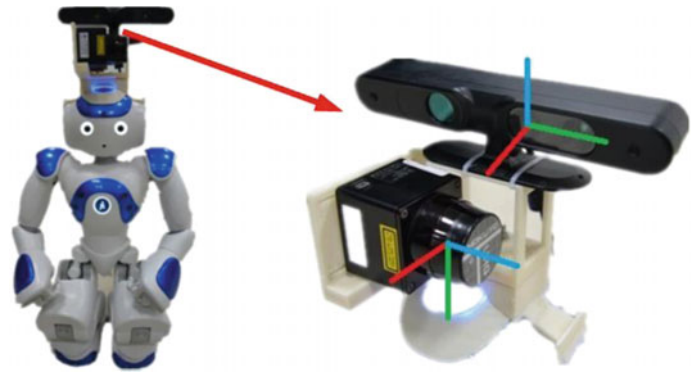


Fig. 5 Robotic platform

3 Experiments

We have conducted a thorough evaluation of the proposed approach. The experiments platform is a humanoid robot NAO. As shown in Fig. 5, the sensors are an ASUS Xtion camera and a Hokuyo URG-04LX laser sensor. The transformation matrix between the laser frame and the torso frame is obtained by calibrating the laser and the RGB-D camera [14].We use open source robot audio software

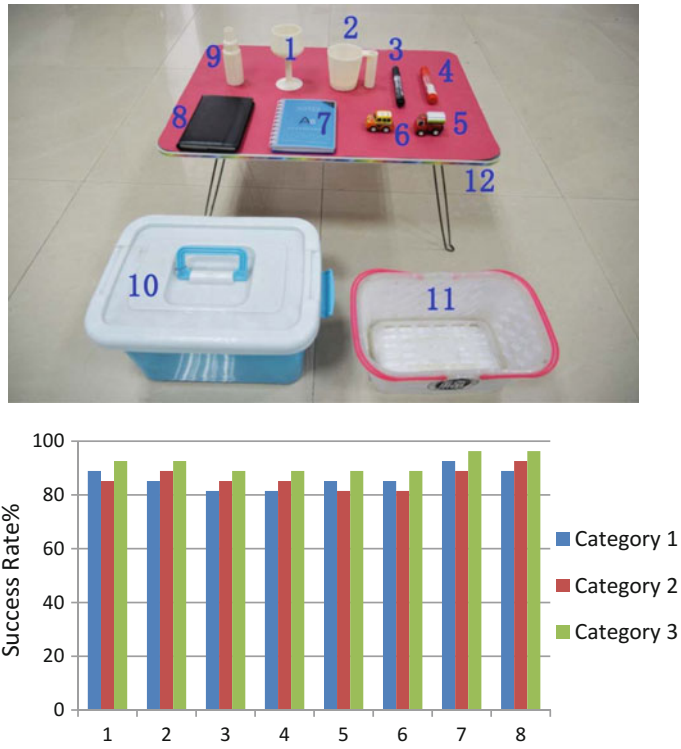


Fig. 6 The rate of the robot successfully manipulating the eight objects

HARK [15] to recognize the sound source to adjust the robot's pose. Then the robot can face to the user. The open source online Baidu speech recognition engine [16] is used for speech recognition.

Our object dataset contains a single bottle (9) and four kinds of objects (1–8) which is shown in Fig. 6. The 3D models of these objects are reconstructed [17] and we apply an algorithm combining 2D SIFT features and 3D OUR-CVFH features [18]. There were nine subjects assigning the three categories of tasks in the experiments. 216 (= 9 × 3 × 8) verbal commands in total were collected and the experiments were conducted in three days. The rate of the robot successfully manipulate the eight objects is over 81% which is shown in Fig. 6.

(a) Guiding the robot (b) Perceiving the environment
(c) Confirming the target (d) Manipulating target objects (e) Receiving the target object

Fig. 7 The process of executing the assigned task

4 Conclusions

The presented approach contains semantic parsing, direction vector forming and ambiguities settling through human-robot communication can be successfully applied on a service robot. Our approach enables the robot to understand the verbal commands by transforming it into verb-clauses and grounding it to low-level robotic instructions. The robot executes tasks by following the robotic instructions step by step. There are ambiguities coming from natural language commands and the actual environment. For instance, the user does not specify a unique object in the verbal commands. More than one object in the environment may be coincident with the object said by the user. To solve the ambiguities, supplementary information including volume, color and relative location should be inquired by the robot to manipulate the unique object (Fig. 7).

Acknowledgements This work was supported by the National Natural Science Foundation of China (Project No. 61673115, 61473090).

References

1. K. Sugiura, K. Zettsu, *Rospeex: A Cloud Robotics Platform for Human-Robot Spoken Dialogues* (Proc, IROS, 2015)
2. S.A. Tellex, T.F. Kollar, S.R. Dickerson et al., *Understanding Natural Language Commands for Robotic Navigation and Mobile Manipulation* (2011)
3. P.E. Rybski, J. Stolarz, K. Yoon et al., Using dialog and human observations to dictate tasks to a learning robot assistant. Intel. Serv. Robot. **1**(2), 159–167 (2008)
4. T. Fong, C. Thorpe, C. Baur, Robot, asker of questions. Robot. Auton. syst. **42**(3), 235–243 (2003)
5. S. Tellex, R. Knepper, A. Li et al., in *Asking for Help Using Inverse Semantics*. Robotics: Science and Systems, vol 2 (2014), p. 3
6. F.M. Carlucci, L. Nardi, L. Iocchi et al., *Explicit Representation of Social Norms for Social Robots* (Proc, IROS, 2015)
7. J. Thomason, S. Zhang, R. Mooney et al., in *Learning to Interpret Natural Language Commands Through Human-Robot Dialog*. Proceedings of the Twenty-Fourth international joint conference on Artificial Intelligence (IJCAI) (2015)
8. H. Khayrallah, S. Trott, J. Feldman in *Natural Language For Human Robot Interaction*. Proc. Workshop on Human-Robot Teaming at the 10th ACM/IEEE International Conference on Human-Robot Interaction (Portland, Oregon, 2015)
9. Open Natural Interaction (OpenNI).http://www.openni.ru/openni-sdk/index.html
10. M. Eldon, D. Whitney, S. Tellex, Interpreting Multimodal Referring Expressions in Real Time
11. C. Perez Quintero, R.T. Fomena, A. Shademan et al., in *SEPO: Selecting by Pointing as an Intuitive Human-Robot Command interface*. Proc. ICRA (2013), pp. 1166–1171
12. P. Kondaxakis, J. Pajarinen, V. Kyrki, *Real-time Recognition of Pointing Gestures for Robot to Robot Interaction* (Proc, IROS, 2014)
13. Real-Time Appearance-Based Mapping. http://introlab.github.io/rtabmap/
14. J. Briales, J. Gonzalez-Jimenez, *A Minimal Solution for the Calibration of a 2D Laser-Rangefinder and a Camera Based on Scene Corners* (Proc, IROS, 2015)
15. K. Nakadai, T. Takahashi, H. Okuno, H. Nakajima, Y. Hasegawa, H. Tsujino, in *Design and Implementation of Robot Audition System HARK Proceedings of Advanced Robotics* (2010), pp. 739–761
16. Baidu Speech Recognition Engine. http://yuyin.baidu.com/docs/asr/
17. J. Prankl, A. Aldoma, A. Svejda et al., in *RGB-D Object Modelling for Object Recognition and Tracking*. Proc. IROS (2015), pp. 96–103
18. A. Aldoma, F. Tombari, J. Prankl et al., in *Multimodal Cue Integration Through Hypotheses Verification for RGB-D Object Recognition and 6DOF Pose Estimation*. Proc. ICRA (2013), pp. 2104–2111

Finger Vein Image Feature Extraction

Jian Lin and Shuncong Zhong

Abstract Aiming at extracting the finger vein patterns in low-quality vein images, a finger vein image pattern extraction algorithm based on Hessian matrix was proposed. Firstly, Hessian matrix was acquired by the convolution of all the second derivative of Gaussian filter and finger vein image. Subsequently, the Hessian matrix eigenvalues of each pixel were calculated and the pixels not belonging to finger vein region according to the requirement of the eigenvalues were filtered. Finally, the maximum of the eigenvalues at each pixel under different scales was chosen as the output, and a finger vein image was segmented effectively after image binarization and morphological filtering. The branching points of the refined finger vein pattern were found as the feature points. The results indicated that the algorithm performs well in separating the vein region from the non-vein region and in extracting the feature points.

Keywords Finger vein recognition · Image segmentation · Hessian matrix

1 Introduction

Finger vein recognition is a new biometrics technology, which can realize identity authentication by extracting the vein features of finger vein images. It has the characteristics of living identification, high security and is easy to use [1]. Finger vein image segmentation is an important part of finger vein recognition technology,

J. Lin
Laboratory of Optics, Terahertz and Non-Destructive Testing & Evaluation,
School of Mechanical Engineering and Automation, Fuzhou University,
Fuzhou 350108, China

S. Zhong (✉)
Fujian Key Laboratory of Medical Instrument and Pharmaceutical Technology,
Fuzhou University, Fuzhou 350108, P. R. China
e-mail: zhongshuncong@hotmail.com

S. Zhong
Shanghai University, Shanghai, P. R. China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_6

which separates venous region from non-vein region, and obtains a binary image containing only finger vein pattern. The segmentation results directly affect the follow-up processing results of feature extraction and matching [2].

Due to the different thickness of bone and muscle tissue in the finger, the finger vein image has the characteristics of small contrast and uneven gray distribution [3]. Therefore, the classical threshold segmentation method can not be applied to the segmentation of finger vein image directly. In recent years, scholars at home and abroad have done a lot of research on finger vein image segmentation algorithm. Miura et al. proposed the maximum curvature algorithm [4], which mainly used the gray level distribution of the vein cross profile to extract the vein pattern and was a classic algorithm for finger vein image segmentation. Yu Chengbo et al. [5] proposed a method based on the direction of the valley-shaped detection to extract finger vein features This method mainly used the convolutions of 8 direction templates and finger vein image to detect the valley area, and obtains the finger vein pattern by multiple threshold segmentation, and the method has good effect of extracting the vein pattern; Chen Liwei. [6] proposed a the finger vein segmentation algorithm based on multi-threshold fuzzy enhancement, which combined the traditional threshold method and fuzzy enhancement algorithm, and the vein pattern was accurate and directional; Wang Baosheng et al. [7] proposed a finger vein image segmentation algorithm based on level set. Current finger vein image segmentation algorithms have their own advantages, but there are also some shortcomings.

The eigenvalues of the Hessian matrix are often used for vascular enhancement in medical images, because they can be used to detect linear structural objects in images [8]. In 1995, Koller et al. [9] proposed a multiscale enhancement method based on Hessian matrix eigenvalues for the detection of vascular structures in 2D and 3D images. Subsequently, many scholars began the detection of blood vessels based on Hessian matrix [10–13]. Frangi et al. [12] constructed a multi-parameter similarity function for good vascular enhancement effect by analyzing the relationship among tubular structures, lamellar structures, spherical structures and Hessian matrix eigenvalues.

Aiming at the characteristics of finger vein image, we proposed a method of automatic segmentation of finger vein image based on Hessian matrix. The experimental results demostrated that this method could segment the finger the vein patterns and could extract their feature points effectively.

2 Finger Vein Pattern Extraction

2.1 Principle of Hessian Matrix

The Hessian matrix [14] is a real symmetric square formed by the second-order partial derivatives of a multivariate function, and is often used for the enhancement of blood vessels in medical images. To judge whether a certain pixel P belongs to the blood vessel region, we need to analyze the local characteristic of the pixel P

neighborhood, which can be expressed approximately by the Taylor expansion in the neighborhood of the pixel P, the mathematical expression is:

$$I(p+\Delta p) \approx I(p) + \Delta p^T \nabla I(p) + \Delta p^T H(p) \Delta(p) \tag{1}$$

where $\nabla I(p)$ and $H(p)$ are the gradient vector and Hessian matrix respectively. For the 2D image $I(x, y)$,the mathematical expression of Hessian matrix at pixel P is:

$$H(x, y) = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{bmatrix} \tag{2}$$

where I_{xx}、I_{xy}、I_{yx} and I_{yy} represent the second-order partial differential of the 2D image in four different directions respectively, which reflect the local second-order structure of the pixel p.

The two eigenvalues of the Hessian matrix and the eigenvectors v_1、v_2 corresponding to the two eigenvalues λ_1、λ_2 are calculated. Let $|\lambda_1| \geq |\lambda_2$, from the geometric meanings of the eigenvalues of the Hessian matrix and its corresponding eigenvectors [15], we know that two orthogonal directions are acquired by eigenvalue decomposition, which means the eigenvectors v_1、v_2 corresponding to λ_1、λ_2 are perpendicular to each other. For the pixel points belonging to the blood vessel region, the eigenvector v_2 correspond to the eigenvalue λ_2 having a small absolute value points to the blood vessel direction, because the local gradation change is very small along the blood vessel direction. The eigenvector v_1 corresponding to the eigenvalue λ_1 with larger absolute value is perpendicular to the direction of the blood vessel, because the gray scale of the vascular cross profile changes more violently.

The eigenvalues of the Hessian matrix can be used to detect the linear structure objects in the image. The combination of eigenvalues satisfying the linear structure shape type is shown in Table 1:

Because the background gray value of the 2D finger vein image in this study is higher than the gray value of the finger vein region, that is, the vein region of the finger is dark in the image. We can see from the Table 1 that, in the ideal circumstance, the Hessian matrix eigenvalues of the region pixels should satisfy the following condition [16]:

$$\lambda_2 \approx 0, \quad \lambda_1 > T_\lambda, \quad |\lambda_1| \geq |\lambda_2|, \quad T_\lambda = 0 \tag{3}$$

In general, the gray distribution of finger vein image is not uniform, and the edge information of finger vein is blurred, which is not the ideal condition. Therefore,

Table 1 Eigenvalue combinations and structure types

Eigenvalue combination	Structure type
$\lambda_1 < 0$, $\lvert\lambda_1\rvert$ much bigger than $\lvert\lambda_2\rvert$	Line structure(dark line)
$\lambda_1 > 0$, $\lvert\lambda_1\rvert$ much bigger than $\lvert\lambda_2\rvert$	Line structure(bright line)

this research makes some improvements to formula (3), as shown in formula (4) and formula (5):

$$\lambda_1 - \lambda_2 > T_\lambda, \quad \lambda_1 \geq \lambda_2, \quad \lambda_1 > 0 \tag{4}$$

$$T_\lambda = \sum (\lambda_1 - \lambda_2)/N \tag{5}$$

where N is the number of pixels in the vein area, after many experiments, T_λ will be set to the average value of the Hessian matrix eigenvalues difference at each pixel. The eigenvalues image formed by the eigenvalues of the pixels in the vein region is obtained by the above equations.

2.2 Multi-scale Analysis of Finger Vein Image

Multi-scale Gaussian filter is constructed by Gaussian function, which can suppress the image noise and get multi-scale images by the convolution of multi-scale Gaussian filter and the image. We can know from the convolution formula of the second derivative that the convolution of the Gaussian filter and the second derivative image is equivalent to the convolution of the second derivative of the Gaussian filter and the original image. Therefore, the formula of the elements in the Hessian matrix are given by Eqs. (6) and (7):

$$I_{xx}(x, y, \sigma) = I(x, y, \sigma) * (\sigma^2 \times \frac{\partial^2}{\partial x^2} G(x, y, \sigma)) \tag{6}$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp(\frac{-(x^2 + y^2)}{2\sigma^2}) \tag{7}$$

where σ is the standard deviation of the Gaussian function, also known as the spatial scale factor, $\sigma^2 \times \frac{\partial^2}{\partial x^2} G(x, y, \sigma)$ is the normalized Gaussian second derivative [17].

The multi-scale analysis of the image is conducted to get finger vein image eigenvalues under different scales, when the Gaussian filter window radius is generally taken as 3σ. In this study, the maximum eigenvalue of each pixel in the vein region is extracted as the output of this point, and the eigenvalue of the non-vein region pixel is 0. Thus, the eigenvalue map of the finger vein pattern at different scales are fused, the mathematical expression is:

$$\lambda(x, y) = \max(\lambda(x, y, \sigma)) \quad \sigma \in [\sigma_{\min}, \sigma_{\max}] \tag{8}$$

where $\lambda(x, y)$ is the eigenvalues of the pixel in the vein region after multi-scale fusion, $\lambda(x, y, \sigma)$ is the maximum eigenvalue of the pixel in the vein region at the single scale. The maximum eigenvalue of the pixel at the multi-scale scale is selected as the output of the point by iterative scaling factor σ. According to the

finger vein width in the image, iteration scale factor interval $[\sigma_{\min}, \sigma_{\max}]$ is selected, and the iteration step is $\triangle step$.

In general, the detection results of the vein with small diameter are better when the scale factor σ is small, and the detection results of the vein with large diameter are better when the scale factor σ is large.

2.3 Extraction of Minutiae Points

The branching points and the ending points in the vein pattern skeleton image are the two type of critical points to be extracted [18]. 3 × 3 pixel-wise operation(as follow Fig. 1) can be used to determine where the minutiae points are.

If P_0 is 1, and the number of transition N_{trans} between 0 and 1 from P_1 to P_8 is equal to or greater than 6, then P_0 is a branching point. When P_0 is an ending point. In this study, the branching points are chosen as the feature points, and the number of transition N_{trans} is equal to 2. N_{trans} is given by Eq. (9):

$$N_{\text{trans}} = \sum_{1}^{8} |p_{i+1} - p_i| \tag{9}$$

3 Experiment and Result Analysis

3.1 Experimental Environment and the Determination of Relevant Parameters

In this study, finger vein images were collected by the finger vein image acquisition module designed by the laboratory. The finger vein image acquisition module is shown in Fig. 2. The module includes 850 nm infrared LED light source and USB camera, where infrared LED light source is at the top of the module, and the USB camera is located at the bottom of the module. Finger vein acquisition module uses the transmission method. The finger is put in the middle of the collection module, where is full of the infrared light perpendicular to the back of the finger. Because

Fig. 1 3 × 3 model

P_1	P_2	P_3
P_8	P_0	P_4
P_7	P_6	P_5

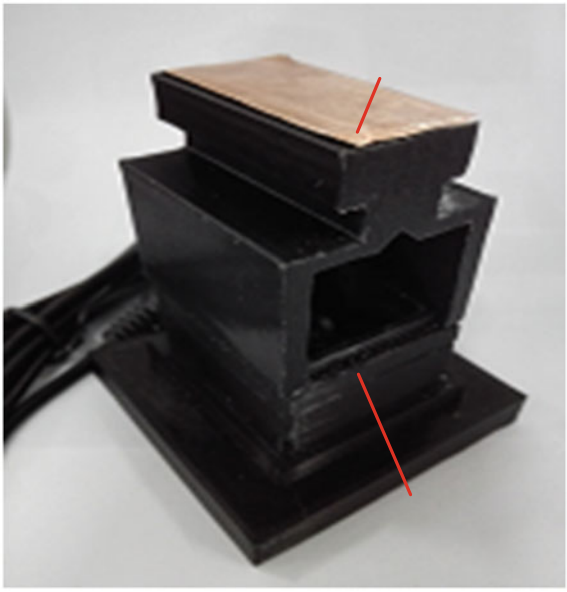


Fig. 2 Image acquisition module

the flowing blood can absorb 850 nm wavelength of infrared light [2], we can get finger vein images from the camera at the bottom of the module.

Scale factor range should be adapted to the vein width of the finger vein image. After several experiments, the scale factor σ range in this study is set to [1.5, 2.5], and the iterative step is 0.1, the total number of iterations is 11 times. The corresponding Gaussian filter window radius is 3σ.

3.2 Finger Vein Image Feature Extraction Process

The resolution of the finger vein image is 320 pix $\times$ 240 pix. As shown in Fig. 3a, the region with smaller gray value is the finger vein region, which means that the dark line is the finger vein. Steps involved for finger vein pattern extraction are given in Algorithm 1. The algorithm is programmed on the MATLAB 2012 platform. All the experimental results are run on a computer with 3.30 GHz CPU and 4 GB memory.

Algorithm 1 finger vein pattern extraction
Input finger vein image $I(x, y)$,the scale factor range σ, the iterative step by step
Output fused eigenvalue map $\lambda(x, y)$

1: do convolution between $I(x, y)$ and all the normalized Gaussian second derivative to get Hessian matrix at each pixel
2: calculate the eigenvalues λ_1、λ_2, let $|\lambda_1| \geq |\lambda_2$
3: **for all** $I(x, y, \sigma)$ **do**

if $\lambda_1 > 0$, $\lambda_1 - \lambda_2 \geq T_\lambda$ **then**

$\lambda(x, y, \sigma) = \lambda_1$

Else

$\lambda(x, y, \sigma) = 0$

end

end

4: $\lambda(x, y) = \max(\lambda(x, y, \sigma)) \quad \sigma \in [\sigma_{\min}, \sigma_{\max}]$
5: **return** $\lambda(x, y)$

The Hessian matrix of each pixel is constructed by convolving the second derivative of the Gaussian filter with the image. From the formula (4), we can know that the maximum value of the Hessian matrix must be greater than zero for the pixels of the vein region, and the difference between the maximum and minimum eigenvalues is larger than the average of the difference of the Hessian eigenvalues. A single-scale finger vein image eigenvalues map is obtained.

In Fig. 3b–e, here shows the eigenvalues map and the binarization results for single-finger vein images with scale factors of 1.5, 1.8, 2 and 2.5 respectively. It can be seen from the figure that the smaller the scale is, the smaller diameter of the vein can be detected.

After eleven iterations, the eigenvalue maps of finger vein images at each scale were obtained, According to Eq. (8), the fusion of eigenvalue maps of multi-scale finger vein images is completed, and the final finger vein pattern with different widths is obtained, as shown in Fig. 3f. After binarization of the threshold value, a finger vein pattern is obtained, as shown in Fig. 3g. It can be seen from the figure, after the segmentation of the finger vein image, there are still some pixels belonging to non-venous structure, and regional discontinuous vein lines, so the author combined median filtering and morphological processing method to filter non-vein structure pixels and connect discontinuous vein lines. The binary image processed by this method is as shown in Fig. 3h below. Figure 3i shows the refined finger vein pattern. And the feature points are shown with red color in Fig. 3j. The time taken to process the finger vein image by Matlab R2012b is about 540 ms.

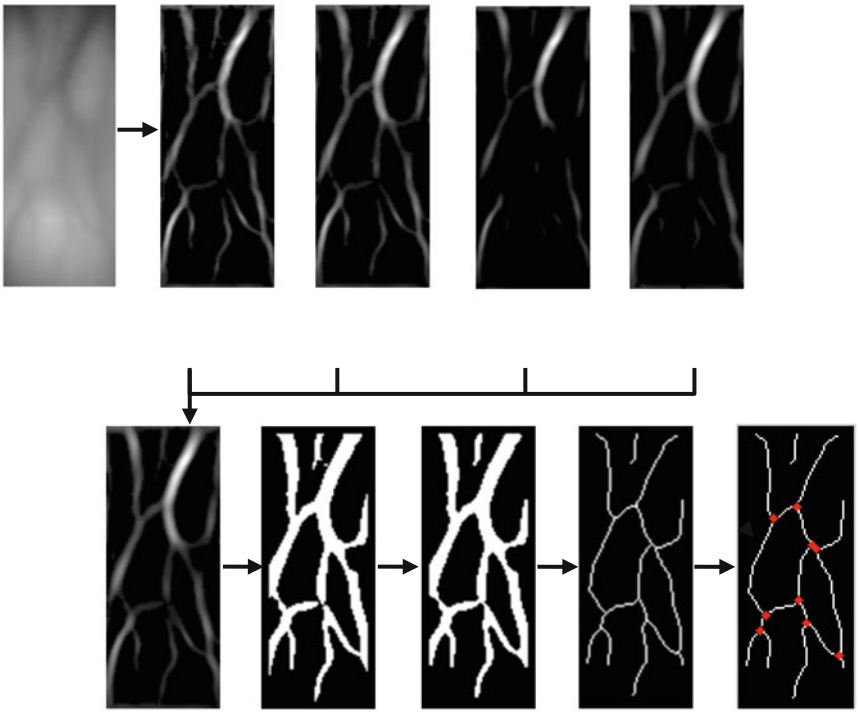


Fig. 3 Extraction and refinement of finger vein patterns

3.3 Finger Vein Image Segmentation Algorithm Effectiveness

In order to illustrate the effectiveness of this method, two additional finger vein images were selected, called img1 and img2. The resolution of each finger vein image was 60 pix $\times$ 150 pix. The finger vein pattern obtained by finger vein image segmentation algorithm was shown in Fig. 4. As can be seen from the figure, the gray value contrast of the finger vein image is low, but the extraction effect of finger vein pattern is still good.

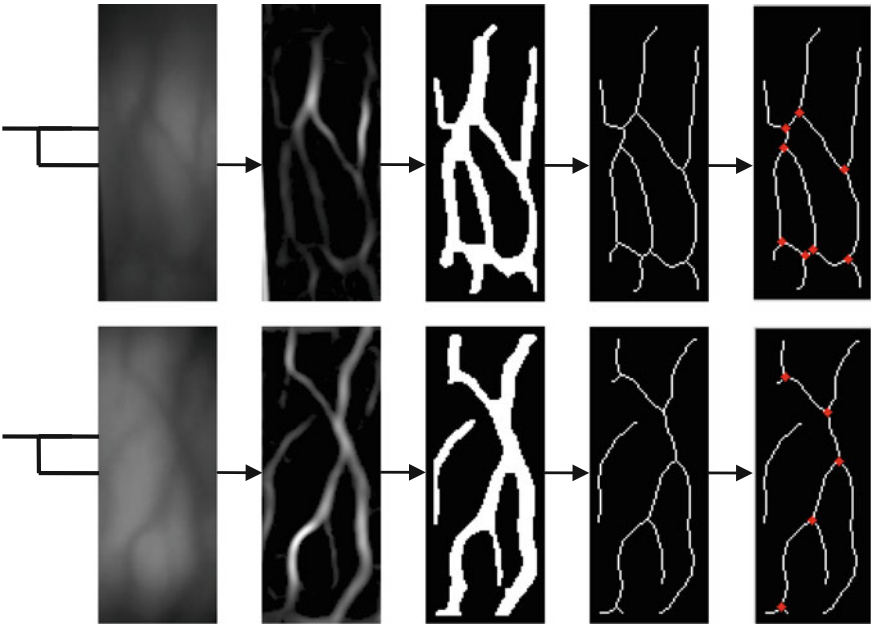


Fig. 4 The other two finger vein image segmentation results

4 Conclusions

In this study, finger vein image segmentation algorithm was developed for gray-scale non-uniform, low-contrast, weak-edge finger vein images. By using the correspondence between the eigenvalues combination of Hessian matrix and the linear structure object, the non-vein region pixels are filtered out, and the region of the finger vein is segmented. Considering the fact that finger vein image is not ideal, the average value of Hessian eigenvalue difference of each pixel are taken as threshold.

The experimental results show that the proposed method can extract finger vein pattern with different widths, and the finger vein pattern is continuous. This method is feasible from the viewpoint of Hessian eigenvalue and segmentation result is good. Subsequent work will focus on finger vein feature matching.

Acknowledgements This work is supported by the Fujian Provincial Excellent Young Scientist Fund (2014J07007), the National Natural Science Foundation of China (51675103), the Training Program of Fujian Excellent Talents in Universities, the Specialized Research Fund for the Doctoral Program of Higher Education, the Ministry of Education, P. R. China (20133514110008) and the Ministry of Health, P.R. China (WKJ-FJ-27).

References

1. Qiu Ying, Tan Dingzhong, Sun Hui, Current situation of the development of the finger vein recognition technology. Exp. Technol. Manag. **27**(3), 39–40 (2010)
2. Yang Ying, Yang Gongping, Review on Finger Vein Recognition. J. Front. Comput. Sci. Technol. **6**(4), 343–354 (2012)
3. Wang Kejun, Ma. Hui, Li Xuefeng et al., Finger vein pattern extraction method using oriented filtering technology. J. Image. Graph. **16**(7), 1206–1212 (2011)
4. N. Miura, A. Nagasaka, T. Miyatake, Extraction of finger-vein patterns using maximum curvature points in image profiles. IEICE TRANS. Inf. Syst. **90**(8), 1185–1194 (2007)
5. Y. Chengbo, Q. Huafeng, Z. Lian, in *A Research on Extracting Low Quality Human Finger Vein Pattern Characteristics[C]*, 2008 2nd International Conference on Bioinformatics and Biomedical Engineering. IEEE (2008), pp. 1876–1879
6. Chen Liwei, Ni Jie, Method for finger vein feature segmentation based on multi-threshold fuzzy enhancement. Appl. Sci. Technol. **38**(4), 14–18 (2011)
7. Wang Baosheng, Chen Yufei, Zhao Weidong et al., Finger-vein image segmentation based on level set. J. Comput. Appl. **36**(2), 526–530 (2016)
8. Cao Rongfei, X. Wang, Wu Zhongke et al., Cerebrovascular segmentation method based on hessian matrix and clustering. J. Syst. Simul. **26**(9), 2104–2109 (2014)
9. T M. Koller, G. Gerig, G. Szekely et al., in *Multiscale Detection of Curvilinear Structures in 2-D and 3-D Image Data[C]*. Computer Vision, 1995. Proceedings., Fifth International Conference on. IEEE (1995), pp. 864–869
10. C. Lorenz, I.C. Carlsen, T.M. Buzug et al., in *Multi-scale Line Segmentation with Automatic Estimation of Width, Contrast and Tangential Direction in 2D and 3D Medical Images[C]*. CVRMed-MRCAS'97 (Springer, Berlin Heidelberg, 1997), pp. 233–242
11. Y. Sato, S. Nakajima, N. Shiraga et al., Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. Med. Image Anal. **2**(2), 143–168 (1998)
12. A.F. Frangi, W.J. Niessen, K.L. Vincken et al., in *Multiscale Vessel Enhancement Filtering [C]*. International Conference on Medical Image Computing and Computer-Assisted Intervention (Springer, Berlin Heidelberg, 1998), pp. 130–137
13. M. Descoteaux, D.L. Collins, K. Siddiqi, A geometric flow for segmenting vasculature in proton-density weighted MRI. Med. Image Anal. **12**(4), 497–513 (2008)
14. Li Xiankui, *Transfer Function Design for the Volume Visualization of MRA Blood Vessels[D]* (School of Electronic, Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, 2015)
15. Chen Qianqing, *Retinal Vessel Segmentation Based on Hessian Matrix and Multiscale Analysis[D]* (Huazhong University of Science & Technology, Wuhan, 2012)
16. Pan Shuo, Yu. Zhang, Wang Kai et al., Automatic 3D lung nodule detection based on enhanced vessel segmentation. Comput. Appl. Softw. **31**(5), 206–209 (2014)
17. Ren Hongyu, *Near Infrared Vein Image Feature Enhancement and Visualization Algorithm [D]* (School of Mechatronics and Information System, China University of Mining and Technology, Beijing, 2014)
18. Yu. Chengbo, Qin Huafeng, Cui Yanzhe et al., Finger-vein image recognition combining modified Hausdorff distance with minutiae feature matching. Interdisc. Sci.: Comput. Life Sci. **1**(4), 280–289 (2009)

A Review of Model Based Online Identification Methods for Robotic Systems

Zhijing Li, Minghao Wang, Jinhua Ye and Haibin Wu

Abstract With the increasing applications of robots in complex environment, the dynamic behaviors of the unknown environment are the major factors which affect the robot system's performance. In robot community, one of the most efficient methods is reasonably modeling the environment to achieve robot force control or safety. Meanwhile, an efficient online parameters estimation method is required in the robot system. Since the parameters of the analytical models are simple and easy to use mathematical formula to describe, a review and comparison of the online identification methods are presented in this paper, according to the linear and nonlinear model type of the analytical models. Moreover, we summarized their application fields and their advantages and disadvantages, and then made a conclusion and the future scope.

Keywords Unknown environment · Online identification method · Robotic system

1 Introduction

In the past decades, industrial robot tasks are usually performed in unconstrained or constrained environment. Traditional methods used to improve the robot system are kinds of compliant devices or advanced algorithms, but the application of those methods are limited to special occasions [1]. With the more applications of robots in complex environment, the working space may include linear, nonlinear or

Z. Li · M. Wang · J. Ye (✉) · H. Wu (✉)
School of Mechanical Engineering and Automation, Fuzhou University,
2 Xueyuan Road, Minhou, Fuzhou 350116, Fujian, China
e-mail: yejinhua@fzu.edu.cn

H. Wu
e-mail: wuhb@fzu.edu.cn

Z. Li
e-mail: lizhijingwei@163.com

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_7

viscoelastic characteristic objects, and it makes the robot hard to be controlled. Moreover, robotics technology is expanding from industrial to service domain. For example, nursing robot is an important research aspect of improving aging people's life. Besides, human-robot interaction control is a challenge field. For example, the coworker robot needs to do the work and sharing work space with workers simultaneously [2], thus the safety of the worker is very important. In addition to industrial and service robots, the medical robot has become a very active research topic in the recent decades, and modeling the soft tissues has become a very important method to adding force feedback control in medical application [3]. The widely used Da Vinci surgical robot does not provide force feedback to the operator [4]. Thus several research groups have studied this field to provide realistic force feedback for surgical applications.

With the development of robot technology, such as elderly care, human-robot cooperation, medical assistant, environment is usually unknown to robots and existing many uncertainties [5]. On the other hand, estimation method has been shown to have much potential to play an important role in robot teleoperation system [6]. To attain those goals, a key component that is able to accurate estimating the properties of the contact environment is needed. Since the environment parameters can change slowly or fast when the robot is working, the robot control system has to track the change of the environment property in time. What's more, a few authors have studied more complex problem which combining the estimation of geometrical and property of the environment. However, this approach will increase the difficulty of the estimation task, and a few methods have been proposed to estimate the environment stiffness with force data. Studies show that modeling and estimating the property of the unstructured environment is an efficient method to improve the performance of robot system.

This paper is structured as follows. We review the existed environment model in robot force control in Sect. 2. Moreover, the theories of the published models are also compared. In Sect. 3, according to the different application areas, the theories and efficiencies of the proposed online parameter identification algorithms are compared and discussed. Finally, conclusions and future works are addressed in Sect. 4.

2 Unknown Environment Modeling

Modeling of unknown environment has been researched for many years. Based on the complexity of environment and feature of the model, multiple researchers have put these models into different categories. At the same time, different models have shown their limitations in different applications and no one model can perfectly describe the property of the unknown environment. Therefore, one possible solution to realize accuracy force feedback is through appropriate modeling based on the properties of the contact environment.

2.1 Category of Model

When the robot working in a constrained environment, contact or impact is inevitable and this is a complex phenomenon according to the literature [7]. Gilardi and Shart reviewed the contact dynamic model and proposed discrete model and continuous model from the energy view-point. The discrete models are based on the assumption that impact time is short and the bodies involved in the impact are mainly rigid, and it limited the model's application. The continuous models, also referred to as compliant contact models, overcome the problems associated with the discrete models [7]. Misra etc. reviewed the computer based surgical simulation in tool-tissue modeling, and broadly classified tool-tissue model as linear elasticity based model, nonlinear elasticity based Finite Element Model, and other techniques based on model [8]. Takacs etc. also reviewed the tool-tissue interaction model for robotics surgery applications [9]. Similarity, Famaey and Sloten provided a review of different continuum mechanical models which aim at facilitating model choice for specific soft tissue modeling applications [10]. In addition, there are three categories of deformation models mentioned in the paper, which are heuristic models, continuums mechanical models and hybrid models. Moreover, Fung proposed the most complete studies of the human tissue mechanical properties in his monograph [11].

Based on different environments and applications, the models are put into different categories by the researchers. A summary of environment model types is listed in Table 1, and more details can be found in corresponding references.

From above we can know that the contact models are usually adopted in industrial control tasks. Tool-tissue model or soft tissue model are always used in medical robots, teleoperation systems and computers based surgical simulator for

Table 1 A summary of environment model types

Operation	Environment modeling type	Literatures
Impact and contact phenomenon	Discrete and continuous models	[7]
Computer-based surgical simulation	Linear elasticity-based, nonlinear (hyperelastic) elasticity-based finite methods, and other techniques	[8]
Soft tissue modeling	Heuristic models, continuum-mechanical models and hybrid models (Hyperelasticity, viscoelasticity, anisotropy, heterogeneity, and damage model)	[9, 10]
Deformation models in surgical simulation	The finite element method and the mass-spring-damper method	[12]
Viscoelastic model based force control	Finite element method and analytical models	[13]

training. Since the increasing applications of the physical human-robot interaction, there is a research that focuses on improving the robot intelligent to identify the unknown environment.

From the structure of the model, the models can be divided into two main groups, Finite Element Method (FEM) and Mass-Spring-Damper Method (MSD) [12]. The FEM is a popular technology and is widely used in computer based surgical simulation. In FEM, the object model is divided into numerous of discrete elements forming a mesh to obtain the accuracy result. However, significant disadvantages of this approach are the vast computational requirements and the influence of strict boundary conditions. These disadvantages make a challenge to use FEM in real time simulations and applications. In the opinion of Moreira, the analytical model also can be classified into three categories, linear elastic models, nonlinear elastic models and viscoelastic models [13]. As the parameters of the analytical model are simple and have explicit mathematical expressions, it's widely used in online identification applications. Therefore, this paper mainly focuses on analytical model which contains spring-damping models and polynomial models.

2.2 The Application of Analytical Model

In the early research works, modeling the unknown environment as elastic spring model is apparently inaccurate. Then the Hertz model is proposed. This is a non-linear model but limited to impact with elastic deformation and its original form does not include damping [14]. The spring-dashpot model is schematically represented with a linear damper for the dissipation of energy in parallel with a linear spring for the elastic behavior [15]. To overcome the problems of the spring-dashpot model and retain the advantage of the Hertz's model, an alternative model for energy dissipation was introduced by Hunt and Crossley [16]. Since the milestone of the impedance control proposed, the property of the environment is integrated in the robot control system. Hogan considered the classical linear time invariant environment as Kelvin-Voigt model [17], therefore the impedance control can compound a dynamic balance between the robot and the environment. Diolaiti etc. have compared the Kelvin-Voigt model and Hunt-Crossley model in the aspect of energy dissipation effects and consistency, and showed that Hunt-Crossley model has more advantages than Kelvin-Voigt model [18]. Behavior of a material described by Kelvin-Voigt and Hunt-Crossley model as showed in Fig. 1 and Fig. 2. Mallapragada etc. modelled the environment as an auto regressive exogenous (ARX) to achieve high performance force control [19].

On the other hand, Yamamoto etc. compared seven possible mathematical models to describe a phantom tissue [6]. The experiment results demonstrated that Hunt-Crossley model is a better choice which provides the minimum force estimation errors in both self-validation and cross-validation and shows a consistent

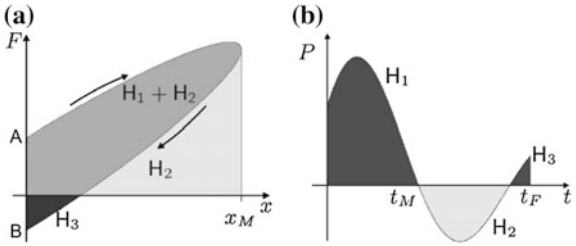


Fig. 1 Behavior of the Kelvin-Voigt model. **a** Hysteresis loop **b** Power exchange [18]

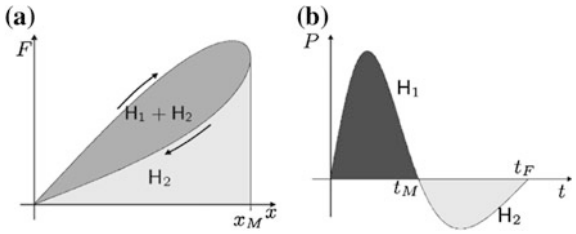
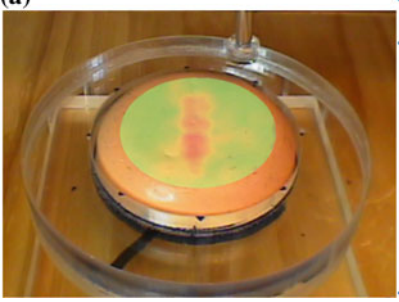


Fig. 2 Behavior of the Hunt-Crossley model. **a** Hysteresis loop **b** Power exchange [18]



(b)

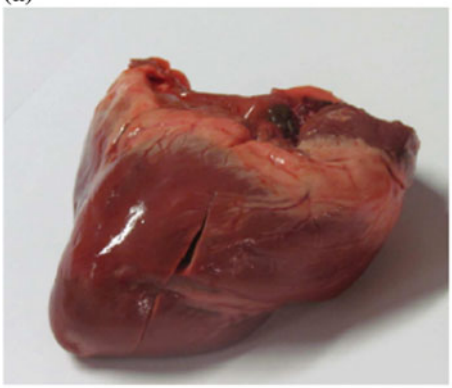
Model	Equation of Model
Kelvin-Voigt	$\hat{f} = \hat{x}_0 + \hat{k}x + \hat{b}\dot{x}$
Mass-damper-spring	$\hat{f} = \hat{x}_0 + \hat{k}x + \hat{b}\dot{x} + \hat{m}\ddot{x}$
Hunt-Crossley	$\hat{f} = \hat{x}_0 + \hat{k}x^{\hat{n}} + \hat{b}x^{\hat{a}}\dot{x}$
2nd order polynomial	$\hat{f} = \hat{a}_0 + \hat{a}_1 x + \hat{a}_2 x^2$
3nd order polynomial	$\hat{f} = \hat{a}_0 + \hat{a}_1 x + \hat{a}_2 x^2 + \hat{a}_3 x^3$
4nd order polynomial	$\hat{f} = \hat{a}_0 + \hat{a}_1 x + \hat{a}_2 x^2 + \hat{a}_3 x^3 + \hat{a}_4 x^4$
2nd order polynomial + velocity-dependent	$\hat{f} = \hat{a}_0 + \hat{a}_1 x + \hat{a}_2 x^2 + \hat{\beta}\dot{x}$

Fig. 3 Model validation test. **a** Phantom tissue with stiff object **b** Model Candidates [6]

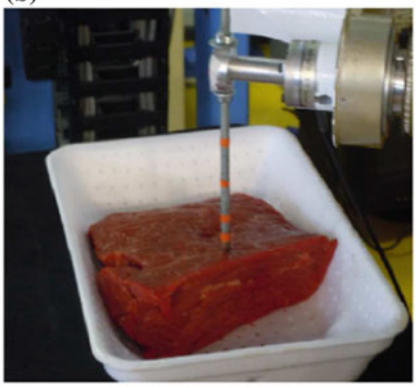
difference between the phantom tissue and the phantom calcified artery. The model validation test as shown in Fig. 3.

A group of six candidate models has been used by Moreira etc. to model the complex behavior of the real soft tissues [13]. The pig heart and beef tissue are taken by the researchers to obtain more reality result as shown in Fig. 4. The relaxation testing results as showed in Fig. 5. Comparing the works of Yamamoto and Moreira, we can know that although various form of viscoelastic models have been developed, only a few have shown a good balance between model mathematical simplicity and experimental fit accuracy.

(a)



(b)



(c)

Model	Equation of Model
Elastic	$f(t) = k \cdot x(t)$
Kelvin-Voigt	$f(t) = k \cdot x(t) + b\dot{x}(t)$
Kelvin-Boltzmann	$f(t) = \beta \cdot x(t) + \alpha\dot{x}(t) - \gamma\dot{f}(t)$
Maxwell	$f(t) = k \cdot x(t) + \alpha\dot{x}(t)$
Hunt-Crossley	$f(t) = k \cdot x^n(t) + \lambda \cdot x^n(t) \cdot \dot{x}$
Fractionl	$f(t) = G \cdot d^r x(t) / dt^r$

Fig. 4 Tissue validation test. **a** Pig heart **b** Beef tissue **c** Model Candidates [6]

Recently, Wang etc. studied the model accuracy of the viscoelastic model, and a list of more accuracy model presented and compared. Based on the standard linear solid model (or Kelvin Boltzmann model), a family of five-element models are constructed, and a double Maxwell-arm Wiechert representative model showing the balance among mathematical simplicity, mathematical loading accuracy and experimental performance. Although the seven or higher element viscoelastic models are more accuracy, but it leads to much more calculations and the more complicated mathematical formula. The theory and details are contained in literature [20]. Takacs etc. also proposed a novel nonlinear mass-spring-damper viscoelastic model and give an explicit mathematical expression [21]. However, the models proposed by Wang and Takacs are more complicate than previous, and have more parameters need to estimate making it more difficult to use in robot force control tasks. As a result, the above mentioned disadvantage limits them in practical applications. A list of the structure of the damping-spring models are shown in Fig. 6. Finite element method and more complicated model are not considered in this paper, because of they are not suitable to use in the real time application scenario.

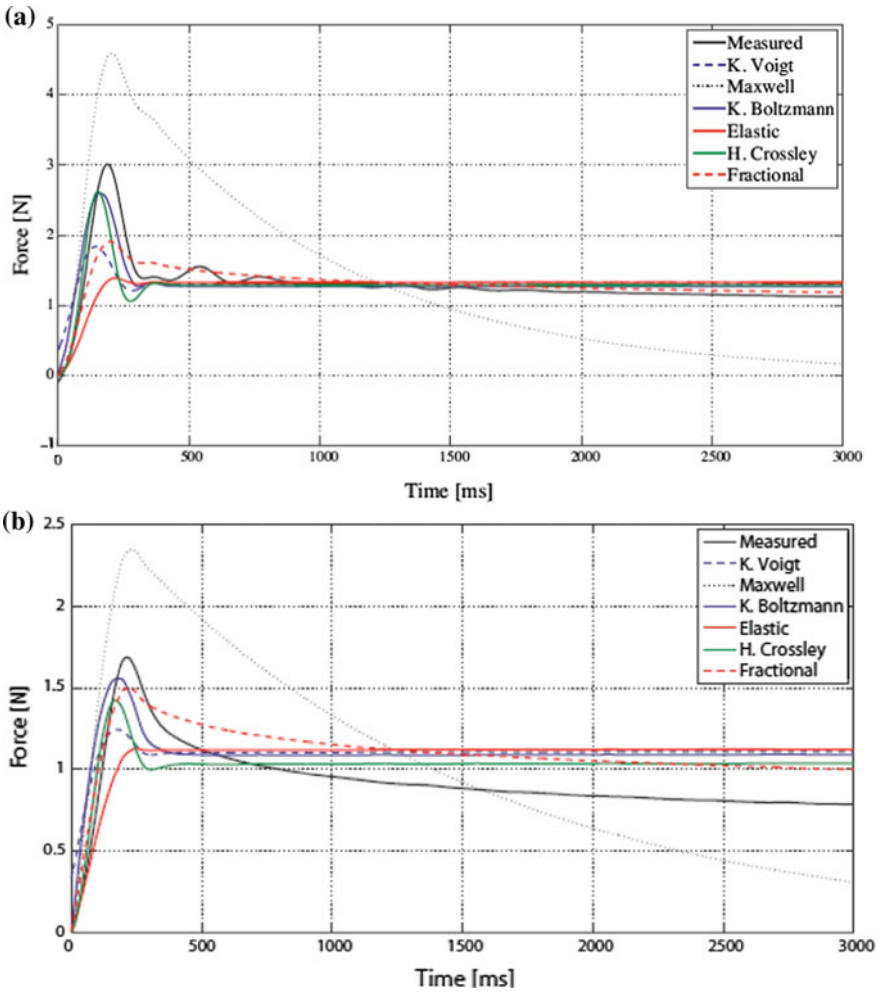


Fig. 5 Measured and estimated forces in relaxation test **a** Pig heat **b** Beef tissue [13]

3 Online Parameter Identification Method

From the modeling and control point of view, the key problem is using an identification algorithm to obtain environment parameters online. Besides, some challenges have been met in unknown environment, and most works take single point contact and simple geometrical deformation consideration. In an intricate scenario, when the robot and human work in the same space, the stability, robustness and bandwidth are the very important features of the control system for safe human-robot interaction safety. If the robot system can online identify the contact object's stiffness and damping, robot force control would be more feasible. From

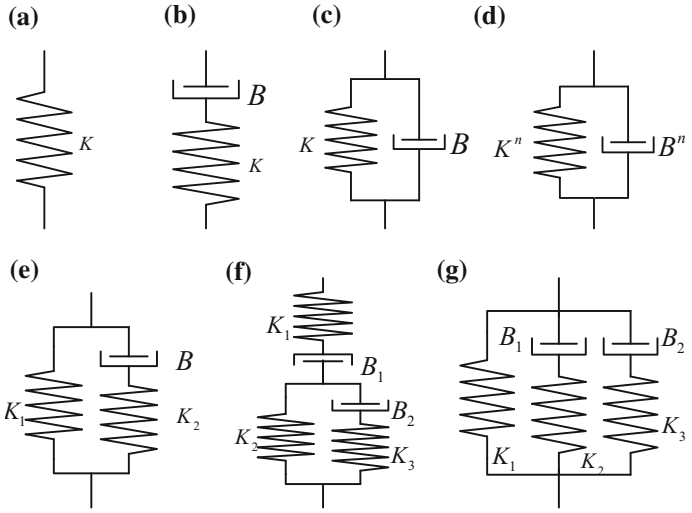
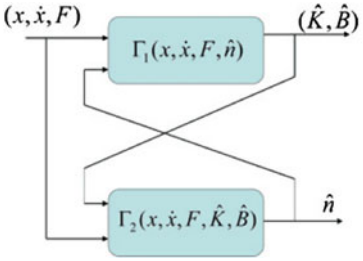


Fig. 6 The structure of the damping-Spring models: **a** Elastic model **b** Maxwell model **c** Kelvin-Voigt model **d** Hunt-Crossley model **e** Kelvin Boltzmann model **f** Maxwell-Kelvin model **g** Wiechert model

this point of view, some methods are proposed. Initial motivation to identify the environment is to improve contact stability and force control, such as impedance control proposed by Hogan [17]. Love and Book used a MIMO recursive identification to improve the robot force control performance both in unconstrained and constrained environment [22]. Seraji and Colbaugh proposed an indirect adaptive control to achieve force control within impedance control [23]. Model reference adaptive control was used by Singh and Popa to invest the robot force control problem [24]. In the literature [25], Erickson etc. review and compared above three algorithms and proposed a signal processing method for identification the contact environment stiffness and damping in constrained motion control. All of the identification algorithms are based on impedance control.

As Diolaiti etc. proved that Hunt-Crossley model has better performance, they proposed a two-stage method to online estimate the model's parameters, and the framework of the algorithm as shown in Fig. 7 [18]. Then Haddadi etc. found the former method has some defects and proposed a new algorithm called one stage identification method [26]. Moreira etc. find Hunt-Crossley and Kelvin Boltzmann models have better consistency with the tissue property than others [13]. While Kelvin Boltzmann model is a linear model, the Hunt-Crossley is nonlinear model, and the former is easier to be used in controllers design. Besides, a novel force control scheme using Active Observer based on Kelvin Boltzmann model is presented, but only one parameter online estimation in the scheme [27]. In addition, Mallapragada etc. model the environment as an Auto Regressive eXogenous (ARX) and used the literature [22] proposed method to obtain the model parameters [19].



$$\mathbf{L}_{k+1} = \frac{\mathbf{P}_k \phi_{k+1}}{\lambda + \phi_{k+1}^T \mathbf{P}_k \phi_{k+1}}$$

$$\mathbf{P}_{k+1} = \frac{1}{\lambda}[\mathbf{P}_k - \mathbf{L}_{k+1} \phi_{k+1}^T \mathbf{P}_k]$$

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \mathbf{L}_{k+1}[F_{k+1} - \phi_{k+1}^T \hat{\theta}_k]$$

Fig. 7 The two-stage identification method for Hunt-Crossley model [18]

From the upon content, we can know that the online identification are most based on Kelvin-Voigt, Hunt-Crossley, and Kelvin Boltzmann's model. Since Recursive Least Squares Method is simple and easy to use, it has been widely employed in online identification method, such as in literatures [18, 22, 25, 26, 28, 29]. In addition, Haddadi and Wang etc. have studied and compared several improved recursive least squares method and proposed a SD-based block least mean squares online identification method [30–32]. Yamamoto also compared the estimation performance of three online estimation techniques which are recursive least-squares, adaptive identification and multi-estimator [6] [33]. Furthermore, Kalman Filter, Neural Network, and Bayesian inference estimation method are also used in parameters identification of environment [34–36].

In complex control systems, unknown disturbance, higher order dynamics, nonlinearities and noise are always present. To deal with those problems, Cortesao introduces the Active Observer (AOB) algorithm in the framework of Kalman Filter [37]. Figure 8 shows the AOB algorithm. A novel method proposed by Coutinho and Cortesao, which contains two force observers to achieve online stiffness estimation in the paper [27]. Then Coutinho and Cortesao proposed to estimate environment stiffness only use force data based on the AOB or Candidate Observer Based Algorithm (COBA) [38–40]. Similarly, the state-space Kalman Filter technology is adopted by Zarrad etc. to estimate the stiffness and geometry of the environment [41]. A summary of model based online parameter estimation methods for unknown environment is listed in Table 2, and the details can be found in corresponding references.

Online parameters estimation technique is useful for improving force control in explicit force control scheme. Numerous applications can be found in industrial, teleoperation and medical robotics. However, most of these methods consider only geometrically simple and single point contact situation. With the development of the robot technology, several new methods are proposed to deal with complex multi-point, unstructured and unknown environment. For example, multi-point contact and geometrical uncertainties scenarios is developed and investigated by Werber etc., but the proposed method is off-line least-square based [42]. Verscheure etc. also have investigated the identification of contact parameters from multi-point contact robotic operations, and extended the former off-line identification method to deal with small geometrical uncertainties [43].

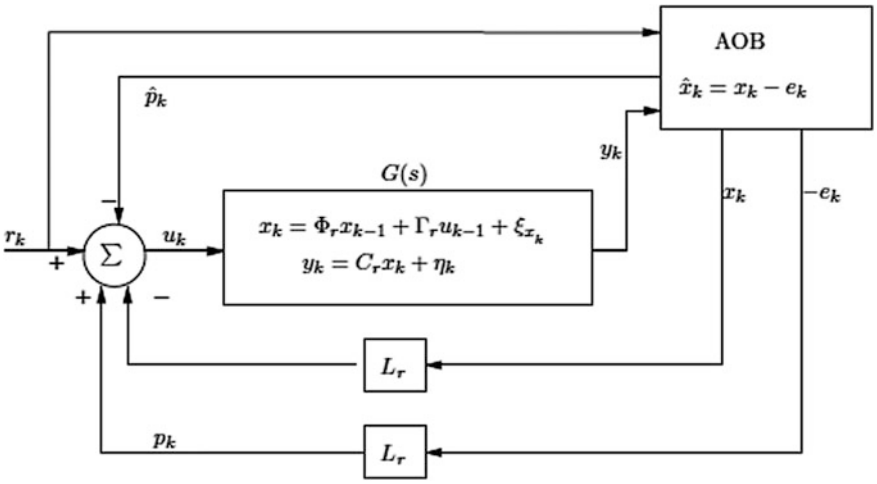


Fig. 8 The Kalman Active Observer (AOB) [37]

Table 2 A summary of model based on online parameter estimation methods and its corresponding applications

Analytical model	Applying environment	Estimation algorithm	Literatures
Spring model	haptic feedback control or force control	Kalman filter	[44]
Kelvin-Voigt	Unconstrained to constrained or force control	Recursive least squares method (RLS)	[18, 22, 25, 26, 28–31]
Kelvin-Voigt	Explicit force control	Model reference adaptive control	[24, 25]
Kelvin-Voigt	Force tracking within impedance control	(Indirect) adaptive control	[23, 25]
Kelvin-Voigt	Impedance control	Signal processing method	[25]
Kelvin-Voigt	Telemanipulation and surgical	Multi-estimator	[33]
Kelvin-Voigt	Deformation force tracking	(Extended) Kalman filter method	[35]
Kelvin-Voigt	Linear time-varying system	Block least mean Pth method	[30]
Kelvin-Voigt	Impedance control	Artificial Neural Network	[45]
Hunt-Crossley	Contact task control	Two stage method (Double RLS method)	[30]
Hunt-Crossley	Contact task control	Signal-stage method	[26]

(continued)

Table 2 (continued)

Hunt-Crossley	Viscoelastic environment	Neural network feedforward method	[34]
Kelvin Boltzmann	Medical or surgical application	Active observer	[27]
Fractional derivative equation model	Robot-assisted surgery	Active observer	[46]
Polynomial model (Second order or higher)	Teleoperation control or surgical application	Adaptive control	[6, 47]
Auto regressive Exogenous model (ARX)	Force control for unknown environment	Weight recursive least square method	[19]

4 Conclusions

Model based force control is one of the most efficient methods in unknown environment robot applications. Considering the characteristics of the FEM and analytical model, the latter is more suitable to be used in the environment modeling to realize online algorithm design. Existing research results show that there is no one model can present the properties of the environment perfectly, but we can choose a suitable model to complete robot tasks as accurate as possible according to different situation. At the same time, online parameter estimation can greatly improve the control performance of the robot system. Therefore, we should balance the model deformation accuracy and computation in practice applications. Existing researches show that, the proposed online identification methods are most based on Kelvin-Voigt, Hunt-Crossley, and Kelvin Boltzmann model. There the more parameters in the model, the more difficult to estimate online. Thus there is no method to applying in the model which the parameters more than five. Consequently, robot force control is facing more challenges, and developing more efficient online parameter estimation algorithm for three and five analytical model to use in the physical human-robot field is the future work.

This paper is mainly focused on analytical model based online parameters estimation, since it is more practical in applications. However, there are some problems in the estimation processing. On the one hand, most estimation methods only take single contact point and one dimension into consideration. In a real environment, the contact situation is more complicated. On the other hand, the approach to combine the contact parameters estimation with geometrical parameters estimation is a challenging work, because it makes the algorithm more complex and cannot be online estimate yet. Moreover, the property of the environment may change on time dramatically. Thus the fast convergence, stability and robustness of the identification method are very important to achieve the goal of robot force control. As a rule of thumb, the above mentioned defects of the identification methods need to be solved in the future research.

Acknowledgements This study is supported by the National Natural Science Foundation of China (No. 51175084, 51575111 and 51605093), and University-Enterprise Cooperation Project (2015H6012).

References

1. T. Brogardh, Present and future robot control development-An industrial perspective. Annu. Rev. Control **31**(1), 69–79 (2007)
2. M. Beetz, G. Bartels, A. Albu-Schaffer, in *Robotic Agents Capable of Natural and Safe Physical Interaction with Human Co-workers*, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (2015), pp. 6528–6535
3. T. Yamamoto, M. Bernhardt, A. Peer etc., in *Tissue Property Estimation and Graphical Display for Teleoperated Robot-assisted Surgery*, IEEE International Conference on Robotics and automation (2009), pp. 4239–4245
4. H.C. Ott, J.O. Bonatti, L. Müller, etc., Computer-enhanced "robotic" cardiac surgery: experience in 148 patients. J. Thorac. Cardiovasc. Surg. **121**(5), 842–853 (2001)
5. S. Sam Ge, Y. Li, C. Wang, Impedance adaptation for optimal robot-environment interaction. Int. J. Control **87**(2), 249–263 (2014)
6. T. Yamamoto, B. Vagvolgyi, K. Balaji, etc., in *Tissue Property Estimation and Graphical Display for Teleoperated Robot-assisted Surgery*, IEEE International Conference on Robotics and Automation (ICRA '09) (2009), pp. 4239–4245
7. G. Gilardi, I. Sharf, Literature survey of contact dynamics modeling. Mech. Mach. Theory **37** (10), 1213–1239 (2002)
8. S. Misra, K.T. Ramesh, A.M. Okamura, Modeling of tool-tissue interactions for computer-based surgical simulation: A literature review. Presence **17**(5), 463–491 (2008)
9. A. Takacs, S. Jordan, R.-E. Precup etc., in *Review of Tool-tissue Interaction Models for Robotics Surgery Applications*, 12th IEEE International Symposium on Applied Machine Intelligence and Informatics (SAMI) (2014), pp. 339–344
10. N. Famaey, J.V. Sloten, Soft tissue modeling for applications in virtual surgery and surgical robotics. Comput. Methods Biomech. Biomed. Eng. **11**(4), 351–366 (2008)
11. Y. Fung, *Biomechanics Mechanical Properties of Living Tissue* (Springer-Verlag, Berlin, 1993)
12. M. Pourhosseini, V. Azimirad, M. Kazemi, A new fast nonlinear modeling of soft tissue for surgical simulation. J. Robot. Surg. **8**(2), 141–148 (2014)
13. P. Moreira, N. Zemiti, C. Liu etc., Viscoelastic model based force control for soft tissue interaction and its application in physiological motion compensation. Comput. Methods Program. Biomed. **116**(2), 52–67 (2014)
14. H. Hertz, *Miscellaneous Papers* (Jones and Schott, London, 1896)
15. W. Goldsmith, *Impact: The theory and physical behavior of colliding solids* (Edward Arnold Publishers Ltd, London, 1960)
16. K.H. Hunt, F.R.E. Crossley, Coefficient of restitution interpreted as damping in vibroimpact. J. Appl. Mech. **42** Series E, 369–376 (1975)
17. N. Hogan, Impedance control: an approach to manipulation. Parts I-III. J. dyn. Syst. Measur. Control. **107**, 1–24 (1985)
18. N. Diolaiti, C. Melchiorri, S. Stramigioli, Contact impedance estimation for robotic systems. IEEE Trans. Rob. **21**(5), 925–935 (2005)
19. V. Mallapragada, D. Erol, N. Sarkar, A new method of force for unknown environment. Int. J. Adv. Rob. Syst. **4**(3), 313–322 (2007)
20. X. Wang, J.A. Schoen, M.E. Rentschler, A quantitative comparison of soft tissue compressive viscoelastic model accuracy. J. Mech. Behav. Biomed. Mater. **20**, 126–136 (2013)

21. A. Takacs, I.J. Rudas, T. Haidegger, Surface deformation and reaction force estimation of liver tissue based on a novel nonlinear mass-spring-damper viscoelastic model. Med. Biol. Eng. 1–10 (2015)
22. L. love and W. Book, in *Environment Estimation for Enhanced Impedance Control*, Proceedings of the IEEE International Conference on Robotics and Automation (Nagoya, Japan, 1995)
23. H. Seraji, R. Colbaugh, Force tracking in impedance control. Int. J. Robot. Res. **16**(1), 97–117 (1997)
24. S.K. Singh, D.O. Popa, An analysis of some fundamental problems in adaptive control of force and impedance behavior: theory and experiments. IEEE Trans. Robot. Autom. **11**(1), 912–921 (1995)
25. D. Erickson, M. Weber, I. Sharf, Contact stiffness and damping estimation for robotics systems. Int. J.Robot. Res. **22**(1), 41–57 (2003)
26. A Haddadi, K. Hashtrudi-Zaad, in *Online Contact Impedance Identification for Robotic Systems*, IEEE International Conference on Intelligent Robot and Systems (2008), pp. 974–980
27. F. Coutinho, R. Cortesão, Online stiffness estimation for robotic tasks with force observer. Control Eng. Pract. **24**, 92–105 (2014)
28. B.S. Ye, B. Song, Z.Y. Li etc., A study of force and position tracking control for robot contact with an arbitrarily inclined plane Int. J. Adv. Robot. Syst. **10**(69), 126–136 (2013)
29. Z.Y. Li, H.M. Cao, Robot impedance control method adapting to unknown or changing environment stiffness and damping parameters. Chin. Mech. Eng. **12**, 1581–1585(Chinese) (2014)
30. A Haddadi, K. Hashtrudi-Zaad, in *Online Contact Impedance Identification for Robbotic Systems*, IEEE International Conference on Intelligent Robots and Systems (2008), pp. 974–1980
31. A Haddadi, K. Hashtrudi-Zaad, in *A New Fast Online Identification Method for Linear Time-varying Systems*, American Control Conference (2008), pp. 1322–1328
32. Z. Wang, A. Peer, M. Buss, in *Fast Online Impedance Estimation for Robot Control*, Proceeding of the 2009 IEEE International Conference on Mechanics (2009), pp. 1–6
33. T. Yamamoto, M. Bernhardt, A. Peer etc., in *Techniques for Environment Parameter Estimation During Telemanipulation*, Proceedings of the 2nd biennial IEEE international conference on biomedical robotics and biomechatronics (2008), pp. 217–223
34. S. Bhasin, K. Dupree, P.M. Patre etc., in *Neural Network Control of a Robot Interacting an Uncertain Hunt-Crossley Viscoelastic Environment*, ASME Dynamic Systems and Control Conference (2008), pp. 875–882
35. L. Roveda, F. Vicentini, L.M. Tosatti, in *Deformation-tracking Impedance Control in Interaction with Uncertain Environments*, IEEE International Conference on Intelligent Robots and Systems (2013), pp. 1992–1997
36. J.D. Gea, F. Kirchner, in *Contact Impedance Adaptation Via Environment Identification*, IEEE International Symposium on Industrial Electronics (2008), pp. 1365–1369
37. R. Cortesao, On Kalman Active Observers. J. Intell. Robot Syst. **48**, 131–155 (2007)
38. R. Cortesao, C. Sousa, P. Queiros, in *Active Impedance Design for Human-robot Comanipulation*, American control conference (2010), pp. 2805–2810
39. F. Coutinho, R. Cortesao, in *Environment Stiffness Estimation with Multiple Observers*, 35th Annual Conference of IEEE Industrial Electronics (2009), pp. 1537–1542
40. F. Coutinho, R. Cortesao, in *System Stiffness Estimation with the Candidate Observers Algorithm*, 18th Mediterranean Conference on Control and Automation (2010), pp. 796–801
41. W. Zarrad, P. Poignet, R. Cortesao, in *Haptic Feedback Control in Medical Robotics Trough Stiffness Estimation with Kalman Filter*, In Proceedings of the IEEE International Conference on Advanced Robotics (2007), pp. 81–87
42. M. Weber, K. Patel, O. ma, etc., Identification of contact dynamics model parameters from constrained robotic operations. J. Dyn. Syst. Meas. Contr. **128**(2), 307–318 (2006)

43. D. Verschure, I. Sharf, H. Bruyninckx etc., Identification of contact parameters from stiff multipoint contact robotic operation. Int. J. Robot. Res. **29**(4), 367–385 (2010)
44. C. Liang, S. Bhasin, K. Dupree, W.E. Dixon, in *A Force Limiting Adaptive Controller for a Robotic System Undergoing a Non-contact to Contact Transition*, 46th IEEE Conference on Decision and Control (2007), pp. 3555–3560
45. H. Wang, K.H. Low, M.Y. Wang, in *Reference Trajectory Generation for Force Tracking Impedance Control by Using Neural Network-based Environment Estimation*, IEEE conference on robotics, automation and mechatronics (2006), pp. 1–6
46. Y. Kobayashi, P. Moreira, C. Liu, etc., in *Haptic Feedback Control in Medical Robots Through Fractional Viscoelastic Tissue Model*, 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society,EMBC (2011), pp. 6704–6708
47. A.M. Okamura, L.N. Verner, C.E. Reiley, etc., in *Haptics for Robot-assisted Minimally Invasive Surgery*, 13th International Symposium of Robotic Research (2007), pp. 26–29

A General Kinetostatic Model Based Stiffness Estimation for Tripod Parallel Kinematic Machines with Prismatic Actuators

Jun Zhang and Tengfei Tang

Abstract A general stiffness modeling methodology for tripod parallel kinematic machines (PKMs) with prismatic actuators is proposed in this paper. With the technique of substructure synthesis, the whole system of a tripod is divided into a platform, a base and three kinematic limb substructures. Each limb assemblage is modeled as a spatial beam constrained by two sets of six degree-of-freedom (6-DOF) virtual lumped springs with equivalent stiffness at their geometric centers. The equilibrium equation of each individual limb assemblage is derived through finite element formulation, while that of the platform is derived with the Newton's 2nd law. The governing stiffness matrix is synthesized by introducing the deformation compatibility conditions between the platform and the limbs. By extracting a 6x6 block matrix from the inversion of the governing compliance matrix, a stiffness matrix of the platform is formulated. Taking the Sprint Z3 Head and the A3 Head as examples, the distributions of stiffness values of these two types of PKM modules are predicted and discussed. It is worth mentioning that the proposed methodology of stiffness modeling can further be applied to other types of PKMs for evaluating the global rigidity performance over entire workplace efficiently with minor revisions.

Keywords Parallel kinematic machine · Tripod · Kinetostatic · Stiffness modeling · Substructure synthesis

1 Introduction

Thanks to the merits of better accuracy and higher rigidity, the tripod parallel kinematic machines (PKMs) with prismatic actuators have been proved as a promising alternative solution for high speed machining (HSM) tasks on extra large

J. Zhang (✉)
School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou, China
e-mail: zhang_jun@fzu.edu.cn

T. Tang
School of Mechanical Engineering, Anhui University of Technology, Ma'anshan, China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_8

scale components with complex geometries. As a successful example, the Sprint Z3 head has been commercially applied in the aeronautical industries [1, 2]. Another newly invented and commercially applied tripod PKM is the Exechon, with an over-constrained 2UPR/1SPR topological architecture [3, 4]. Herein, 'R' and 'S' denote a revolute joint and a spherical joint respectively while 'P' represents an active prismatic joint. Inspired by the 3-PRS topology of the Sprint Z3 head, a similar 3-RPS tripod-based PKM named the A3 head was proposed as a multiple-axis spindle head to form a hybrid 5-axis high-speed machining unit [5, 6]. Other investigations on the tripod PKMs can also be traced in recent publications [7–11].

In the early design stage for the tripod PKMs that are designed for manipulation with high rigidity and high positioning accuracy requirements, stiffness is one of the most overwhelming concerns. However, due to the complex kinematics and structural features, the derivation for the stiffness matrix of such kinds of PKM modules is nevertheless, a tough task, not mentioning the challenge of estimating stiffness throughout the workspace with accuracy and efficiency. Therefore, the estimation for rigidity performance of a tripod PKM still remains as a challenge unless a computational efficient as well as accurate stiffness modeling method is proposed.

Numerous efforts have been contributed to the stiffness modeling and estimation for various PKMs in the past decades. Among all these efforts, the finite element method (FEM) [12, 13], the matrix structure method (MSM) [14, 15], the virtual joint method (VJM) [16–18] and the screw-based method (SBM) [19–22] are the most common used approaches. For example, Pham and his co-workers [13] proposed an analytical finite element model for a flexure parallel mechanism. The analytical results were then compared with the experimental tests to validate the computation accuracy of the proposed stiffness model. As to the matrix structure method, a Jacobian-based stiffness model [15] was proposed by Bi and his co-workers. By using the lumped-parameter method, Zhang et al. [16] established a kinetostatic model for an enhanced tripod mechanism. Li and Xu [19] employed the screw theory to develop a systematic and analytical stiffness model for a family of 3-DOF parallel mechanisms with three prismatic limbs. Wang et al. [22] presented a semi-analytical approach to investigate the stiffness of a tripod-based robot named the TriVariant-B.

It is worth noting that the above stiffness models were specially established for specific PKMs. In other words, it lacks of versatility for these stiffness models being applied to different types of tripod PKMs. Motivated by this thought, the authors aim to present a general stiffness modelling methodology for different types of tripod PKMs with prismatic actuators. To achieve an acceptable balance between the computational efficiency and accuracy, a kinetostatic model that considers the compliances of both limbs and joints is adopted in the present study. For this purpose, the limbs are modeled as spatial beams with corresponding cross-sections constrained by passive joints which are simplified as lumped virtual spring units with equivalent stiffness coefficients.

The reminder of the paper is organized as follows. In Sect. 2, a general kinetostatic stiffness model is established to yield an analytical formulation of the platform's stiffness for the tripod PKMs. In Sect. 3, a general algorithm principle of numerical simulation is proposed to estimate the stiffness performance of two types of tripod PKMs. The stiffness mapping of the two typical tripods are predicted and discussed in details. Finally, some conclusions and remarks are drawn in Sect. 4.

2 Stiffness Modeling Methodology

2.1 Kinematic Definitions

Figure 1 shows the schematic diagram of a general tripod PKM consisting of a platform, a base and three prismatic actuated kinematic limbs. Each limb connects the platform to the base through two passive joints whose geometric centers are denoted as A_i and B_i (i=1, 2, 3), respectively.

As depicted in Fig. 1, the following Cartesian coordinate systems are defined: a global coordinate system B-xyz is attached to the base with its origin B being recommended to set at the geometric center of the base; similarly, a body-fixed coordinate system A-uvw is defined at the platform; and three body-fixed limb reference frames B_i-$x_iy_iz_i$ are arranged at the geometric centers of passive joints B_i. For the convenient of formulation, let z_i point in the direction of vector from B_i to A_i.

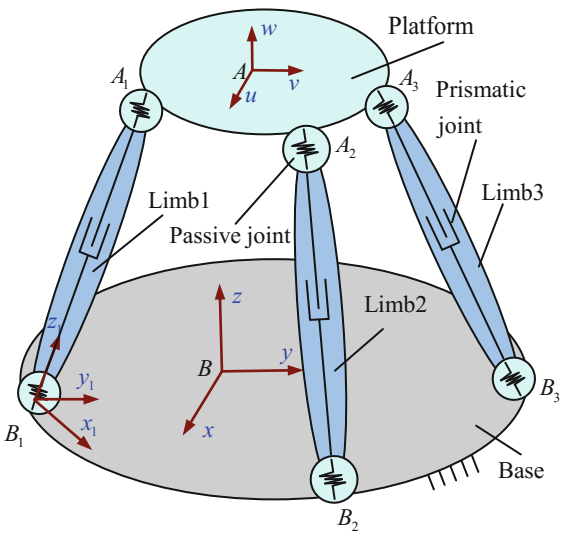


Fig. 1 Schematic diagram of a general tripod PKM

Assume the transformation matrix of A-uvw with respect to B-xyz is ${}^{B}\boldsymbol{T}_{A}$

$$ {}^{B}\boldsymbol{T}_{A} : \mathrm{Trans}(A - uvw \rightarrow B - xyz) \tag{1} $$

Similarly, assume the transformation matrix of B_i-$x_iy_iz_i$ with respect to B-xyz is ${}^{B}\boldsymbol{T}_{Bi}$

$$ {}^{B}\boldsymbol{T}_{Bi} : \mathrm{Trans}(B_i - x_iy_iz_i \rightarrow B - xyz) \tag{2} $$

2.2 Finite Element Formulation of the Limb Assemblage

According to the kinematic motion of limbs in a tripod PKM, they can be roughly classified into two categories, i.e., Case A and Case B.

Figure 2 shows the assembling scheme of Case A where the limb length between the two passive joints is a constant value when the limb undergoes kinematic motions.

In this case, the limb body is constrained by the two passive joints at A_i and B_i through the front bearing and the rear bearing, respectively. The passive joints in the limb assemblage are simplified into two sets of 6-DOF lumped virtual spring units with equivalent linear/angular stiffness constants denoting as $\boldsymbol{k}_{A1i}/\boldsymbol{k}_{A2i}$ and $\boldsymbol{k}_{B1i}/\boldsymbol{k}_{B2i}$. These spring constants can be determined either by finite element computation or by semi-analytical analyses. With the finite element method, the limb can be meshed into finite elements with each node having three linear and three angular coordinates along and about three perpendicular axes [23].

To facilitate the formulation, assume each limb body is divided into n elements with A_i and B_i being one node of the 1st and the nth beam element, respectively. For clarity, one may denote the element nodes in the discrete spatial beam as e_{i1}, e_{i2}, …, $e_{i(n+1)}$ and define a set of nodal reference frame e_{ij}-$x_{ij}y_{ij}z_{ij}$ at the center of element e_{ij} with its three axes parallel to those in the limb frame B_i-$x_iy_iz_i$. With the boundary conditions aroused from the passive joints, a set of static equilibrium equations for the limb can be formulated in the frame of B_i-$x_iy_iz_i$ as

$$ \boldsymbol{k}_i\boldsymbol{\xi}_i = \boldsymbol{w}_i \tag{3} $$

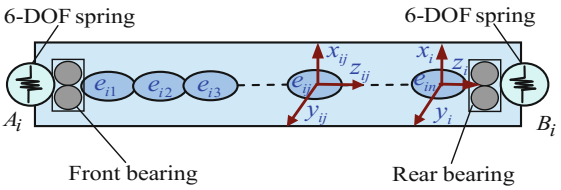


Fig. 2 Assembling scheme and finite element model of an individual limb in Case A

where $\boldsymbol{k}_i$, $\boldsymbol{\xi}_i$ and $\boldsymbol{w}_i$ are the stiffness matrix, the general coordinates and the general load vector of the ith limb body in B_i-$x_iy_iz_i$ and can be further expressed as

$$\boldsymbol{w}_i = [\boldsymbol{f}_{Ai}^T, \boldsymbol{m}_{Ai}^T, \ldots, \boldsymbol{f}_{Bi}^T, \boldsymbol{m}_{Bi}^T]^{\mathrm{T}} \tag{4}$$

where $\boldsymbol{f}_{Ai}/\boldsymbol{m}_A$ and $\boldsymbol{f}_{Bi}/\boldsymbol{m}_{Bi}$ denote the forces/moments acting at the nodes A_i and B_i measured in the B_i-$x_iy_iz_i$, respectively.

The general coordinates of an individual limb measured in the limb frame B_i-$x_iy_iz_i$ can thus be expressed as

$$\boldsymbol{\xi}_i = [\boldsymbol{\delta}_{Ai}^{\mathrm{T}}, \boldsymbol{\rho}_{Ai}^{\mathrm{T}}, \ldots, \boldsymbol{\delta}_{Bi}^{\mathrm{T}}, \boldsymbol{\rho}_{Bi}^{\mathrm{T}}]^{\mathrm{T}} \tag{5}$$

where $\boldsymbol{\delta}_{Ai}/\boldsymbol{\rho}_{Ai}$, $\boldsymbol{\delta}_{Bi}/\boldsymbol{\rho}_{Bi}$ are the linear/angular nodal coordinates of nodes A_i and B_i, respectively.

Therefore, the nodal coordinates of A_i and B_i can be referred to $\boldsymbol{\xi}_i$ by the following transformation matrices

$$\boldsymbol{\delta}_{Ai} = \boldsymbol{N}_{Bi}^{A1}\boldsymbol{\xi}_i, \quad \boldsymbol{\rho}_{Ai} = \boldsymbol{N}_{Bi}^{A2}\boldsymbol{\xi}_i \tag{6}$$

$$\boldsymbol{\delta}_{Bi} = \boldsymbol{N}_{Bi}^{B1}\boldsymbol{\xi}_i, \quad \boldsymbol{\rho}_{Bi} = \boldsymbol{N}_{Bi}^{B2}\boldsymbol{\xi}_i \tag{7}$$

$$\boldsymbol{N}_{Bi}^{A1} = [\mathbf{I}\ \mathbf{0}\ \underbrace{\mathbf{0}\ldots\mathbf{0}}_{6n}], \quad \boldsymbol{N}_{Bi}^{A2} = [\mathbf{0}\ \mathbf{I}\ \underbrace{\mathbf{0}\ldots\mathbf{0}}_{6n}] \tag{8}$$

$$\boldsymbol{N}_{Bi}^{B1} = [\underbrace{\mathbf{0}\ldots\mathbf{0}}_{6n}\ \mathbf{I}\ \mathbf{0}], \quad \boldsymbol{N}_{Bi}^{B2} = [\underbrace{\mathbf{0}\ldots\mathbf{0}}_{6n}\ \mathbf{0}\ \mathbf{I}] \tag{9}$$

where $\mathbf{0}$ and $\mathbf{I}$ denote a zero matrix and an identity matrix in 3×3, respectively.

Figure 3 shows the assembling scheme of Case B where the limb length between the two passive joints is a varying value when the limb undergoes kinematic motions.

Similar to the derivation in Case A, one may assume each limb is divided into n elements with A_i, B_i and C_i being the 1st, the $(j + 1)$th and the $(n + 1)$th nodes of the limb, respectively. Accordingly, a set of equilibrium equations of the ith limb in frame B_i-$x_iy_iz_i$ can be formulated in the matrix form as Eq. (3).

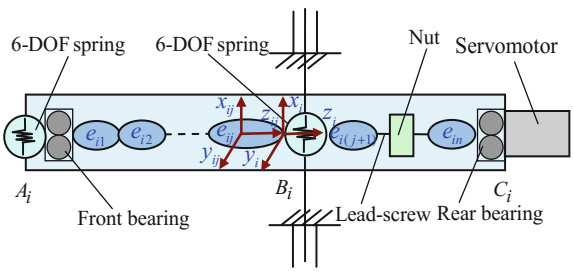


Fig. 3 Assembling scheme and finite element model of an individual limb in Case B

In this case, $\boldsymbol{w}_i$ and $\boldsymbol{\xi}_i$ can be expressed as

$$\boldsymbol{w}_i = [\boldsymbol{f}_{Ai}^{\mathrm{T}}, \boldsymbol{m}_{Ai}^{\mathrm{T}}, \ldots \boldsymbol{f}_{Bi}^{\mathrm{T}}, \boldsymbol{m}_{Bi}^{\mathrm{T}}, \ldots \boldsymbol{f}_{Ci}^{\mathrm{T}}, \boldsymbol{m}_{Ci}^{\mathrm{T}}]^{\mathrm{T}}, \quad \boldsymbol{\xi}_i = [\boldsymbol{\delta}_{Ai}^{\mathrm{T}}, \boldsymbol{\rho}_{Ai}^{\mathrm{T}}, \ldots, \boldsymbol{\delta}_{Bi}^{\mathrm{T}}, \boldsymbol{\rho}_{Bi}^{\mathrm{T}}, \ldots, \boldsymbol{\delta}_{Ci}^{\mathrm{T}}, \boldsymbol{\rho}_{Ci}^{\mathrm{T}}]^{\mathrm{T}} \tag{10}$$

where $\boldsymbol{f}_{Ci}/\boldsymbol{m}_{Ci}$ denotes the forces/moments acting on the node C_i measured in B_i-$x_iy_iz_i$. $\boldsymbol{\delta}_{Ci}/\boldsymbol{\rho}_{Ci}$ is the three linear/angular nodal coordinates of C_i. The nodal coordinates of A_i, B_i and C_i can be related to $\boldsymbol{\xi}_i$ by the following transformations

$$\boldsymbol{\delta}_{Ai} = \boldsymbol{N}_{Bi}^{A1}\boldsymbol{\xi}_i, \quad \boldsymbol{\rho}_{Ai} = \boldsymbol{N}_{Bi}^{A2}\boldsymbol{\xi}_i \tag{11}$$

$$\boldsymbol{\delta}_{Bi} = \boldsymbol{N}_{Bi}^{B1}\boldsymbol{\xi}_i, \quad \boldsymbol{\rho}_{Bi} = \boldsymbol{N}_{Bi}^{B2}\boldsymbol{\xi}_i \tag{12}$$

$$\boldsymbol{\delta}_{Ci} = \boldsymbol{N}_{Bi}^{C1}\boldsymbol{\xi}_i, \quad \boldsymbol{\rho}_{Ci} = \boldsymbol{N}_{Bi}^{C2}\boldsymbol{\xi}_i \tag{13}$$

$$N_{Bi}^{A1} = [\mathbf{I}\,\mathbf{0}\,\underbrace{\mathbf{0}\ldots\mathbf{0}}_{6n}], \quad N_{Bi}^{A2} = [\mathbf{0}\,\mathbf{I}\,\underbrace{\mathbf{0}\ldots\mathbf{0}}_{6n}] \tag{14}$$

$$\begin{aligned} \boldsymbol{N}_{Bi}^{B1} &= \left[\underbrace{\mathbf{0}\ \ldots\ \mathbf{0}}_{6j}\ \ \mathbf{I}\ \ \mathbf{0}\ \ \underbrace{\mathbf{0}\ \ldots\ \mathbf{0}}_{6(n-j)}\right], \\ \boldsymbol{N}_{Bi}^{B2} &= \left[\underbrace{\mathbf{0}\ \ldots\ \mathbf{0}}_{6j}\ \ \mathbf{0}\ \ \mathbf{I}\ \ \underbrace{\mathbf{0}\ \ldots\ \mathbf{0}}_{6(n-j)}\right] \end{aligned} \tag{15}$$

$$\boldsymbol{N}_{Bi}^{C1} = \left[\underbrace{\mathbf{0}\ \ldots\ \mathbf{0}}_{6n}\ \ \mathbf{I}\ \ \mathbf{0}\right], \quad \boldsymbol{N}_{Bi}^{C2} = \left[\underbrace{\mathbf{0}\ \ldots\ \mathbf{0}}_{6n}\ \ \mathbf{0}\ \ \mathbf{I}\right] \tag{16}$$

Equation (3) can be transformed into the global coordinate system B-xyz as

$$\boldsymbol{K}_i\boldsymbol{U}_i = \boldsymbol{W}_i \tag{17}$$

where $\boldsymbol{K}_i$, $\boldsymbol{U}_i$ and $\boldsymbol{W}_i$ are the stiffness matrix, the general coordinates vector and the external loads vector of limb i measured in B-xyz. And there exist

$$\boldsymbol{K}_i = \boldsymbol{T}_i\boldsymbol{k}_i\boldsymbol{T}_i^{\mathrm{T}}, \quad \boldsymbol{U}_i = \boldsymbol{T}_i\boldsymbol{\xi}_i, \quad \boldsymbol{W}_i = \boldsymbol{T}_i\boldsymbol{w}_i \tag{18}$$

$$\boldsymbol{T}_i = \mathrm{diag}[{}^{B}\boldsymbol{T}_{Bi}, {}^{B}\boldsymbol{T}_{Bi}, \ldots, {}^{B}\boldsymbol{T}_{Bi}, {}^{B}\boldsymbol{T}_{Bi}] \tag{19}$$

where ${}^{B}\boldsymbol{T}_{Bi}$ is the transformation matrix of B_i-$x_iy_iz_i$ with respect to B-xyz as defined in Eq. (2) and can be determined by inverse kinematics.

2.3 Equilibrium Equation of the Platform

The force diagram of the platform is depicted in Fig. 4. Herein, $\boldsymbol{F}_{Ai}/\boldsymbol{M}_{Ai}$ denotes the forces/moments provided by the passive joints A_i; $\boldsymbol{F}_{\mathrm{P}}/\boldsymbol{M}_{\mathrm{P}}$ represents the external forces/moments acting on the platform.

With the Newton's 2nd law, the following static equations can be formulated

$$-\sum \boldsymbol{F}_{Ai} + \boldsymbol{F}_{\mathrm{P}} = \boldsymbol{0}, -\sum \boldsymbol{r}_{Ai} \times \boldsymbol{F}_{Ai} + \boldsymbol{srritrmsrritrmM}_{\mathrm{P}} - \sum \boldsymbol{M}_{Ai} = \boldsymbol{0} \quad (20)$$

$$\boldsymbol{F}_{Ai} = {}^{B}\boldsymbol{T}_{Bi}\boldsymbol{f}_{Ai}, \quad \boldsymbol{M}_{Ai} = {}^{B}\boldsymbol{T}_{Bi}\boldsymbol{m}_{Ai} \quad (21)$$

where $\boldsymbol{r}_{Ai}$ is the vector pointing from A to A_i measured in B-xyz.

2.4 Deformation Compatibility Conditions

The displacement relationships between the platform/base and the limb is depicted in Fig. 5, in which $A_{i\mathrm{M}}$ and $A_{i\mathrm{L}}$ denote the interface points associated with the platform and the ith limb, while $B_{i\mathrm{B}}$ and $B_{i\mathrm{L}}$ denote the interface points associated with the base and the ith limb, respectively. $\nabla \boldsymbol{A}_i/\boldsymbol{\delta A}_i$ and $\boldsymbol{\delta B}_i/\boldsymbol{\rho}_{Bi}$ are the linear/angular displacements of $A_{i\mathrm{M}}$ and $B_{i\mathrm{L}}$ measured in the limb coordinate system B_i-$x_iy_iz_i$.

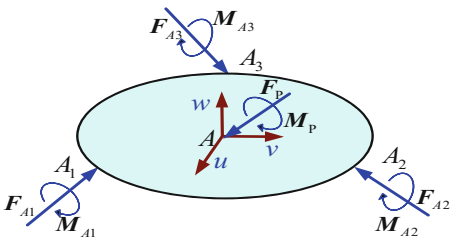


Fig. 4 Force diagram of the platform

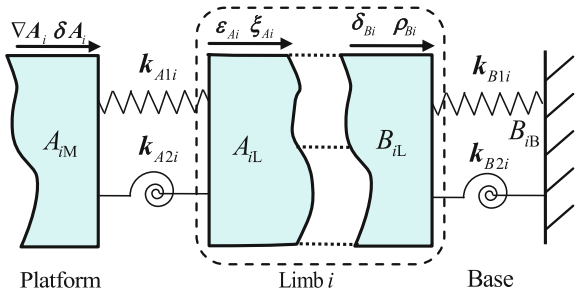


Fig. 5 Two types of deformation compatibility conditions

Assume the elastic motion of the platform caused by the deflections of the three flexible limb assemblages is $\boldsymbol{U}_{\mathrm{P}}$, the elastic displacements of $A_{i\mathrm{M}}$ (attached to the platform) can be derived as the followings.

$$\nabla \boldsymbol{A}_i = {}^{B}\boldsymbol{T}_{Bi}^{T}\boldsymbol{D}^{ri}\boldsymbol{U}_{\mathrm{P}}, \quad \delta \boldsymbol{A}_i = {}^{B}\boldsymbol{T}_{Bi}^{\mathrm{T}}\boldsymbol{D}^{ai}\boldsymbol{U}_{\mathrm{P}} \tag{22}$$

$$\boldsymbol{D}^{ri} = \begin{bmatrix} 1 & 0 & 0 & 0 & z_{Ai} & -y_{Ai} \\ 0 & 1 & 0 & -z_{Ai} & 0 & x_{Ai} \\ 0 & 0 & 1 & y_{Ai} & -x_{Ai} & 0 \end{bmatrix}, \quad \boldsymbol{D}^{ai} = \begin{bmatrix} \boldsymbol{0}_{3\times 3} & \mathbf{I}_{3\times 3} \end{bmatrix} \tag{23}$$

where x_{Ai}, y_{Ai} and z_{Ai}, are the coordinates of point A_i measured in B-xyz.

As a result, the reaction forces and moments of the passive joints A_i and B_i measured in B_i-$x_iy_iz_i$ can be obtained as

$$\boldsymbol{f}_{Ai} = -\boldsymbol{k}_{A1i}(\boldsymbol{N}_{Bi}^{A1}\boldsymbol{T}_i^{\mathrm{T}}\boldsymbol{U}_i - {}^{B}\boldsymbol{T}_{Bi}^{\mathrm{T}}\boldsymbol{D}^{ri}\boldsymbol{U}_{\mathrm{P}}), \quad \boldsymbol{m}_{Ai} = -\boldsymbol{k}_{A2i}(\boldsymbol{N}_{Bi}^{A2}\boldsymbol{T}_i^{\mathrm{T}}\boldsymbol{U}_i - {}^{B}\boldsymbol{T}_{Bi}^{\mathrm{T}}\boldsymbol{D}^{ai}\boldsymbol{U}_{\mathrm{P}}) \tag{24}$$

$$\boldsymbol{f}_{Bi} = -\boldsymbol{k}_{B1i}\boldsymbol{N}_{Bi}^{B1}\boldsymbol{T}_i^{\mathrm{T}}\boldsymbol{U}_i, \quad \boldsymbol{m}_{Bi} = -\boldsymbol{k}_{B2i}\boldsymbol{N}_{Bi}^{B2}\boldsymbol{T}_i^{\mathrm{T}}\boldsymbol{U}_i \tag{25}$$

2.5 Stiffness Matrix of the Platform

By assembling the equilibrium equations of the limbs and the platform, one may derive the governing equations of a general tripod PKM in the matrix form as

$$\boldsymbol{K}\boldsymbol{U} = \boldsymbol{W} \tag{26}$$

where $\boldsymbol{K}$, $\boldsymbol{U}$ and $\boldsymbol{W}$ are the governing stiffness matrix; coordinates vector and load vector. And there exist

$$\boldsymbol{K} = \begin{bmatrix} \boldsymbol{K}_{1,1} & & & \boldsymbol{K}_{1,4} \\ & \boldsymbol{K}_{2,2} & & \boldsymbol{K}_{2,4} \\ & & \boldsymbol{K}_{3,3} & \boldsymbol{K}_{3,4} \\ \boldsymbol{K}_{4,1} & \boldsymbol{K}_{4,2} & \boldsymbol{K}_{4,3} & \boldsymbol{K}_{4,4} \end{bmatrix} \tag{27}$$

$$\boldsymbol{U} = \begin{bmatrix} \boldsymbol{U}_1^{\mathrm{T}} & \boldsymbol{U}_2^{\mathrm{T}} & \boldsymbol{U}_3^{\mathrm{T}} & \boldsymbol{U}_{\mathrm{P}}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \tag{28}$$

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{W}_1^{\mathrm{T}} & \boldsymbol{W}_2^{\mathrm{T}} & \boldsymbol{W}_3^{\mathrm{T}} & \boldsymbol{W}_{\mathrm{P}}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}, \quad \boldsymbol{W}_{\mathrm{P}} = \begin{bmatrix} \boldsymbol{F}_{\mathrm{P}}^{\mathrm{T}} & \boldsymbol{M}_{\mathrm{P}}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \tag{29}$$

The stiffness matrix of the platform expressed in the body-fixed frame A-uvw can be further formulated as

$$\boldsymbol{K}_{\mathrm{p}} = \boldsymbol{T}_0^{\mathrm{T}}\{[\boldsymbol{K}^{-1}]_{(H-18(n+1))\times(H-18(n+1))}\}^{-1}\boldsymbol{T}_0, \quad \boldsymbol{T}_0 = \mathrm{diag}\left[{}^{B}\boldsymbol{T}_A \quad {}^{B}\boldsymbol{T}_A\right] \tag{30}$$

where H=18n + 24 is the dimension of the governing stiffness matrix.

3 Stiffness Estimation

In this section, two typical tripod PKMs, namely the Sprint Z3 head module (Case A) and the A3 head module (Case B) are taken as examples to demonstrate the versatility of the proposed general stiffness modeling methodology. The stiffness mapping of the two example systems over a typical work plane is plotted and briefly discussed.

The structures of the two example tripods are depicted in Fig. 6.

As can be observed from Fig. 6, the topological architecture behind the two tripods are a 3-PRS parallel mechanism and a 3-RPS parallel mechanism, respectively [24, 25].

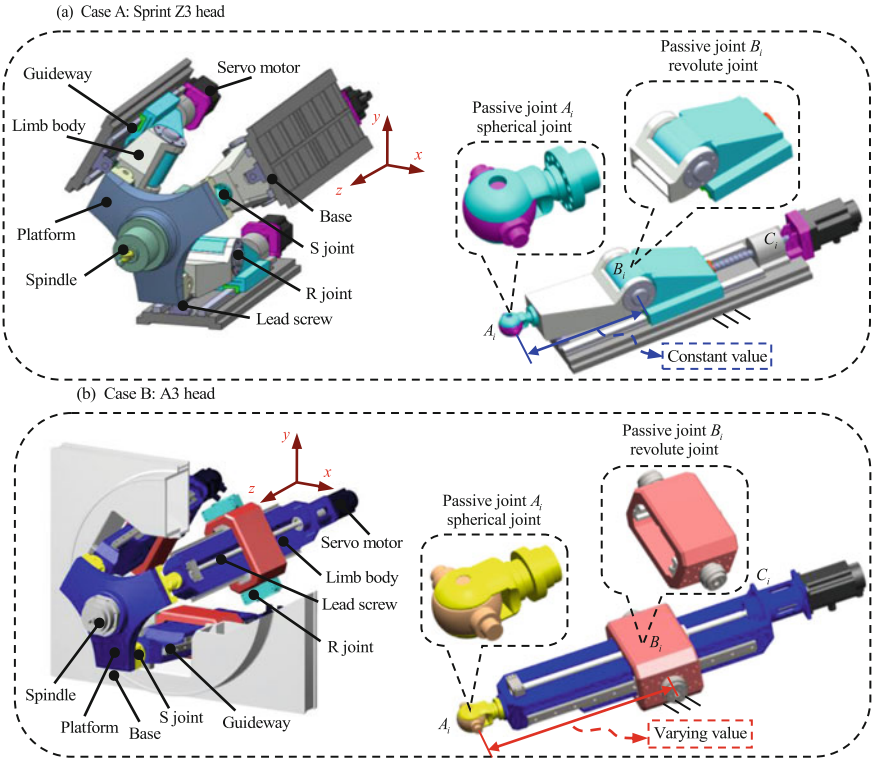


Fig. 6 Structures of two typical tripods

The major geometric parameters and stiffness coefficients of the two typical tripods are listed in Table 1.

Based on the above parameters and the derivations described in Sect. 2, one can obtain the stiffness mapping of the two tripods throughout the workspace. For the sake of generality, the following illustrates the mappings of the six principle stiffness values over a given typical work plane of $p_z = 570$ mm. Herein, p_z denotes the central distance between the base and the platform in z direction; ψ and θ are the Euler angles in terms of precession and nutation; k_{11}, k_{22} and k_{33} denote the three linear principle stiffness values along u, v and w axes while k_{44}, k_{55} and k_{66} represent the three angular principle stiffness values about u, v and w axes.

As shown in Figs. 7 and 8, the stiffness mapping of the two tripods is strongly position-dependent and is coincident with the tripod's structural features. To be specific, the distribution of stiffness values over the work plane for the two tripods are both symmetric with respect to θ, which can be physically explained by the structural symmetry of limb 1, limb 2 and limb 3.

Further observations show that ψ and θ have different impacts on the stiffness distributions of the two tripods. Taking the Sprint Z3 head for instance, the precession ψ has a 'stronger' impact on the linear principle stiffness values along u and v axes, while has a 'weaker' impact on the linear principle stiffness along w axis. In addition, for both the Sprint Z3 head and the A3 head, the stiffness value along

Table 1 Parameters of the Sprint Z3 head and A3 head

Nomenclature	Z3 head	A3 head
Radius of the platform r_p	250	250
Radius of the base r_b	250	250
Stroke of the tripod s	200	200
Elastic modulus of the limb body E	200	200
Shear modulus of the limb body G	80	80
Stiffness of short axis in u direction k_{su}	23	23
Stiffness of short axis in v direction k_{sv}	23	23
Stiffness of short axis in w direction k_{sw}	623	623
Stiffness of long axis in u direction k_{lu}	112	112
Stiffness of long axis in v direction k_{lv}	214	214
Stiffness of long axis in w direction k_{lw}	100	100
Stiffness of cross axis in u direction k_{cu}	676	676
Stiffness of cross axis in v direction k_{cv}	446	446
Stiffness of cross axis in w direction k_{cw}	348	348
Stiffness of a revolute joint along x direction k_{rx}	280	380
Stiffness of a revolute joint along y direction k_{ry}	330	530
Stiffness of a revolute joint along z direction k_{rz}	330	530
Stiffness of a revolute joint about y direction k_{rv}	20	18
Stiffness of a revolute joint about z direction k_{rw}	20	18

Units mm, Gpa, N μm^{-1}, N m rad^{-1}

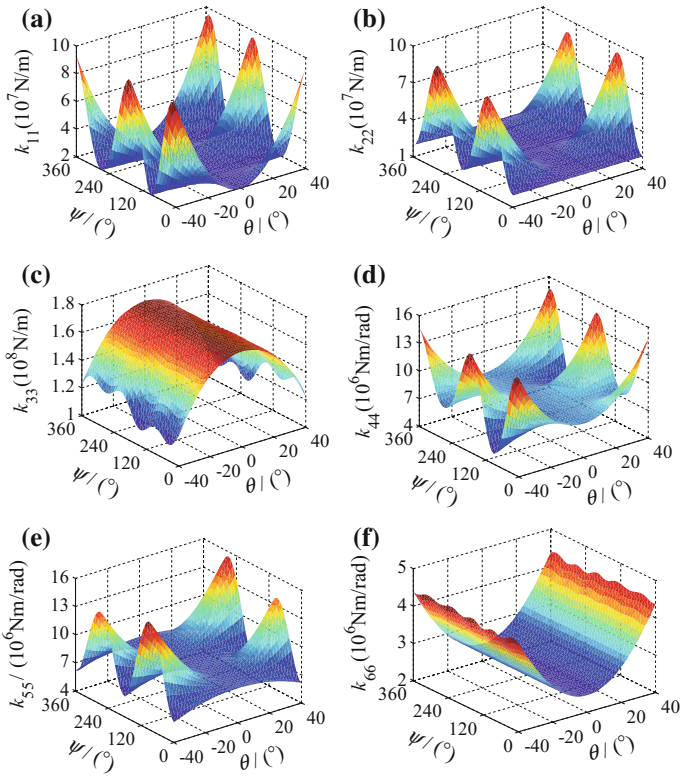


Fig. 7 Six principle stiffness values of the Sprint Z3 head over the work plane of p_z = 570 mm

w axis is the largest among the three linear principle stiffness values and the stiffness value about w axis is the smallest among the three angular principle stiffness values. This may imply that the rigidity performance about w direction must be paid more attention when design and apply these two kinds of tripods.

By comparing Figs. 7 with 8, one can find that the A3 head claims a competitive stiffness performance to the Sprint Z3 head along w axis. However, the rigidities of the A3 head along and about other axes are smaller than those of the Sprint Z3 head.

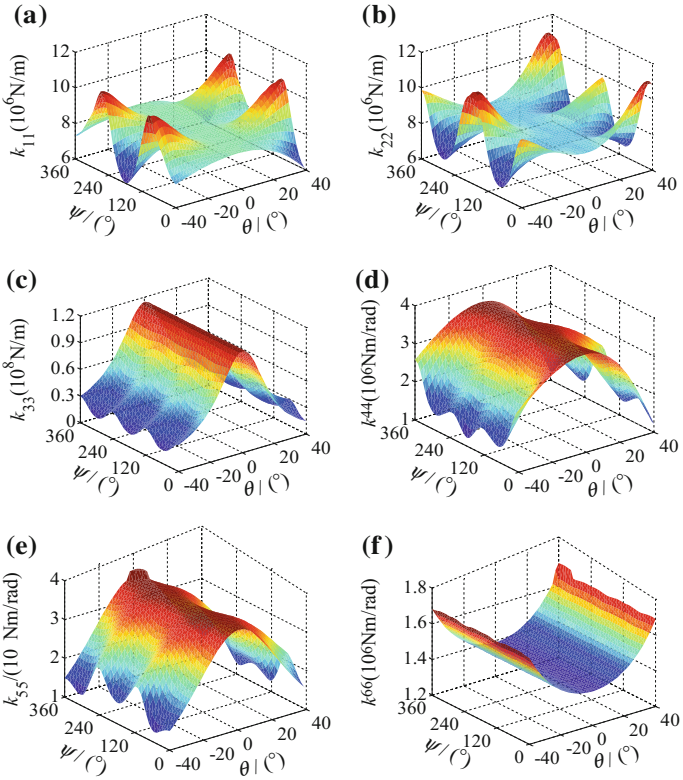


Fig. 8 Six principle stiffness values of the A3 head over the work plane of p_z = 570 mm

4 Conclusions

With the studies carried out in this paper, the following conclusions can be drawn:

(1) The general tripod PKMs are classified into two groups according to the motion patterns of the kinematic limbs. With this classification, a kinetostatic model based methodology is proposed to derive the analytical stiffness matrix of the platform for the general tripod PKMs with prismatic actuators.
(2) The proposed kinetostatic model considers the deflections of limb structures as well as the joint assemblages, leading to a satisfactory computation accuracy of the stiffness property for the tripod's platform.
(3) The stiffness mapping of two typical tripods named the Sprint Z3 head and the A3 head are predicted numerically to show strong position-dependency and structural symmetry.

(4) The present study provides a general framework for the stiffness evaluation of tripod-based PKMs with prismatic actuators. Also it is expected that the present kinetostatic model can be further expanded to an elastodynamic model by adding the mass and coriolis terms, with which the dynamic analyses for the tripods can be conducted.

Acknowledgements This work was supported by the Open Fund of the State Key Laboratory for Manufacturing Systems Engineering (Xi'an Jiaotong University) with Grant No. sklms2015004 and Open Fund of Shanghai Key Laboratory of Digital Manufacture for Thin-walled Structures with Grant No. 2014001. The second author would like to acknowledge for Innovation Research Fund for Postgraduates of Anhui University of Technology (Grant No. 2015032).

References

1. Y. Li, Q. Xu, Kinematic analysis of a 3-PRS parallel manipulator. Robot. Comput.-Integr. Manuf. **23**(4), 395–408 (2007)
2. N. Hennes, D. Staimer, Application of PKM in aerospace manufacturing-high performance machining centers ECOSPEED, ECOSPEED-F and ECOLINER, in *Proceedings of the 4th Chemnitz Parallel Kinematics Seminar* (2004), pp. 557–577
3. Z.M. Bi, Kinetostatic modeling of Exechon parallel kinematic machine for stiffness analysis. Int. J. Adv. Manuf. Technol. **71**(1–4), 325–335 (2014)
4. T. Bonnemains, H. Chanal, B.C. Bouzgarrou et al., Dynamic model of an overconstrained PKM with compliances: The Tripteor X7. Robot. Comput.-Integr. Manuf. **29**(1), 180–191 (2013)
5. Y.G. Li, H.T. Liu, X.M. Zhao et al., Design of a 3-DOF PKM module for large structural component machining. Mech. Mach. Theory **45**(6), 941–954 (2010)
6. Z.J. Tian et al., Dynamic modeling and eigenvalue evaluation of a 3-DOF PKM module. Chin. J. Mech. Eng. **23**(2), 166–173 (2010)
7. Y. Jin, Z. Bi, C. Higgins, et al., Optimal design of a new parallel kinematic machine for large volume machining. *Advances in Reconfigurable Mechanisms and Robots I*, (Springer, London, 2012), pp. 343–354
8. D. Zlatanov, M. Zoppi, R. Molfino, Constraint and singularity analysis of the exechon tripod, in *ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, (American Society of Mechanical Engineers, 2012), pp. 679–688
9. Z.M. Bi, Y. Jin, Kinematic modeling of Exechon parallel kinematic machine. Robot. Comput.-Integr. Manuf. **27**(1), 186–193 (2011)
10. X. Chen, X.J. Liu, F.G. Xie et al., A comparison study on motion/force transmissibility of two typical 3-DOF parallel manipulators: the sprint Z3 and A3 tool heads. Int. J. Adv. Rob. Syst. **11**(1), 74–89 (2014)
11. F. Xi, D. Zhang, C.M. Mechefske et al., Global kinetostatic modelling of tripod-based parallel kinematic machine. Mech. Mach. Theory **39**(4), 357–377 (2004)
12. A. Ahmad, K. Andersson, U. Sellgren, S. Khan, A stiffness modeling methodology for simulation-driven design of haptic devices. Eng. Comput. **30**(1), 125–141 (2012)
13. H.H. Pham, I.M. Chen, Stiffness modeling of flexure parallel mechanism. Precision Eng. **29** (4), 467–478 (2005)
14. T. Huang, X. Zhao, D.J. Whitehouse, Stiffness estimation of a tripod-based parallel kinematic machine. IEEE Trans. Robot. Autom. **18**(1), 50–58 (2002)

15. Z.M. Bi, S.Y.T. Lang, D. Zhang, Stiffness analysis of a tripod with a passive link, in *ASME 2005 International Mechanical Engineering Congress and Exposition*, (American Society of Mechanical Engineers, 2005), pp. 1665–1671
16. D. Zhang, L. Wang, Conceptual development of an enhanced tripod mechanism for machine tool. Robot. Comput.-Integr. Manuf. **21**(4), 318–327 (2005)
17. F. Majou, C. Gosselin, P. Wenger et al., Parametric stiffness analysis of the Orthoglide. Mech. Mach. Theory **42**(3), 296–311 (2007)
18. M. Ceccarelli, G. Carbone, A stiffness analysis for CaPaMan (Cassino parallel manipulator). Mech. Mach. Theory **37**(5), 427–439 (2002)
19. Y. Li, Q. Xu, Stiffness analysis for a 3-PUU parallel kinematic machine. Mech. Mach. Theory **43**(2), 186–200 (2008)
20. B. Li, H. Yu, Z. Deng et al., Stiffness modeling of a family of 6-DoF parallel mechanisms with three limbs based on screw theory. J. Mech. Sci. Technol. **24**(1), 373–382 (2010)
21. J.S. Dai, X. Ding, Compliance analysis of a three-legged rigidly-connected platform device. J. Mech. Des. **128**(4), 755–764 (2006)
22. Y.Y. Wang, T. Huang, X.M. Zhao et al., A semi-analytical approach for stiffness modeling of PKM by considering compliance of machine frame with complex geometry. Sci. Bull. **53**(16), 2565–2574 (2008)
23. H. David, *Fundamentals of Finite Element Analysis* (McGraw-Hill, NewYork, 2004)
24. Y.Q. Zhao, J. Zhang, L.Y. Ruan et al., A modified elasto-dynamic model based static stiffness evaluation for a 3-PRS PKM. Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci. **230**(3), 353–366 (2016)
25. J. Zhang, Y.Q. Zhao, H.W. Luo, Hybrid-model-based stiffness analysis of a three-revolute-prismatic-spherical parallel kinematic machine. Proc. Inst. Mech. Eng. Part B J. Eng. Manuf. doi:10.1177/0954405416634257(2016)

Part II
Sensors and Measurements

Magnetic Memory Testing for Fillet Weld of Boiler Tube

Kejia Wang, Ligang Yao, Yongwu Cai and Zhiwei Liao

Abstract Magnetic memory detection technology was applied to detect the fillet weld of boiler tube. Through the summary and analysis of the magnetic field intensity and the graphical distribution for each fillet weld of boiler tube, we could find out the more serious stress concentration problem of nozzle fillet weld, which provides the basis for the inspection of fillet weld of boiler tube.

Keywords Magnetic memory testing · Stress concentration · Fillet weld defects

1 Introduction

The safety of special equipment such as boiler, pressure vessel, and pressure pipeline is directly related to the national economic development and people's life and property safety. Therefore, the nondestructive testing method applied to quality testing, operation safety evaluation of these devices can prevent major accidents, and has great social and economic benefits [1, 2].

The use of conventional nondestructive testing methods to one hundred percent testing of boiler nozzle welds, including magnetic particle testing, penetrant testing, radiographic testing, ultrasonic testing, eddy current testing method, can effectively find the developed most macro or micro defects (such as cracks, corrosion, folding, inclusion etc.). However, to the welding stress concentration and early damage, especially the tacit not forming discontinuous change, it is difficult to implement effective evaluation, thus unable to avoid accidents caused by stress concentration and early damage development of equipment.

The magnetic memory testing technology can be used to accurately and reliably detect the serious dangerous area of the stress concentration problem on the weld seam of the boiler pipe, which is an effective method for the early diagnosis of the

K. Wang · L. Yao (✉) · Y. Cai · Z. Liao
School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou, Fujian 350116, P. R. China
e-mail: ylgyao@fzu.edu.cn

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_9

ferromagnetic metal equipment [3]. The magnetic memory testing method can accurately judge the defects in the weld joint of the boiler pipe, which is a good foundation for the later routine nondestructive testing, safety evaluation and life assessment [4].

2 Principle of Magnetic Memory Testing

Under the action of the stress, the metal will produce dislocation movement and form slip. When the dislocation motion is blocked, it will produce stress concentration. Even if the stress is removed, due to elastoplastic constraints, there still exists residual stress. At the same time, the residual stress in the metal, make the free energy in the structure increase. According to mechanics, spontaneous process of the system is the free energy reducing process of the system. So for the ferromagnetic components, in order to reduce the free energy of the system, the stress concentration area under the action of the earth's magnetic field will result in spontaneous rotation to form the magnetic pole, and offset part elastic energy caused by the stress concentration in the form of magnetic energy. The result is that the leakage magnetic field will be formed on the metal surface to cause magnetic memory. For the ferromagnetic material work piece under the action of geomagnetic and alternating loads, defects or the stress permeability is smallest, tangential component of magnetic field has a maximum, and the normal component changes the sign, and has zero point. It will form a leakage magnetic field in the surface of the work piece, the surface leakage magnetic field "remember" the defect of the components or the position of the stress concentration, that is, "magnetic memory" effect. Because the magnetic normal component on the surface of ferromagnetic material to zero is corresponding to the defects or the position of the stress concentration, the scanning and detection of leakage magnetic normal components on the surface of the work piece can indirectly determine the work piece whether has defect or position of stress concentration [5].

3 Damage Form of Fillet Weld of Boiler Tube

In the accident of boiler failure, the result of nozzle weld defects accounted for a large proportion. The main reason causing over the defect and damage is: welding or by various microscopic and macroscopic stress concentration failed to timely treatment, and the defects formation such as slag, porosity, lack of penetration and cracks caused by fatigue and failure. The basic characteristic is that under the continuous or alternating action of external stress, bending moment and torque, various types of micro defects (such as dislocations, slip and stress induced transformation) on the weld seam surface or inside, and gradually evolved into a macro crack, as well as the process of failure due to crack propagation, as shown in Fig. 1.

Fig. 1 Defects of fillet weld of boiler tube

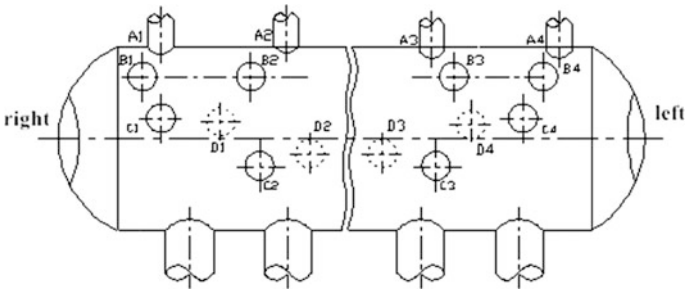


Fig. 2 Distribution of the rear nozzle of the 1# boiler barrel. Note: *A1*, *A2*, *A3*, *A4* are respectively the 1, 2, 7 and 8 fillet welds of saturated steam outlet tube seat from the *rear right*; *B1*, *B2*, *B3*, *B4* are respectively 1, 2, 11, 12 fillet welds of water inlet tube seat from the drum *rear right*; *C1*, *C2*, *C3*, *C4* are respectively 1, 2, 19, 20 fillet welds of increasing tube seat from the drum *rear left*; *D1*, *D2*, *D3*, *D4* are respectively 1, 2, 13, 14 fillet welds of increasing tube seat from the drum *front right*

4 Detection of the Basic Situation

The 1# boiler barrel material of one power plant is P355GH, the nozzle material is 20G, the whole boiler puts into use at the end of June 2006, the specific nozzle distribution of boiler barrel is shown in Fig. 2, and the barrel right view is shown in Fig. 3.

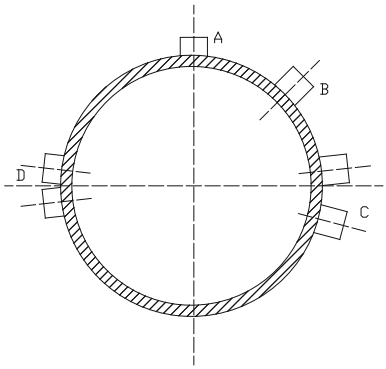


Fig. 3 Right view of boiler tube. Note: *A* is the saturated steam outlet tube; *B* is the water inlet tube; *C* is the rear increasing tube; *D* is the front increasing tube

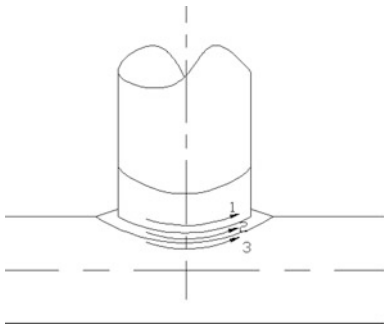


Fig. 4 Checking location of boiler nozzle weld

5 Field Test

The equipment used in this test is TSC-1 M-4 stress concentration detector which produced by Russian power diagnostic company, using the 1# sensor [6]. The test object is 1# boiler cylinder nozzle welds, planning to test 54 boiler tubes, including 8 saturated steam outlet tubes, 12 water supply tubes, 14 front increasing tubes, 20 rear increasing tubes. Specific detection method: firstly, numbering of the boiler nozzle weld, continuously scan the fillet weld of boiler barrel and nozzle and butt weld by using detecting instrument probe; detecting the size of magnetic strong sense H of the weld leakage magnetic field normal component, and number to store in the instrument. Then, analyzing the results by instrumentation and computer software to find out the stress concentration area, and make a judgment of defects. The specific checking location can be seen in Fig. 4.

6 Magnetic Memory Testing Results and Analysis

After the magnetic memory testing for 54 nozzles, a total of 39 nozzles have been found to have large stress concentration. 6 roots from 8 saturated steam outlet tubes exist stress concentration (1, 2, 3, 5, 6, 7 from the right); 10 roots from 14 front increasing tubes exist stress concentration (2 4, 5, 6, 7, 8, 9, 10 12, 13 from the right); 11 roots from 12 water supply tubes exist stress concentration (1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12 from the right); 12 roots from 20 rear increasing tubes exist stress concentration (2, 5, 6, 7, 8, 10, 12, 13, 15, 16, 17, 18 from the right), following to analyze the examples as 10th water supply tube (from the right) and the 5th saturated steam outlet tube (from the right).

6.1 The Fillet Weld Detection of the 10th Water Supply Tube (from the Right)

Through magnetic memory testing, we found that there exists more stress concentration in the 10th water supply tube from the right, and the detection part is shown in Fig. 5.

The specific magnetic memory test is shown in Figs. 6 and 7.

From the magnetic memory testing graphics can be seen: the image in the range of $0^{\circ} \sim 0^{\circ} + 30^{\circ}$, magnetic field strength Hp increases sharply from a very small value (i.e. gradient K value changes soon), specific values are shown in Table 1. And 2, 3 channel probes through the 0 line and changes the sign, so here is the larger stress concentration. In the range of $180^{\circ} + 35^{\circ} \sim 180^{\circ} + 55^{\circ}$, the magnetic field strength value Hp of the 2, 3 channel probes pass 0 line many times, and has a large gradient value K, so there also has a certain stress concentration.

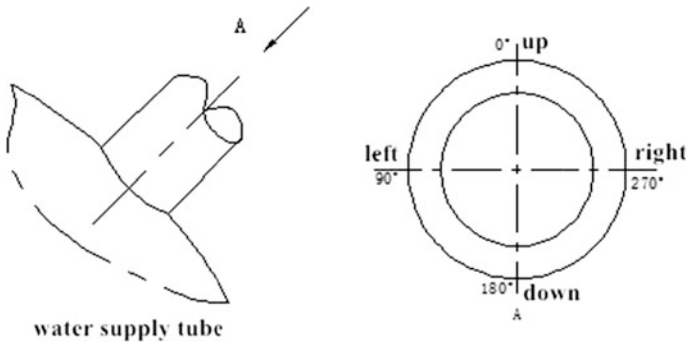


Fig. 5 Schematic diagram of water supply tube detecting position

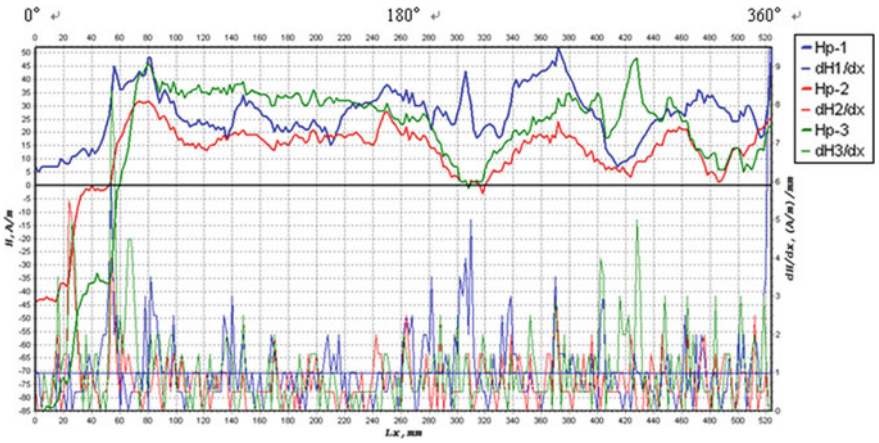
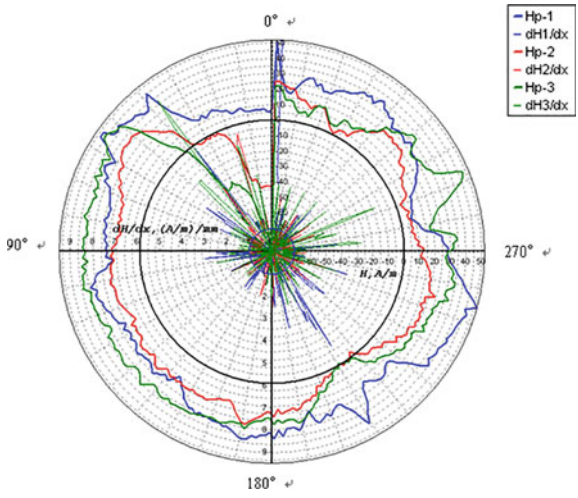


Fig. 6 2D graph of magnetic memory testing of fillet weld of the 10th water supply tube (from the *right*)

Fig. 7 Polar coordinate of magnetic memory testing of fillet weld of the 10th water supply tube (from the *right*)



6.2 *The Fillet Weld Detection of the 5th Saturated Steam Outlet Tube (from the Right)*

The detection position of the saturated steam outlet tube is shown in Fig. 8.

The specific magnetic memory test is shown in Figs. 9 and 10.

From the magnetic memory testing graphics can be seen: the image in the range of $0^{\circ} - 60^{\circ} \sim 0^{\circ} + 10^{\circ}$, magnetic field strength Hp of 1, 2 and 3 channel probes pass 0 line many times, the sign of magnetic field strength Hp changes constantly, and the magnetic field gradient value K is large, so here should have greater the stress

Table 1 Result parameters of magnetic memory testing of fillet weld of the 10th water supply tube (from the right)

Detection result							
Number	Sensor channel	Location of stress concentration area SCZ	Limit curve value of maximum stress concentration field Hp (min/max) (A/m)	Parameter value			Remarks
				K^{med}	K^{max}	m	
Gs_y10	1	$0^{\circ} \sim 0^{\circ} + 30^{\circ}$	5.0/52.0	1.105	9.500	8.601	
	2	$180^{\circ} + 35^{\circ}$	−44.0/32.0	0.804	5.500	6.839	
	3	$\sim 180^{\circ} + 55^{\circ}$	−85.0/48.0	1.063	8.500	7.998	

Note K^{med} is the average value of magnetic field gradient, K^{max} is the maximum value of magnetic field gradient, $m = K^{max}/K^{med}$

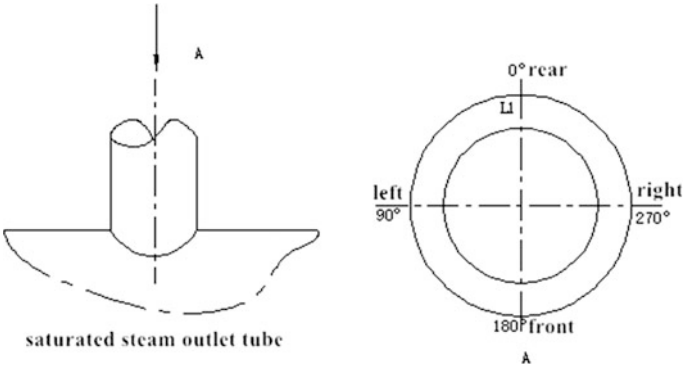


Fig. 8 Schematic diagram of the saturated steam outlet tube

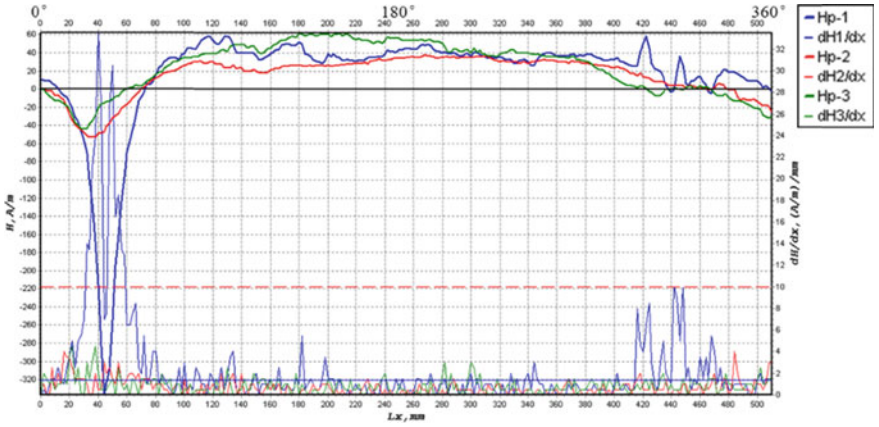


Fig. 9 2D graph of magnetic memory testing of fillet weld of the 5th saturated steam outlet tube (from the *right*)

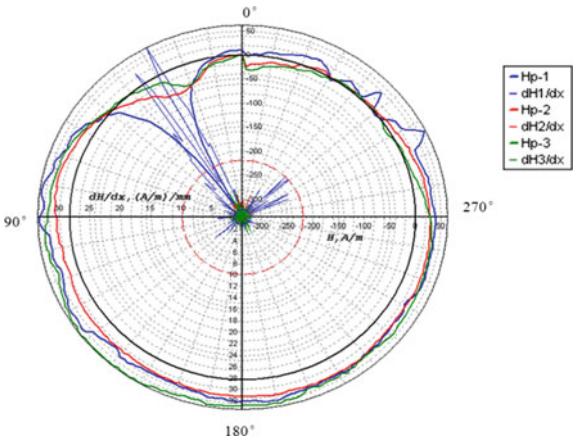


Fig. 10 Polar coordinate of magnetic memory testing of fillet weld of the 5th saturated steam outlet tube (from the *right*)

Table 2 Result parameters of magnetic memory testing of fillet weld

Detection result							
Number	Sensor channel	Location of stress concentration area SCZ	Limit curve value of maximum stress concentration field Hp (min/max) (A/m)	Parameter value			Remarks
				K^{med}	K^{max}	m	
Bh_y5	1	$0^{\circ} - 60^{\circ}$	337.0/59.0	2.486	33.500	13.474	
	2	$\sim 0^{\circ} + 10^{\circ}$	53.0/38.0	0.730	4.000	5.476	
	3	$0^{\circ} + 25^{\circ}$ $\sim 0^{\circ} + 40^{\circ}$	44.0/62.0	0.861	4.500	5.224	

Note K^{med} is the average value of magnetic field gradient, K^{max} is the maximum value of magnetic field gradient, $m = K^{max}/K^{med}$

concentration. In the range of $0^{\circ} + 25^{\circ} \sim 0^{\circ} + 40^{\circ}$, the size of magnetic field strength value Hp of 1 channel probe changes dramatically, and has a very large gradient value K, so, here also has a certain stress concentration. The specific parameter values are shown in Table 2.

7 Conclusions

Magnetic memory testing instrument was used for the field detection of boiler nozzle fillet weld and finds out weld joint whose stress concentration is more serious. Through the analysis of two nozzle welds, it explains the location and

extent of stress concentration to lay the foundation for further defect analysis, detection and safety evaluation. Practice has proved that the magnetic memory testing method can accurately find the location of stress concentration, and can be more intuitive to determine the extent of stress concentration. The magnetic memory method for the early detection of pressure bearing equipment can be used to judge the safety of equipment, and take reasonable measures as soon as possible, which has a strong application value.

References

1. R. Jilin, L. Junming et al., Metal Magnetic Memory Testing Technology. (China Electric Power Press, Beijing). **23**(1), 47–48 (2000)
2. Z. Guang, *Nondestructive Testing Technology Analysis Based on Magnetic Memory*. (China University of Petroleum, Beijing, 2005)
3. Z. Weichang, Theory of metal magnetic memory diagnosis method. Nondestr. Test. **23**(10), 424–426 (2001)
4. A.A. Dubov, Monitoring the stressed-strained state of pipelines. Tyazheloe Mashinostroenie **11**, 33–35 (2004)
5. Z. Junhua, *Research on the Mechanism of Magnetic Memory Testing*. (Master Thesis, Zhengzhou University, 2003)
6. Dobbrev. *Metal Magnetic Memory Testing Methods and Instruments*. (Russian Power Diagnostics Corporation, 2004)

Development of Projection Fringes Phase Shifting Technology Based Optical System for Small Object Surface Profile Fast Scanning

Ti-Yu Lai, Shu-Sheng Lee, Sheng Chi, Chia-Ming Jan and Po-Chi Hu

Abstract During the decades, the application of fringe projection techniques for rebuilding the three-dimensional(3D) surface has become more useful. The fringe projection techniques can measures the 3D surface as small as MEMS components, and as large as large panel (2.5 m × 0.45 m). The technique can also be applied in versatile fields, such as 3D intra-oral dental measurements [1], non-invasive 3D imaging and monitoring of vascular wall deformations [2], industrial and scientific applications such as characterization of MEMS components [3, 4]. Aiming to measure the small size of object, we have designed a system which can measure the object minimized to 2.5 cm × 2.5 cm × 2.5 cm. The components in the system were the computer, mini protector and the camera which were used to generate different phases of fringe patterns, project the patterns on the object, and record the fringe patterns quickly. The data were processed by filtering, phase-shifting, and unwrapping algorithms. Finally, the object's 3D surface profile has been rebuilt from the phase data. It's no need to use extra device to produce shifted phase images; therefore, it can shorten the time of measurement.

Keywords Projection fringes · Phase shifting technology · Surface profile scanning · Non-contact measurement

In this research work, we chooses Aiptek MobileCinema A50P DLP Pico Projector as the projector and BASLER acA2000-50gm as the camera [5]. The controllable parameters in this system are the width of the fringe, the angle of the projector, and

T.-Y. Lai · S.-S. Lee (✉)
Department. of Systems Engineering and Naval Architecture, National Taiwan Ocean University, No.2, Beining Rd, Keelung City 202, Taiwan
e-mail: sslee@email.ntou.edu.tw

S. Chi
Department. of Mechanical and Mechatronic Engineering, National Taiwan Ocean University, No.2, Beining Rd, Keelung City 202, Taiwan

C.-M. Jan · P.-C. Hu
Micro/Meso Mechanical Manufacturing R &D Department, Metal Industries Research & Development Centre, No.1001, Gaonan Highway, Nanzi Dist, Kaohsiung City 811, Taiwan

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_10

the angle of the camera [6]. According to these, a sequence of experiments using different focal lengths lenses, and changing different angles of projector and camera has done. Although the projected pattern on the object was about 2.5 cm × 2.5 cm, the camera can focus to the clear and smaller fringes and ignore the aberration ones. The 3D surface of the object smaller than 1 cm × 1 cm could be measured as well as the others.

1 Introduction

Nowadays many countries are committed to develop the non-contact 3D profile scanning system. The most used technology in the system is one by one point scanning or line scanning. It takes much time to finish scanning the whole object. When it is applied to dental treatment and used to measure the patient's teeth 3D profile, the dentist's hand shaking will introduce some error to the measurement result. The long scanning time also increases the discomforts to the patient. Therefore, the projection structure light technology becomes more and more popular to the scanning system. It can obtain the whole object's surface profile at once and reduce the scanning time. The periodic fringes such as binary fringes or interference fringes are simple structural pattern and often used to be projected to the object. The deformed fringes are recorded by the CCD/CMOS camera, and are analyzed and calculated to get the surface height of the object. In our research, we use the three-step phase shifting method to analyze the phase changes of the deformed fringes. This method needs fewer pictures, less equations and less calculation to get the result of measurement compared with other phase shifting method. As well as it can achieve the results more quickly, the error caused by the user manipulation might be reduced during the measurement.

2 Principle of the Method

The optical system herein is based on the projection fringes and the phase shifting method. The relationship between the object surface profile and the projected fringes is going to be described in this section. The 3-step phase shifting and the related equations for retrieving the wrapped phase are also illustrated.

2.1 The Projection Fringe Phase on Object Surface

The projection fringes used here is similar to interference fringes. They are designed and calculated according to the sinusoidal function and the pixel number by

personal computer. We denote it as "digital fringe" because it is not produced from the real optical system. The digital fringe intensity function $I(x, y)$ is

$$I(x, y) = a(x, y) + b(x, y) \cos \phi(x, y) \tag{1}$$

where $a(x, y)$ and $b(x, y)$ are the constants, and $\phi(x, y)$ is the phase of the reference plan. The fringe on the object is deformed and a phase related to the height $\psi(x, y)$ is added to the Eq. (1):

$$I_{obj}(x, y) = a(x, y) + b(x, y) \cos[\phi(x, y) + \psi(x, y)]. \tag{2}$$

Our optical system schematic is shown in Fig. 1. As mentioned in previous section, the projector projects the fringes to the object and CCD/CMOS camera records the images of the phase-changed fringes. The projection angle of the projector is θ_1 and the recording angle between the camera optical axis and the optical normal line is θ_2. The object height is h (Figure 2).

The distance between "A" and "C" recorded by CCD/CMOS camera can be expressed as the phase difference as

$$\overline{AC} = w \frac{\Delta\psi}{2\pi} \tag{3}$$

where w is the fringe width, and $\Delta\psi$ is the phase difference between A and C. The height h can be calculated by the following equation:

$$h = \frac{\overline{AC}}{\tan\theta_1 + \tan\theta_2} = w \frac{\Delta\psi}{2\pi} \frac{1}{(\tan\theta_1 + \tan\theta_2)} = k\Delta\psi \tag{4}$$

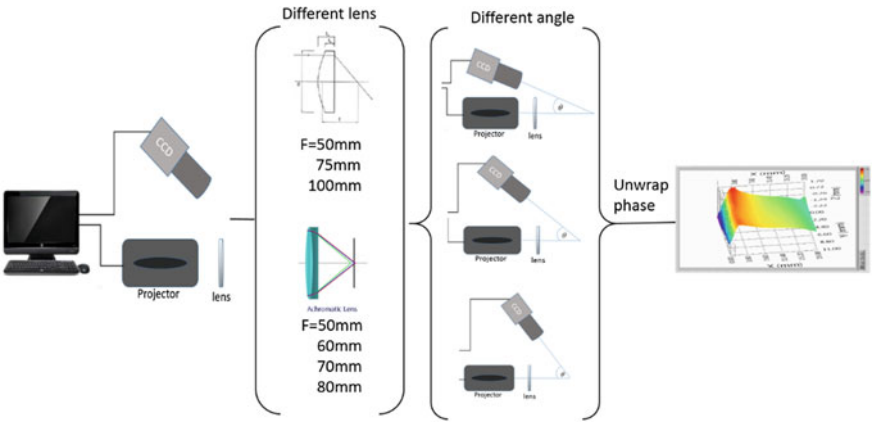


Fig. 1 The projecting fringes optical measurement system and controllable parameters

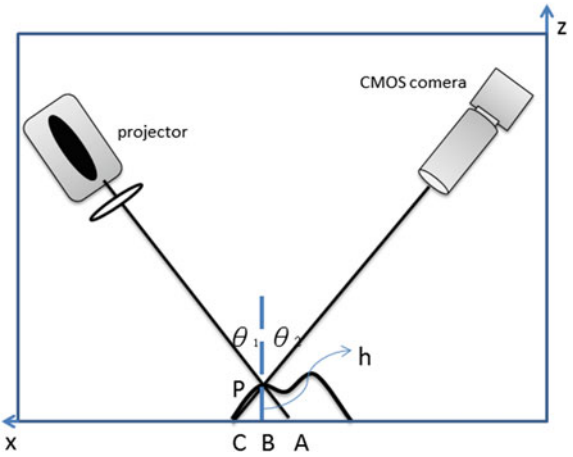


Fig. 2 Schematic of projection fringe optical system

where k is denoted as the height parameter and its representation equation is

$$k = \frac{w}{2\pi} \frac{1}{(\tan\theta_1 + \tan\theta_2)} \tag{5}$$

2.2 *Three-Step Phase Shifting Method*

Three different phase fringes images will be projected onto the object in three-step phase shifting method. The equations for the three fringes images are

$$I_0(x,y) = a_0(x,y) + b_0(x,y)\cos[\phi(x,y) + \psi(x,y)] \tag{6}$$

$$I_{\pi/2}(x,y) = a_{\pi/2}(x,y) + b_{\pi/2}(x,y)\cos[\phi(x,y) + \psi(x,y) + \frac{\pi}{2}] \tag{7}$$

$$I_{\pi}(x,y) = a_{\pi}(x,y) + b_{\pi}(x,y)\cos[\phi(x,y) + \psi(x,y) + \pi] \tag{8}$$

From the Eqs. (6) to (8), the representation for every pixel's phase $\phi(x,y)$ of the reference plane is obtained as

$$\phi(x,y) = \arctan\frac{I_0(x,y) - 2I_{\pi/2}(x,y) + I_{\pi}(x,y)}{I_0(x,y) - I_{\pi}(x,y)}, \tag{9}$$

By Eq. (9), we can calculate the wrapped phase $\phi(x,y) + \psi(x,y)$ and reconstruct it into continuous phase by Fourier Transformation. Then, substrate the continuous phase $\phi(x,y) + \psi(x,y)$ with the designed $\phi(x,y)$ to obtain $\psi(x,y)$. Finally, multiply $\psi(x,y)$ with the experimental optical system parameter k, and the object's 3D profile can be completed.

3 Experimental Arrangement and Experiment Result

The experimental setup, the results of the experiment with different parameters, and consequence discussion will be stated in this section.

3.1 The Optical Setup

The photo of our optical system is show in Fig. 3. A digital light processing (DLP) micro projector (MobileCinema A50P, AIPTEK International Inc.) was implemented to project the digital fringes image on the object on reference plane. The projector's optical axis was parallel to the normal direction of the reference plane. It means that the angle of the projector θ_1 is 0°. The distance between the reference plane and projector was 85.5 mm. A lens was put between the projector and object to reduce the image size. The distance between the reference plane and the CMOS camera is 91.5 mm.

3.2 The Experimental Result

When the optical system has been set up well, three different phases of the fringes images were projected onto the object in sequence. The CMOS camera has recorded the three deformed fringes images as Fig. 4. To verify the affection of the different recording angles, we have repeated the experiment at the angles 30.0°, 32.5°, 35.0°, 37.5°, 40.0°, 42.5° and 45.0°.

The grayscale values of these three images were transferred from the bmp files with Matlab program. The wrapped phase $\phi(x, y) + \psi(x, y)$ can be calculated from

Fig. 3 The experimental optical system

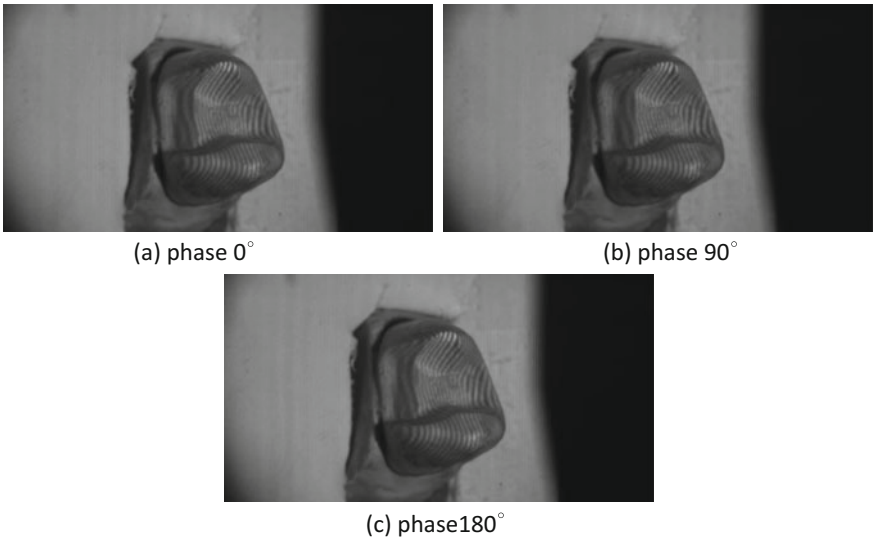


Fig. 4 Three deformed fringes images with different phases, 0°(**a**), 90°(**b**), and 180°(**c**), at the recording angle 30.0°

the three grayscale images data according to the Eq. (9) and be reconstructed to the unwrapped phases. Multiplied the unwrapped phases with the height parameter k and the height of the object referred to the reference plane was gotten. The distance between every pixel in *x* and *y* direction can be known from the experiment. Hence, the object 3D profile has been reconstructed from the subtraction between the previous height data and the reference plane's height shown as in Fig. 5.

With the assist of the LabVIEW program, we could easily identify the highest point and the second highest one. We have compared these two points and got the height difference between them at different recording angles that were listed in Table 1.

4 Discussion and Conclusion

In this paper, we have demonstrated a selves-developed optical system, which can acquired the 3D surface profile of the small object quickly. From the experiment results, the average of the height difference between the two fixed points obtained from different recording angles was 1.08041 mm and most data were near this value. This verified that our system owned a good accuracy and can repeat the measurement result. So, we have achieved the goal to complete a small object surface profile fast scanning optical system based on the projection fringes phase shifting technology.

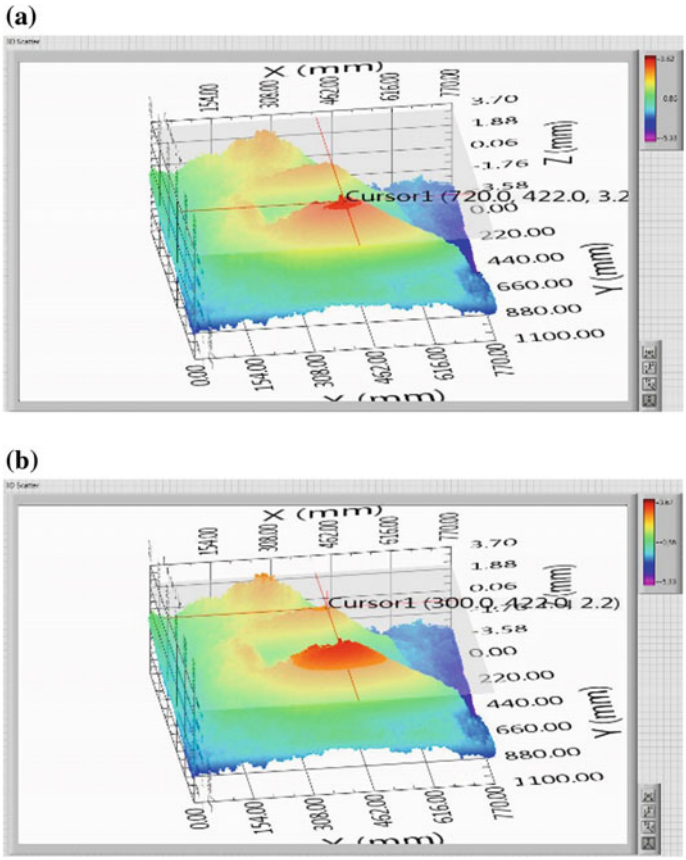


Fig. 5 The 3D profile of the object recorded at angle 35.0°: the highest point (**a**) and the second highest point (**b**)

Table 1 The difference between the highest and the second one

Recording angle (°)	Height difference (mm)
30.0	1.00594
32.5	0.14186
35.0	1.02254
37.5	1.06526
40.0	1.12797
42.5	1.71394
45.0	1.48534

Acknowledgements This research work is financially supported by Metal Industries Research & Development Centre of Taiwan.

References

1. L. Chen, C. Huang, Miniaturized 3D surface profilometer using digital fringe projection. Meas. Sci. Technol. **16**(5), 1061–1068 (2005)
2. K. Genovese, C. Pappalettere, Whole 3D shape reconstruction of vascular segments under pressure via fringe projection techniques. Opt. Laser Eng. **44**(12), 1311–1323 (2006)
3. C. Quan, C.J. Tay, X.Y. He, X. Kang, H.M. Shang, Microscopic surface contouring by fringe projection method. Opt. Laser Technol. **34**(7), 547–552 (2002)
4. X. He, W. Sun, X. Zheng, M. Nie, Static and dynamic deformation measure- ments of micro beams by the technique of digital image correlation. Key Eng. Mater. **326–328**, 211–214 (2006)
5. S. Chi, S.S. Lee, J. Jen-Yu Huang, T.Y. Lai, C.M. Jan, P.C. Hu, Design and verification of the miniature optical system for small object surface profile fast scanning, Proc. SPIE 9899, Optical Sensing and Detection IV, 989928 (2016)
6. S.Y Chou, C.H. Cho, M.J. Du, S. Chi, Projection moire profilometry with high-dynamic range image, Opti. Eng. **51**(2), (2012)

Sensitivity Effect of Single Load Cell on Total Output of a Combinatorial Structure

Guiyong Guo, Shuncong Zhong and Ligang Yao

Abstract A rule of the effect between the output of a single load cell and a combinatorial structure was proposed by changing the sensitivity of the single load cell. The relevant relationship between the sensitivity of single load cell and the variation of total output of the combinatorial structure was analyzed by the numerical derivation method and experimental validation. In the cases, the output of a single load cell is linear or nonlinear, when the same load was acted on the combinatorial structure. The results showed that the effect was limited and the influence quantity is due to the accuracy and number of load cell. However, it must be controlled in a certain range to avoid additional measurement errors to the combinatorial structure. Moreover, the rule provides an evaluation method with improving the measurement accuracy for the combinatorial structure. Consequently, the rule can be employed as a useful guidance for the product quality when this evaluation method is applied to the load cell selection for the electronic truck scales.

Keywords Sensitivity · Effect · Load cell · Combinatorial structure · Accuracy

G. Guo · S. Zhong (✉) · L. Yao
School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350108, People's Republic of China
e-mail: zhongshuncong@hotmail.com

G. Guo
Fujian Province Institute of Metrology, Fuzhou 350003, People's Republic of China

S. Zhong
Fujian Key Laboratory of Medical Instrument and Pharmaceurical Technology, No.2 Campus Road, 350108 Fuzhou, People's Republic of China

S. Zhong
Shanghai University, Shanghai, P. R. China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_11

1 Introduction

The importance of load cell in modern science and technology has been more and more recognized. Currently, various applications cannot be separated from load cells, whether in the field of industrial, biological engineering, health care, environmental protection or other fields [1, 2]. Different accuracies and specifications have been needed in different fields. However, it is an inevitable trend towards the load cells with better and better characteristics. One of the basic characteristics of load cell is 'input-output' relevant relationship. Under the action of the steady state signal, the characteristic is a static characteristic [3–5]. The important indicators of the static characteristics are the linearity, sensitivity, hysteresis and repeatability. For the combinatorial structure, it composes of $2 \sim n$ (n is a positive integer) single load cells, bearing platen, balanced platen and other components. According to different uses, it can be assembled into different combinations to achieve different accuracies, measurement ranges and dimensions of technical requirements.

Bray. A and Prof. G. Barbato's orientation [6] discussed compensating effect of combinatorial load cell but not mentioned sensitivity effect. A new calibration system has been developed by Fujian Province Institute of Metrology, but the key technologies of load cells are mainly about the combination with kinds of components [7]. According to statistics, other researchers also had not focused on this topic. Consequently, there are only a few relevant research articles about the sensitivity effect of single load cell to the total output of a combinatorial structure.

In the light of numerical derivation method, the linearity is an important factor for a group of regular numbers. Based on the simple principle, the paper will analyze the sensitivity effect of the load cell with different accuracies, linearity and discuss the application in the electronic truck scales. Based on the rule, it is a useful method to weaken the magnitude of sensitivity effect for load cell manufactories.

2 Model and Basic Principle

The sensitivity effect rules of the load cell discussed in the paper is similar to the rules of numerical derivation method as follows. As shown in Eqs. (1) and (2), the algebraic sum results of several linear arrays are also linear. However, the algebraic sum results may accordingly vary with the changes of some arrays if they are nonlinear.

$$\begin{cases} l_1 = (a_1, a_2, \ldots, a_j)^T \\ l_2 = (b_1, b_2, \ldots, b_j)^T \\ \cdots \\ l_m = (m_1, m_2, \ldots, m_j)^T \end{cases} \tag{1}$$

$$\begin{aligned} l &= l_1 + l_2 + \cdots + l_m \\ &= (a_1 + b_1 + \cdots + m_1, a_2 + b_2 + \cdots + m_2, \cdots, a_j + b_j + \cdots + m_j)^T \end{aligned} \tag{2}$$

In the paper, a combinatorial structure has some favorable functions in practical applications. We usually wanted to select the smaller height, larger measurement rang and higher accuracy of load cell. However, the single load cell cannot meet the technical specifications of some load cells with high-accuracy. Hence, combinatorial structure is proposed to solve the problems. There has been a conventional method for the combinatorial structure by adjusting potentiometer of circuits but it is not suitable for long term stability. So we selected three single load cells combined together as a triple-pair structure as shown in the Fig. 1, which is including bearing platen, balance platen, three single load cells and base platen.

The circuits of a combinatorial structure are combined together with several load cells, the equivalent circuit is as shown in Fig. 2. As we know, the total voltage is equal to the voltage of every branch in the voltage of parallel circuit.

The nodal voltage of the equivalent circuit can be expressed as follows,

$$E = \sum_{i=1}^{n} (e_i/R_i) / \sum_{i=1}^{n} (1/R_i) \tag{3}$$

where R_i $(i = 1, 2, \ldots, n)$ and e_i are the resistance and the voltage of every branch circuit, respectively. From the Eq. (3), we can see that for the equivalent circuit, the value of $\sum_{i=1}^{n} (1/R_i)$ is a constant. In order to obtain a high-accuracy measurement capability, every load cell should have the same value of R_i and e_i, respectively. Another possibility is that the value of e_i/R_i is also a constant.

According to the calibration specification ISO 376:2011(E) [8] and verification regulation JJG 144-2007 [9], the linearity can be defined as follows, $L = \Delta\theta_L/\Delta\theta_f$, where $\Delta\theta_L$ is the maximum deviation value of the process average calibration curve and the average end-point line. And $\Delta\theta_L$ is the rated output value, for the load cell the sensitivity denotes the proportion S between output Δy and input Δx that is $S = \Delta y/\Delta x$. In this paper, the sensitivities of three single load cells (specification: RTN C3/10T) are S_1, S_2 and S_3, which are obtained by acting a standard force F_i ($F_i = 10^4$ kgf, and $i = 1, 2, 3$) on every single load cell, respectively. Although the

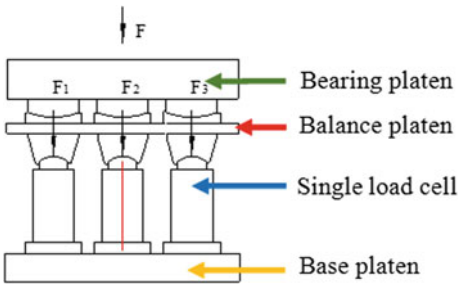


Fig. 1 Load bearing model of a combinatorial structure, which is a triple-pair structure

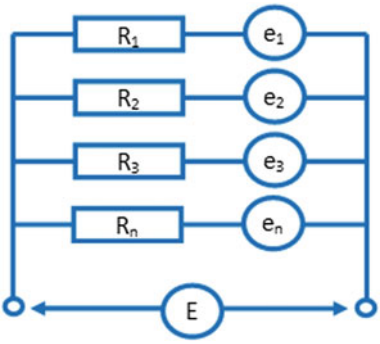


Fig. 2 Equivalent circuit of a combinatorial structure

sensitivity value S_0 ($S_0 = S_{n0} \pm \Delta S_{n0} = 2.85 \pm 2.85 \times 0.1\%$) of every load cell is provided by the manufacturer, which is only a nominal sensitivity coefficient. Actually, the sensitivity value is equal to S_0' ($S_0' = S_{n0} \pm \Delta S_{n0}' = 2.85 \pm 2.85 \times 0.2\%$).

A standard force F_i ($F_i = w_i \cdot F$, $n = 1, 2, 3$) was acted on the combinatorial structure, where w_i denotes weight coefficients of every load cell. Therefore, the total output of the combinatorial structure as follows,

$$S = \sum_{i=1}^{3} F_i \cdot S_i / 10^4 = F \times 10^{-4} \times (w_1 \cdot S_1 + w_2 \cdot S_2 + w_2 \cdot S_2) \tag{4}$$

In order to calculate the deviation $\Delta\theta_{k,i}$ of output, it can be obtained from Eq. (5) as follows,

$$\Delta\theta_{k,i} = \theta_{k,i} - \theta_{k,ir} \tag{5}$$

$$\theta_{k,ir} = S_{0,i} \cdot F_{k,i} / F_i \tag{6}$$

where the subscript i, k denote the k th standard force acting on the i th load cell, and r is the referential or theoretical value of the load cells above.

3 Testing Result and Analysis

The dead-weight force standard machine (span = 30,000 kgf, accuracy = 0.005%) was used as the standard test equipment for the measurement of sensitivity. And a DMP40 indicating instrument made by HBM company was used to connect the load cells and displayed the sensitivity value in the testing. The environmental conditions as follow: temperature = (21.0 ± 0.5) °C, humidity = (55 ± 5)% RH. The axis of combinatorial load cell was concentric with the axis of the dead-weight force standard machine by putting a mark in the same position. According to the

calibration specification ISO 376:2011(E), three positions distributed uniformly on the dead-weight force standard machine. So three positions had been chosen for calibration to obtain for more accuracy result,which included 0, 120° and 240° (Figs 3, 4, 5 and 6).

4 Application in the Load Cell Selection for the Electronic Truck Scales

As shown in Fig. 7, electronic truck scale is used to weight for trucks and materials, which consists of N $(N \geq 3)$ load cells and several weighing platforms, and installed on the foundation made of reinforced concrete.

The accuracy of electronic truck scale is related to the stiffness of weighing platforms, mechanical installation conditions, accuracy level of the load cells and other factors, but also with the sensitivity coefficient of every load cell. The condition of realizing the full parallel circuit of every load cell as follows,

$$p = \frac{S_1}{R_1} = \frac{S_2}{R_2} = \cdots = \frac{S_N}{R_N} \tag{7}$$

Where R_N $(N = 1, 2, \ldots, n)$ and S_N are the resistance and sensitivity value of every load cell, respectively. From the Eq. (7), we can see that for the electric truck scale, the proportion value of every load cell should be the same as each other so as to obtain a normal accuracy, certainly it also depends on the accuracy of the electric truck scale. For quite some time, the manufacturers always adjust potentiometer of circuits so as to ensure the accuracy. But it is not suitable for long term stability, especially when the long term stability is so important to the consumers.

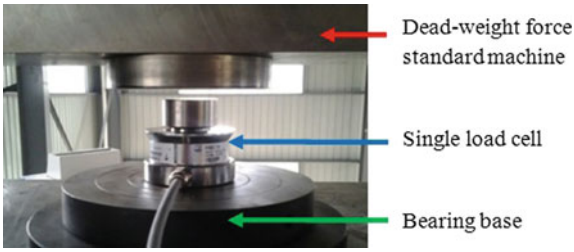


Fig. 3 Testing of a single load cell. The dead-weight force standard machine (span = 30,000 kgf, accuracy = 0.005%) was used as the standard test equipment for the measurement of sensitivity

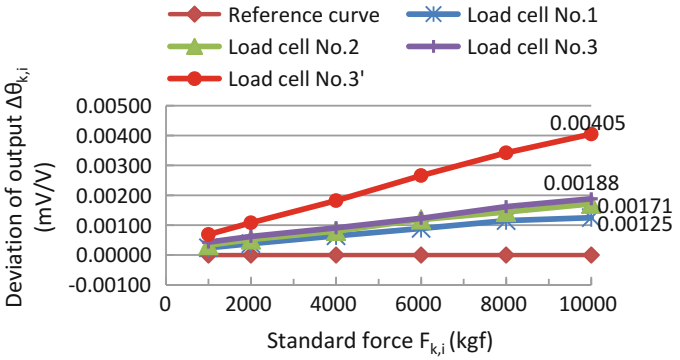


Fig. 4 Deviation of four load cells with different sensitivities: the deviation of load cell no. 1, no. 2 and no. 3 is less than 0.03% whist the deviation of load cell no. 1, no. 2 and no. 3' is nerarly equal to 0.10%

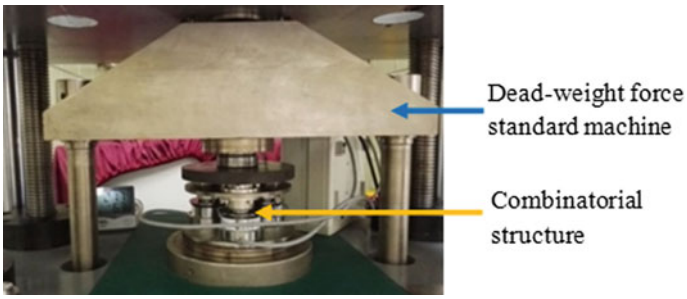


Fig. 5 Testing of a combinatorial structure. The axis of combinatorial structure was concentric with the axis of the dead-weight force standard machine by putting a mark in the same positions

Fig. 6 Deviation of output a combinatorial structure combined with different load cells. The 1st condition of combinatorial structure is linear curve (L = 0.029%), which consists of load cells no. 1, no. 2 and no. 3. The 2nd condition is nonlinear curve (L = 0.054%), which consists of load cells no. 1, no. 2 and no. 3'

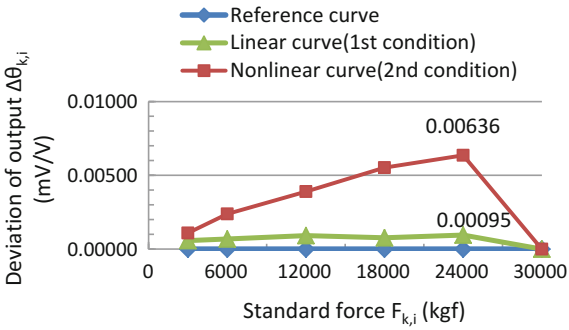


Fig. 7 The electric truck used to weight for trucks and materials. This is one of the applications by employing the evaluation method as a useful guidance for the product quality

5 Conclusions

In the work, a new consideration was proposed for load cells, which is an important factor to a combinatorial structure, especially to the linearity and accuracy. For a combinatorial structure, when the output deviation is controlled in a small value such as 0.03%, the linearity will be in a reasonable range such as 0.029%. On the contrary, when the output deviation value such as 0.10% is directly selected from the specifications, the linearity will obviously not meet technical specifications of some load cells with high-accuracy.

Acknowledgements We gratefully acknowledge support from the National Key Scientific Instruments and Equipment Development Projects of China under Grant No. 2011YQ090009.

References

1. H. Zhang, J. Cui, D. Chen, Y. Chen. Study on a new type of adopted structure for three dimensions strain gauge force sensor. Chin. J. Sens. Actuators. **27**(2), 162–167 (2014)
2. Lifan Meng, *Jinhui Lan* (Principle and application of sensor, Beijing, 2011)
3. T. Wei, X. Yang, J .Yao, H. Xu. Rotary effect on composite loading sensors. Chin. J. Cons. Machinery. **12**(1), 67–72 (2014)
4. T.P. Wei, X.X. YANG, J.H. YAO et al., The additional side force on the force transducer in the combinatorial load cell [C]. Appl. Mech. Mater. **541**, 1327–1332 (2014)
5. T.P. Wei, X.X. YANG, J.H. YAO et al., The influence of a balanced structure on the rotation effect of a build-up system [J]. Measurement **61**(2), 162–168 (2015)
6. A. Bray, G. Barbato, R. Levi, in *Theory and Practice of Force Measurement* (Academic Press, Italy, 1990)
7. J.H Yao, China Metrology vol 5(2011), p. 048
8. International Organization for Standardization: ISO 376-2011 Metallic Materials—Calibration of Force Proving Instruments Used for the Verification of Uniaxial Testing Machines (Switzerland 2011)
9. General Administration of Quality Supervision Inspection and Quarantine of the PRC: JJG 144-2007 Verification Regulation of Standard Force Measuring Instrument (Chinese Metrology Press, Beijing 2007)

Curvature Radius Measurement of Lens by Using Line Field Spectral Domain Optical Coherence Tomography

Jiewen Lin, Shuncong Zhong, Tengfei You, Qiukun Zhang and Junhai Tong

Abstract A non-contact and non-invasion method to measure lens curvature radius was proposed. In the experiment, we developed a homemade Line Field Spectral Domain Optical Coherence Tomography (LF-OCT) System to measure lens curvature and focal length, using a broadband super luminescent diode (SLD) light source having a center wavelength of 830 nm. By using the methods of spectrum correction, the frequency, phase and amplitude can be corrected automatically and we can reach a higher accuracy. The result demonstrated that the LF-OCT has high accuracy and reliability and therefore it could be recommended to use in optical lens measurement due to its non-contact and non-invasion in nature.

Keywords Line Field Spectral Domain Optical Coherence Tomography (LF-OCT) · Spectrum correction · Non-contact · Non-invasion

1 Introduction

Some measurement methods for lens curvature radius has been proposed, such equal thickness interference [1], autocollimation method [2], differential confocal microscopy technology [3] and spherometers [4] and so on. Optical Coherence Tomography (OCT) [5] was proposed by Huang et al. and it is currently widely used in medicine imaging [6] and the detection of composite materials [7] etc. Zhong et al. extent the function of OCT for vibration monitoring [8, 9]. The vibration signal is colletected from the direct displacement of structure surfaces. Therefore, OCT is capable of surface curve measurement. In addition, line field spectral domain optical coherence tomography (LF-OCT) system [10] was

J. Lin · T. You · Q. Zhang · J. Tong
Laboratory of Optics, Terahertz and Non-Destructive Testing,
School of Mechanical Engineering and Automation, Fuzhou University,
Fuzhou 350108, People's Republic of China

S. Zhong
Shanghai University, Shanghai, P. R. China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_12

proposed to increase the OCT imaging speed, which does not need to scan the samples. In the present work, a home-made LF-OCT was proposed to measure the curvature radius of Lens.

2 LF-OCT Measurement System and Its Main Parameters

Figure 1 shows the schematic diagram of home-made LF-OCT system, including SLD light source, Michelson Interferometer, and CCD camera etc. [9, 11]. A SLD light source having a center wavelength of 830 nm were used. As shown in Fig. 1, when SLD light source (low coherence light) pass the convex lens with a focal length of 50 mm, it becomes a beam of parallel light. Before the parallel light entering Michelson Interferometer, it should be focused into a line by a cylindrical lens. The light will be divided into one sample beam and one reference beam by a cube beam splitter. The light reflected from the sample (convex lens), so we used a convex lens to collect the divergent light [8, 9]. The interference light passes through the light grating, it will be dispersed according to the wavelength and then we can used CCD camera to capture it. From the interference pattern, we can obtain the curvature information of the lens.

The light source could be expressed as

$$E(k, \omega, r) = A(k) \times \exp[-i(\omega t - kr)] \tag{1}$$

where A(k) is the amplitude spectrum of SLD light source's electric field strength, ω is the angular frequency, and k is the number of waves. Due to the detector signal

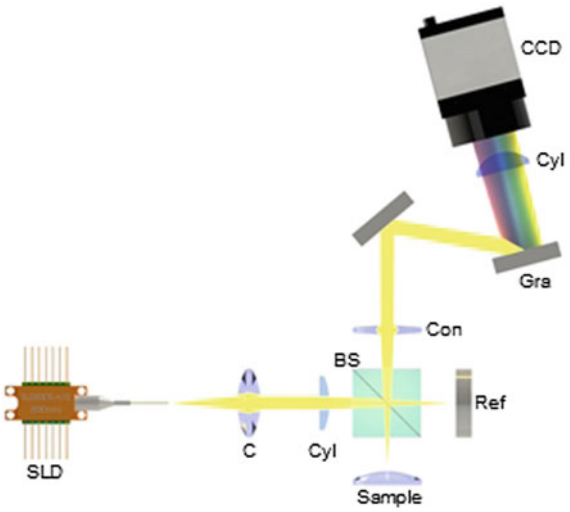


Fig. 1 The schematic diagram of home-made LF-OCT system. *C* Collimator, *Cyl* Cylindrical Lens, *BS* Cube Beam Splitter, *Ref* Reference Interface (Flat glass surface), *Con* Convex Lens, *Gra* Optical Grating

only related to the optical path difference of sample light and reference light but not detection time, so the light source can be expressed as

$$E(k, \omega, r) = A(k) \times \exp(ikr) \tag{2}$$

The reference light source can be expressed as

$$E_r(k) = A(k) \times \exp(i2kr) \tag{3}$$

where 2r is the optical path length of reference light, 2kr is the phase.

The reflected light from sample beam at different depths can be expressed as

$$E_s(k) = \int_{-\infty}^{+\infty} A_s(z) \ \exp(i2k(r+nz))dz \tag{4}$$

where $E_S(K)$ is the amplitude of reflected light at different depths, n is the refractive index. If we ignored multiple reflection and scattering of light in a sample, we can make $2(r + nz)$ as the optical path at depth z. When the reference light and the sample light interfere, the interference optical signal $I(k)$ can expressed as

$$I(k) = S(k)\left[E_r^2 + 2E_s \int_{-\infty}^{+\infty} A_s(z)\cos(2knz)dz + \iint_{-\infty}^{+\infty} A_s(z)A_s(z')\exp(i2k(nz - nz'))dzdz'\right] \tag{5}$$

where the first item is the DC component of interference spectrum, the second item is the superposition of multiple light reflected from different depth and the third item is the self-coherence in different depth of the sample. The third item is much smaller than the reference light intensity, so $I(k)$ can briefly expressed as

$$I(k) = S(k)[E_r^2 + 2E_s \int_{-\infty}^{+\infty} A_s(z)\cos(2knz)dz] \tag{6}$$

By doing Fourier transform for Eq. (6), we can obtain

$$F\{I(k)\} = F\{S(k)\} * \{E_r^2\delta(z) + \left[E_r + E_s^*(-z)\right]\} \tag{7}$$

3 Experiment and Discussions

We chose three plane convex lens (Thorlabs LA1433, LA1131, and LA1608) with focal length of 150 mm, 50 mm and 75 mm as the samples. The parameters of the Lens were list in the Table 1. Figure 2a showed the interference fringes captured using the Lens 1433 and Fig. 2b is the interference spectrum on vertical pixel of 1375. Generally, fast Fourier transform (FFT) will be used as spectral analysis of the spectrum, but the discrete spectrum information obtained by this method may generate large error in amplitude, frequency and phase due to energy leakage [12, 13].

Table 1 Parameters of lens

Sample model	f_n/mm	R/mm	Material
LA1131	50	25.8	N-NK7
LA1608	75	38.6	N-NK7
LA1433	150	77.3	N-NK7

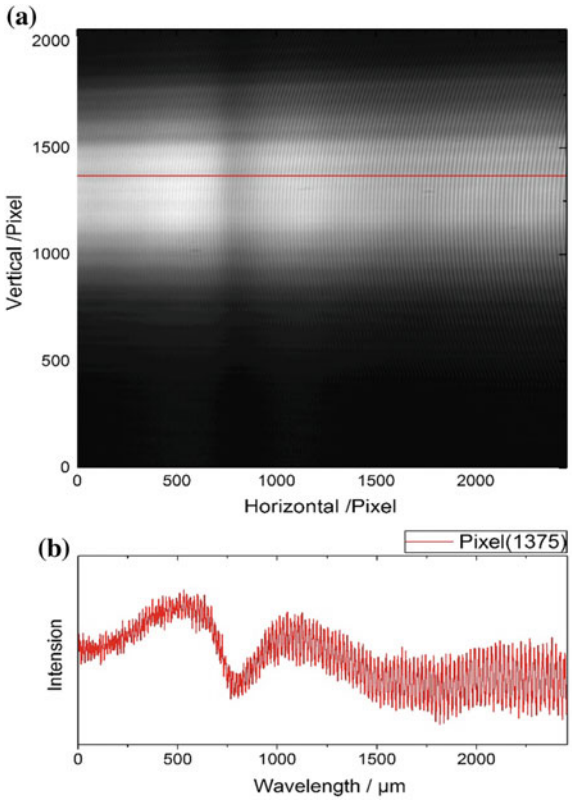


Fig. 2 **a** Interference fringes and **b** interference spectrum

In order to reduce the energy leakage effect due to the FFT computation, spectrum correction technique [12, 13] will used to correct the errors caused by leakage of energy. In this experiment, we used synthesized correction method [12] of spectrum phase difference to correct the error. We could improve the frequency resolution of Fourier transform and improve measurement accuracy. The main principle of this technique is that the phase difference $\Delta\varnothing$ can be obtained by the use of two consecutive signals after translation and adding different windows.

$$\Delta\varnothing = \varnothing - \varnothing_0 = -2\pi a_1 T(f_1 - \Delta f) \tag{8}$$

where $\varnothing$ and $\varnothing_0$ is the phase of two segment signal, T is the length of symmetric window. We can get a signal $x_0(t)$ by transfer a_1T of the continuous signal $x(t)$ ($a_1 > 0$). f_1 is the discrete spectral peak frequency before correction, Δf is the frequency correction. So we can obtain the following equation

$$\Delta f = \frac{-2\pi a_1 T f_1 - \Delta\varnothing}{2\pi a_1 T} \tag{9}$$

Figure 3 showed Lens surface curves obtained the developed LF-OCT with and without the spectrum correction technique. The solid line is much smooth than the dash line, which means that the result with spectrum correction method has better performance than the one without the correction method. Subsequently, the plano convex lens curvature radius could be achieved by doing curve fitting on the Lens surface curves obtained by the LF-OCT system. For the Lens LA1608, the curvature radius measured by the LF-OCT system is 77.327 mm.

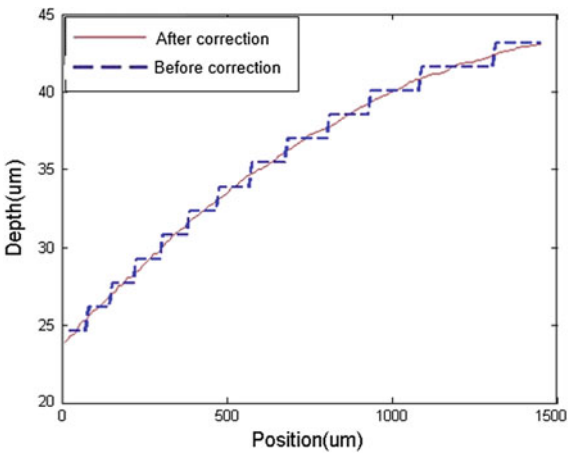


Fig. 3 Curve extraction and spectrum correction

Table 2 Experimental results

Sample model	Manufacturer data	Measured result
LA1131	25.8	25.785
LA1608	38.6	38.606
LA1433	77.3	77.327

All data units are *mm*

Table 2 showed the the curvature radius comparisons of manufacture data and the measured results for three Thorlabs Lens. As seen from the table, it was demonstrated the high precision performance of the developed LF-OCT system.

4 Conclusions

By using the developed LF-OCT system and the spectral correction technique, the curvature radius of lens with different focal length was measured precisely. The measured results are consistent with the actual dimensions provide by the manufacture. It was demonstrated that the LF-OCT has high accuracy and reliability and therefore it could be recommended to use in optical lens measurement due to its non-contact and non-invasion in nature.

Acknowledgements This work is supported by the Fujian Provincial Excellent Young Scientist Fund (2014J07007), the National Natural Science Foundation of China (51675103), the Training Program of Fujian Excellent Talents in Universities, the Specialized Research Fund for the Doctoral Program of Higher Education, the Ministry of Education, P. R. China (20133514110008) and the Ministry of Health, P.R. China (WKJ-FJ-27).

References

1. J. Hu, D. Xiang, P. Guo, L. Guan, Photoelectric experimental instrument for equal thickness interference. Coll. Phys. **23**, 43–45 (2004)
2. L. Lixin, Measurement of the Focal Length of the Concave Lens by Improved Autocollimation Method. J. Jishou Univ. **36**, 30–32 (2015)
3. X. Wang, Z. Wang, W. Zhao, L. Qiu, Uncertainty evaluation in laser differential confocal radius-of-curvature measurement. J. Appl. Opt. **35**, 155–158 (2014)
4. T. Zhihui, Z. Shi, W. Liu, H. Yang, Y. Sui, High-accuracy measurement for radius of curvature and its uncertainties. Opt. Precis. Eng. **21**, 2495–2501 (2013)
5. D. Huang, E.A. Swanson, C.P. Lin et al., Opt. Coherence Tomogr. **254**(5035), 1178–1181 (2005)
6. S. Zhong, L. HO, Y. Shen et al., Non-destructive quantification of pharmaceutical tablet coatings using terahertz pulsed imaging and optical coherence tomography. Opt. Lasers Eng. **49**(3), 361–365 (2011)
7. S. Zhong, Y. Yan, Y. Shen, Non-destructive testing of GFRP materials by Fourier-domain infrared optical coherence tomography. Int. Conf. Autom. Control Artif. Intell. 1407–1410 (2012)

8. S. Zhong, H. Shen, Y. Shen, Real-time monitoring of structural vibration using spectral-domain optical coherence tomography. Opt. Lasers Eng. **49**(1), 127–131 (2011)
9. Jianfeng Zhong, Shuncong Zhong, Qiukun Zhang et al., Two-dimensional optical coherence tomography for real-time structural dynamical characterization. Opt. Lasers Eng. **66**(3), 74–79 (2015)
10. S. Lawman, Y. Dong, B.M. Williams, V. Romano, S. Kave, S.P. Harding, C. Willoughby, Y. Shen, Y. Zheng, High resolution corneal and single pulse imaging with line field spectral domain optical coherence tomography. Opt. Express **24**(11), 12396 (2016)
11. T. You, S. Zhong, J. Zhong et al., Development of two-dimensional optical coherence tomography. J. Mech. Electr. Eng. **32**(5), 602–606 (2015). (in Chinese)
12. K. Ding, X. Zhu, M. Xie, S. Zhong, J. Luo, Synthesized correcting method of phase difference on discrete spectrum. J. Vibr. Eng. **15**(1), 118–122 (2002). (in Chinese)
13. K. Ding, S. Zhong, A universal phase difference correcting methods on discrete spectrum. pp. 142–145 (2003) (in Chinese)

Noninvasive Blood Pressure Measurement Based on Photoplethysmography

Yudong Wu and Shuncong Zhong

Abstract Stationary wavelet transform (SWT) and the Artificial Neural Networks (ANN) Algorithm were proposed for continuous measurement of blood pressure. A total number of 26,900 photoplethysmography (PPG) signals from Mimic Database were analyzed using SWT decomposition. The fifth layer high frequency coefficient was employed to reconstruct the signals. Ten characteristic parameters obtained from the reconstructed signals, as ANN input vector, and the PPG corresponding to the systolic blood pressure (SBP) and diastolic blood pressure (DBP), as the ANN vector output, were used to train the ANN model. The averaging error between the systolic pressure estimated by the proposed method and actual pressure value is 3.98 mmHg, and the standard deviation is 3.32 mmHg. The averaging error between the diastolic pressure estimated by the proposed method and actual pressure value is 3.81 mmHg, and the standard deviation is 3.63 mmHg. These results satisfy the American National Standards of the Association for the Advancement of Medical Instrumentation.

Keywords Blood pressure · Photoplethysmography · Stationary wavelet transform · The artificial neural networks

1 Introduction

Blood pressure is an important indicator of the physical condition of the human body. The level of blood pressure reflects the cardiovascular function. Hypertension is a common chronic disease, and it is also the main risk factor for cardiovascular disease. Therefore, continuous blood pressure monitoring is of great significance for the diagnosis and monitoring of the disease.

Y. Wu · S. Zhong (✉)
Laboratory of Optics, Terahertz and Non-Destructive Testing, School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350108, People's Republic of China
e-mail: zhongshuncong@hotmail.com

S. Zhong
Shanghai University, Shanghai, P. R. China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_13

In clinical, blood pressure measurement technology is divided into direct and indirect methods. The direct method is to insert a catheter with a pressure sensor into the main artery to detect the blood pressure signal. This method not only has high technical requirements, but also has a certain traumatic. Indirect methods include Offset sound method, oscillometry [1], pulse delay method and the pulse wave characteristic parameters method. Offset sound method is often used in clinical measurement. But it is susceptible to external noise. Most electronic sphygmomanometers use oscillometry to measure blood pressure. But, continuous measurement cannot be achieved by measuring the blood pressure with oscillographic. The pulse wave delay method is based on the relationship between the pulse transit time (PTT) and blood pressure [2–4]. The PTT is calculated from the ECG [5] signal and the PPG [6, 7] signal. This method requires the acquisition of two signals. So the measurement equipment is more complex. The principle of the pulse wave characteristic parameter method [8–11] is to calculate the blood pressure indirectly through the relationship between the characteristic parameters and blood pressure. The method is suitable for intelligent wear and other equipment to achieve continuous monitoring of blood pressure.

The core of the pulse wave characteristic parameter method is to extract the characteristic points of PPG signals accurately. However, clinical studies have shown the waveform characteristics of human pulse wave varying with the changes of peripheral vascular resistance, blood vessel wall elasticity and blood viscosity [12]. Fig. 1 shows the different types of PPG signals, from which we can see that

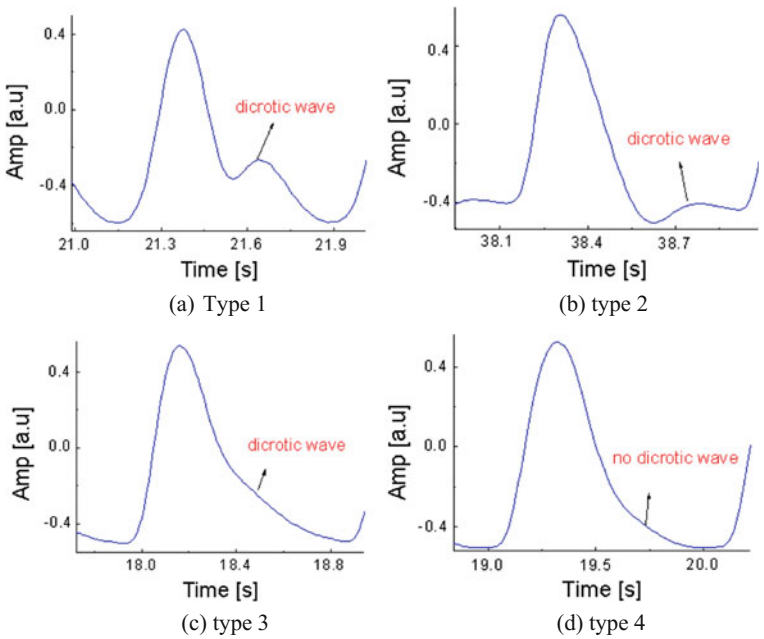


Fig. 1 Different types of PPG signals

different people's PPG signal is different. Especially, the dicrotic wave is not obvious. The change of waveform feature will affect the extraction of feature points, which will affect the calculation of blood pressure.

In the present work, different types of PPG signals were decomposed by SWT, and then the high frequency coefficients of each layer were reconstructed. We find that the fifth layer reconstruction signal (D5 layer) has the same characteristics. Therefore, we extract the feature parameters of the D5 layer as the ANN vector input to train the modal. The method can accurately identify the characteristic parameters and provide reference value for noninvasive continuous measurement of blood pressure.

2 Experimental Principle

2.1 Stationary Wavelet Transform Principle

The wavelet transform [13] is similar to the Fourier transform. Continuous wavelet transform is defined as:

$$WT_f(ab) = (f, \psi_{a,b}) = \int_{-\infty}^{+\infty} f(\lambda) \frac{1}{\sqrt{a}} \psi \bullet \left(\frac{\lambda - b}{a} \right) d\lambda \tag{1}$$

In the above formula, $\psi_{a,b} = \frac{1}{\sqrt{a}} \psi\left(\frac{\lambda - b}{a}\right)$ is the wavelet basis function. It is obtained by translating $(\lambda - \lambda/a)$ of a wavelet function $\varphi(\lambda)$ [14, 15].

The principle of Stationary Wavelet Algorithm:The first step it is to divide the signals of lengths N into two coefficients, which is the low frequency coefficient cA1 and the high frequency coefficient cD1. The former is obtained by convolution of S and low pass filter L_0_D. The latter is obtained by convolution with a high pass filter Hi_D. The length of cA1 and cD1 is N instead of N/2 in discrete wavelet transform. As show in Fig. 2:

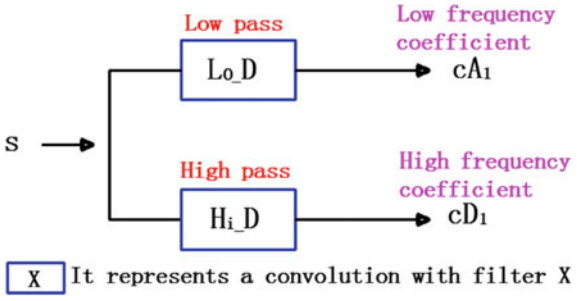


Fig. 2 SWT decomposition for the first step

The next step is to divide the low-frequency coefficients cA1 into two parts using the same method. But, in here, cA1 is used for instead of s. The new filter is interpolated by the previous filter. The SWT generates cA2 and cD2. As show in Fig. 3.

Here, different types of PPG signals were decomposed by SWT, and then the high frequency coefficients of each layer were reconstructed. We find that the fifth layer reconstruction signal (D5 layer) has the same characteristics. The Fig. 4 is relative to Fig. 1. According to Fig. 4, most of the useful high-frequency components of the original signal are distributed in the D5 layer. The stroke-by-beat signal of the D5 layer corresponds to the original PPG signal in the time scale. Some of dicrotic wave is not obvious in the original PPG signal, can be reflected in the D5 layer. Therefore, in this paper, D5 layer as the research object. We extract the characteristic parameters of D5 layer as the input of ANN.

Figure 5 shows the ten feature parameters extracted from the D5 layer. First of all, the aortic opening point A1, systolic maximum pressure point B1, the gorge point C1 and dicrotic wave point D1 [16–19] were calculated using the differential threshold method. Secondly, the 10 characteristic parameters were extracted from D5 layer. The amplitude parameter mainly includes the amplitude Hb of the point B1, the amplitude Hc of the point C1 and the amplitude Hd of the point D1. The time parameter mainly includes the time Tb of the point B1 to the point A1, the time Tc of the point C1 to the point A1 and the time Td of the replay of the point D1 to the next point A1. The area parameter mainly includes the area S from the point A1 to the point B1, the area S1 from the point B1 to the next wave starting point and the total area s. The ten characteristic parameters are used as vector input of ANN.

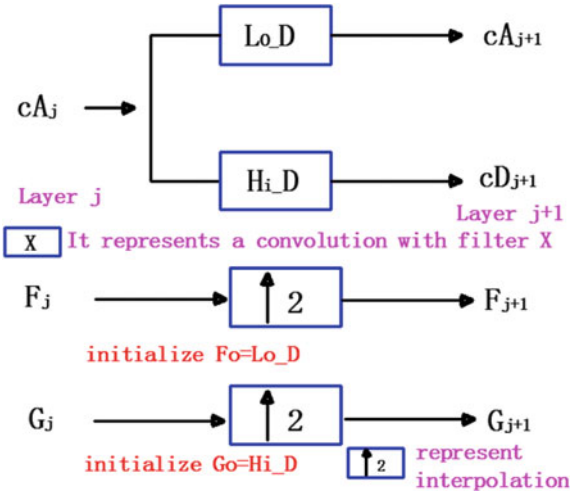


Fig. 3 SWT decomposition for the second step

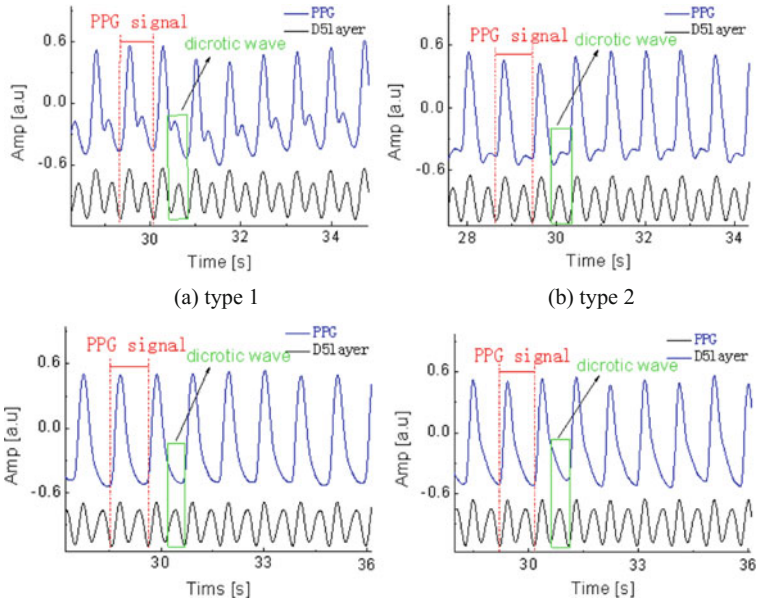


Fig. 4 Different characteristics of the PPG signals and the corresponding D5 layer

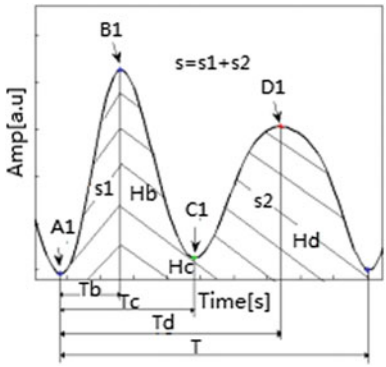


Fig. 5 Extraction of D5 layer signal characteristic parameters of 10

2.2 BP Neural Network Theory

As shown in Fig. 6 below, the BP neural network structure includes input layer, hidden layer and output layer. The hidden layer use sigmoid transfer function. The output layer was linear and the hidden layer was consists of 30 neurons. The Levenberg-Marquardt algorithm [20, 21] and the trainlm function were used to the network. The 10 characteristic parameters extracted by the D5 layer taken as the

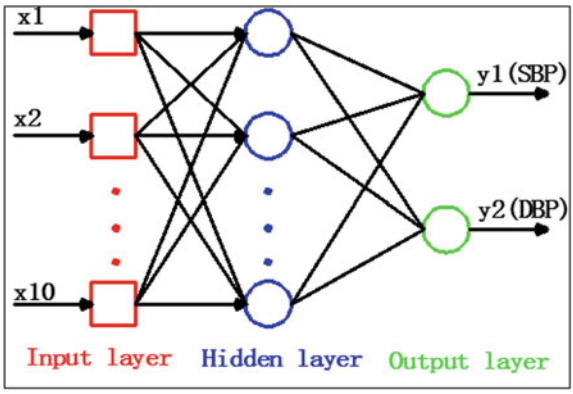


Fig. 6 The BP neural network structure

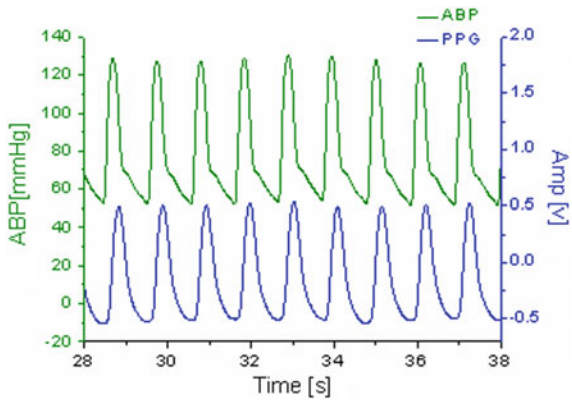


Fig. 7 The ABP signal of the database and the corresponding PPG signal

input of the input layer. SBP and DBP corresponding to PPG signal were used as output layer.

3 Experiment and Analysis

In the present work, a total of 26,900 PPG signals were calculated. These data were divided into two parts: training and testing. Training section contains 24,000 data and the test section contains 2000 data.

The PPG signal and the corresponding SBP and DBP used in the experiment were from THE MIMIC DATABASE [22–27]. The database records a variety of physiological signals and vital signs. These signals include electrocardiogram, blood pressure and PPG signals. Figure 7 shows the ABP signals and the corresponding PPG signal. Ten characteristic parameters is obtained from the D5 layer as the ANN input. Meanwhile, the SBP and DBP were recorded in the database as the ANN output to train model.

3.1 Model Training

During the training of ANN model, Training samples accounted for 70% of the whole data, verification samples and test samples each accounted for 15%. The performance of the trained model is determined by the error histogram and the R value of the training data.

Figure 8a shows the error histogram of the training sample, the verification sample, and the test sample. Error calculation formula is: error = target output − the actual output. As we can see, most of the error is centered around 0, which shows that the target output is close to the actual output. Figure 8b shows the regression model of the training sample, the verification sample and the test sample. As shown in the figure, the R values of the training samples, the verification samples, the test samples and all the data were 0.96614, 0.96415, 0.96008, 0.96493. According to the error histogram and the regression graph, we can see that the training model was well.

3.2 Model Testing

There are 2000 PPG signal data were used for the test estimation of the model. Figure 9 shows the line chart for the actual and estimated blood pressure values. Figure 9a shows the line graph of the SBP actual value and the estimated value calculated by the model. Figure 9b shows the line graph of the DBP actual value and the estimated value calculated by the model. It can be seen from Fig. 9 that the blood pressure estimated by the model is very close to the actual value.

Figure 10 shows the error histogram of SBP and DBP in Fig. 9.The error e is calculated by the following formula:

$$e = BP_{est} - BP \tag{2}$$

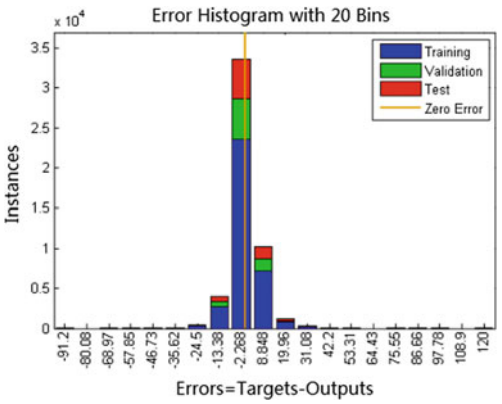
BP_{est} is the SBP or DBP estimated by the ANN training model. BP is the actual value in the corresponding database.

The statistics of the test results were shown in Table 1. The absolute error e_a calculated by the following formula:

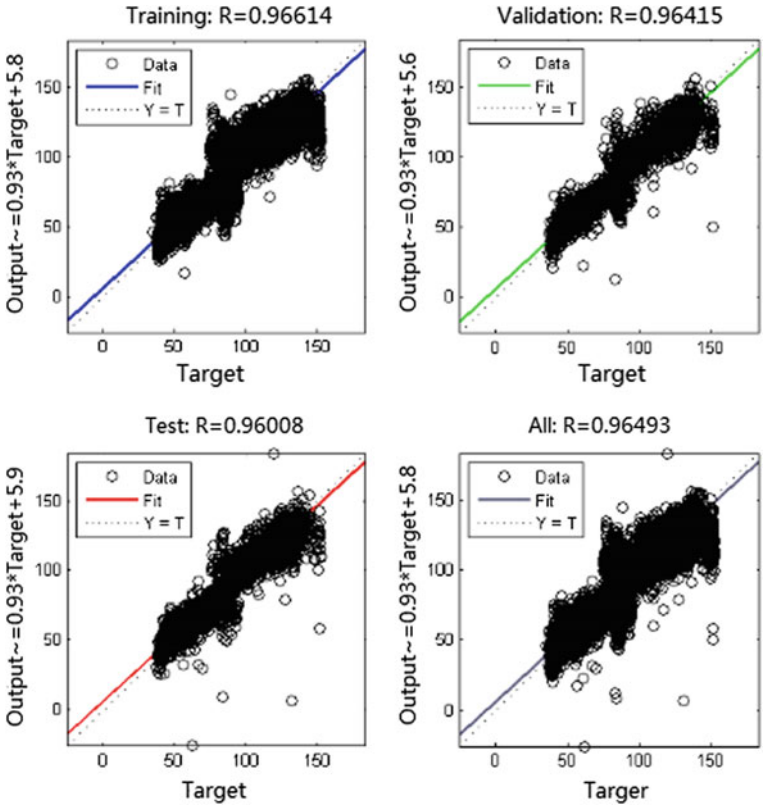
$$e_a = |e| \tag{3}$$

The mean value and standard deviation of the absolute error of e_a was used in Table 1 for performance analysis (Mean value ± Standard deviation).

According to Table 1, the averaging error between the systolic pressure estimated by the proposed method and actual pressure value is 3.98 mmHg, and the standard deviation is 3.32 mmHg. The averaging error between the diastolic pressure estimated by the proposed method and actual pressure value is 3.81 mmHg, and the standard deviation is 3.63 mmHg. These results satisfy the

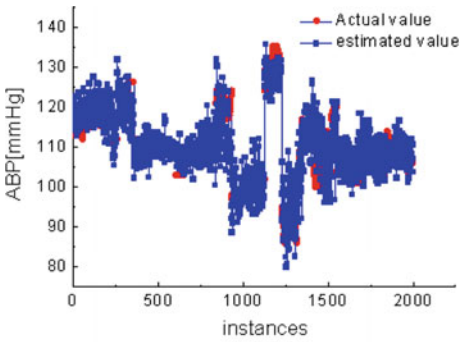


(a) Training data histogram

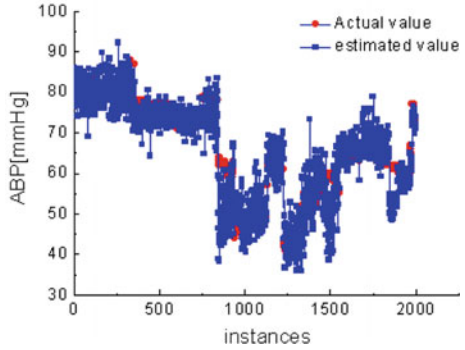


(b) Training data regression diagram

Fig. 8 Training evaluation model



(a) systolic blood pressure



(b) diastolic blood pressure

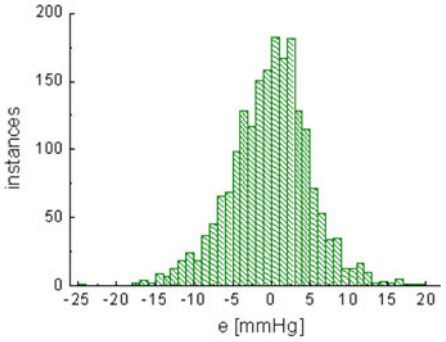
Fig. 9 Blood pressure line chart real and estimated values

American National Standards of the Association for the Advancement of Medical Instrumentation, where the maximal accepted errors is 5 ± 8 mmHg [28].

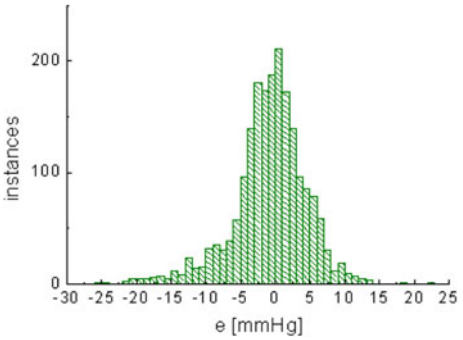
4 Conclusions

In this paper, an algorithm based on Stationary wavelet transform (SWT) and Artificial Neural Networks (ANN) was proposed to continuous and noninvasive blood pressure estimation from the PPG signal. The MIMIC database was used for training the ANN, and a total of more than 24,900 pulsations were analyzed by SWT to extract the 10 parameters. The obtained results 3.98 ± 3.32 mmHg for systolic and 3.81 ± 3.63 mmHg for diastolic pressure fulfills with the American National Standards of the Association for the Advancement of Medical Instrumentation, where the maximal accepted errors is 5 ± 8 mmHg. Therefore, this method can achieve non-invasive continuous blood pressure measurement.

Acknowledgements This work is supported by the Fujian Provincial Excellent Young Scientist Fund (2014J07007), the National Natural Science Foundation of China (51675103), the Training



(a) systolic blood pressure



(b) diastolic blood pressure

Fig. 10 Blood pressure error histogram

Table 1 Experiment and results

The number of PPG for training	The number of PPG per for test	SBP:e_a (mmHg)	DBP:e_a (mmHg)
24,900	2000	3.98 ± 3.32	3.81 ± 3.63

Program of Fujian Excellent Talents in Universities, the Specialized Research Fund for the Doctoral Program of Higher Education, the Ministry of Education, P. R. China (20133514110008) and the Ministry of Health, P.R. China (WKJ-FJ-27).

References

1. L. Wen, L. Gongzhi, C. Long, Design of electronic blood pressure monitor based on oscillographic method. Electron. Des. Technol. **21**(1), 1023–7364 (2014)
2. L. Xu, K. Gao, Continuous cuffless arterial blood pressure measurement based on PPG quality assessment. Int. J. Comput. Biol. Drug Des. **8**(2), 150–158 (2015)
3. L-h Bai, C. Wang, W. Miao et al., Continuous blood pressure monitoring system based on pulse wave conduction time. Biomed. Eng. Res. **33**(4), 221–225 (2014)
4. S. Wang, Pulse conduction time and systolic blood pressure correlation study. Shenyang: China Medical University (2013)

5. H.D. Lin, Y.S. Lee, B.N. Chuang, Using dual-antenna nanosecond pulse near-field sensing technology for non-contact and continuous blood pressure measurement, in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (IEEE, 2012), pp. 219–222
6. Z. Yurun, Z. Chen, J. Zhu, Pneumatic volume wave based wave breathing. J. Chin. Biomed. Eng. **32**(4), 508–512 (2013)
7. L.-h. Guo, Study on continuous blood pressure measurement of sleeves with PPG signals. Zhejiang University (2011)
8. XF Teng, YT Zhang, Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach, in *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, vol. 4 (IEEE, 2003), pp. 3153–3156
9. Y.Z. Yoon, G.W. Yoon, Nonconstrained blood pressure measurement by photoplethysmography. J. Opt. Soc. Korea **10**(2), 91–95 (2006)
10. G. Fortino, V. Giampa, PPG-based methods for non invasive and continuous blood pressure measurement: an overview and development issues in body sensor networks, in *2010 IEEE International Workshop on Medical Measurements and Applications Proceedings (MeMeA)* (IEEE, 2010), pp. 10–13
11. Z. Li, C. Wang, H. Zhu, et al., Non-invasive continuous blood pressure measurement based on photoplethysmography. J. Chin. Biomed. Eng. Soc. **31**(4): 607–614 (2012)
12. Z. Luo, S. Zhang, Y. Yang, *Engineering analysis and clinical application of pulse wave* (Science Press, Beijing, 2006)
13. J. Guo, W. Sun, Wavelet Analysis Theory and MATLAB7 Implementation. 261–267 (2005)
14. T. Roman, T. You, Q. Zhang et al., Time domain optical coherence tomography system based on stationary wavelet transform. J. Appl. Opt. **4**, 543–549 (2015)
15. Z. Wang, K-x Yin, H. Zhao et al., Pulsed signal denoising based on wavelet transform. Commun. Technol. **44**(5), 151–153 (2011)
16. M.-Y. Tang, K. Li, X.-T. Ma, Time domain feature extraction and algorithm of pulse wave signal, **4**, 15–17 (2010)
17. K. Xu, J. Wang, H. Yu, et al., Objective to study the correlation between pulse wave time-domain characteristics and blood pressure. Chin. J. Med. Device **8**, 42–45 (2009)
18. P. Shi, PDG based cardiovascular system function evaluation and application. Shanghai: Shanghai Jiaotong University (2010)
19. Zhu Zhong, Xu Liu, Research on pulse wave feature point recognition based on waveform and wavelet. J. Instrum. **37**(2), 379–386 (2016)
20. Zhou Pin, *Design and application of MATLAB neural network* (Tsinghua University Press, Beijing, 2013)
21. M. Chen, *Principles of MATLAB neural networks and examples of fine solution* (Tsinghua University Press, Beijing, 2013)
22. G.B. Moody, R.G. Mark, A database to support development and evaluation of intelligent intensive care monitoring. Computers in Cardiology (IEEE, 1996), pp. 657–660
23. K.V. Madhav, M.R. Ram, E.H. Krishna et al., Robust extraction of respiratory activity from ppg signals using modified mspca. IEEE Trans. Instrum. Meas. **62**(5), 1094–1106 (2013)
24. P. Li, M. Liu, X. Zhang et al., Novel wavelet neural network algorithm for continuous and noninvasive dynamic estimation of blood pressure from photoplethysmography. Sci. Chin. Inform. Sci. **59**(4), 1–10 (2016)
25. Y. Choi, Q. Zhang, S. Ko, Noninvasive cuffless blood pressure estimation using pulse transit time and Hilbert-Huang transform. Comput. Electr. Eng. **39**(1), 103–111 (2013)
26. X. He, R.A. Goubran, X.P. Liu, Secondary peak detection of PPG signal for continuous cuffless arterial blood pressure measurement. IEEE Trans. Instrum. Meas. **63**(6), 1431–1439 (2014)
27. O. Sayadi, M.B. Shamsollahi, Utility of a nonlinear joint dynamical framework to model a pair of coupled cardiovascular signals. IEEE J. Biomed. Health Inform. **17**(4), 881–890 (2013)
28. American National Standard for Electronic or Automated sphygmomanometers, *ANSI/AAMI SP10, Arlington* (Association for the Advancement of Medical Instrumentation, VA, 1992)

Terahertz Chiral SSPPs Mode on the Helically Grooved Metal Wire

Haizi Yao and Shuncong Zhong

Abstract Terahertz wave (λ = 30 μm–3 mm@f = 0.1–10 THz) is such a band of electromagnetic wave that locates between light and microwave. Many chiral biomolecules, such as protein and amino acid, have their own unique structure vibrating frequencies at Terahertz region, therefore they have the "fingerprint" terahertz absorption spectrum. It is meaningful to study the chirality of chiral macromolecules through their terahertz circular dichroism (CD) response which is defined as the differential absorption of two opposite chiral terahertz wave. In the present work, we used the Finite-element-method (FEM) to investigate terahertz spoof surface plasmon polariton (SSPP) wave propagating on the helically grooved metal wire. Different from the metal wire decorated by the periodic grooves the helically grooved metal wire can support a kind of chiral SSPP surface mode. Due to the chirality of helical grooves, the normal degenerate HE1 SSPP mode on helically grooved metal wire decompose into two opposite chiral modes, i.e., HE−1 and HE+1 SSPP surface modes. Mode analysis showed that the dispersion curves of HE+1 mode deviated from that of HE−1 modes, and thus form a broad frequency band where the chiral HE+1 mode could exists exclusively. Super-chiral terahertz field concentration that may be achieved on the helically grooved metal wire through chiral HE+1 mode can find important applications in near-field circular dichroism spectroscopy in terahertz frequencies.

Keywords Terahertz circular dichroism · Terahertz spoof surface plasmon polariton

H. Yao · S. Zhong (✉)
Laboratory of Optics, Terahertz and Non-Destructive Testing, School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350108, People's Republic of China
e-mail: zhongshuncong@hotmail.com

S. Zhong
Fujian Key Laboratory of Medical Instrument and Pharmaceutical Technology, Fuzhou 350108, People's Republic of China

S. Zhong
Shanghai University, Shanghai, P. R. China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_14

1 Introduction

As the collective excitations of conduction electrons that exist on the metal surface, surface plasmon polaritons (SPP) enable light waves to be localized into sub-wavelength region [1], giving a new approach to manipulate the light. Terahertz wave (λ = 30 μm–3 mm@f = 0.1–10 THz) is a special band of electromagnetic wave between light wave and microwave, having many important applications in a variety of field [2]. In the THz frequencies, most noble metal show the characteristics of a perfectly electric conductor (PEC), thus the flat metal surface does not support SPP. In 2004, Pendry et al. [3] theoretically demonstrated that the dispersion of the surface electromagnetic modes on metal surface can be engineered by texturing metal surface. They showed that the effective plasma frequency of the textured surface can be lowered and was determined completely by the geometry of the metal surface; so the structured PEC surfaces could also support surface modes resembling SPP of light, namely spoof SPPs (SSPPs) [4]. SSPP facilitates an extension of the SPP for subwavelength confinement and enhancement of electromagnetic wave in microwave and terahertz wave. The capability of confining electromagnetic wave makes SSPP important in terahertz waveguide applications. A variety of plasmonic terahertz waveguides based on SSPP have been proposed, such as ultrathin and flexible films [5, 6], metal surfaces textured by arrays of grooves or pillars [7–9] and so on. Comparing to the planar waveguide, there has been an increasing interest in cylindrical wire waveguide of terahertz [10–12] since it can give rise to focusing of the electric field that may has potential application in the terahertz near-field microscope. However, in these works almost all the discussions only concern axisymmetric structure and fundamental waveguide mode, and the complicated waveguide structure and high-order SSPP modes had been rarely reported. As an example of non-axisymmetric structure, the helically grooved wire has been demonstrated to be able to guide the SSPP [13]. Recently, another great work demonstrated that the helically grooved metal wire could support a kind of chiral surface wave with tuneable orbital angular momentum [14], this interesting finding can find a potential for some special application for THz.

In the present work, we had a further discussion about the chiral SSPP on the helically grooved metal wire on the basis of the former literature [14]. we used the Finite-element method (FEM) to investigate waveguide mode of terahertz SSPP propagating on the helically grooved metal wire. The results showed that for the metal wire decorated by helical groove, the dispersion curve of initial degenerate HE1 waveguide mode begin to divaricate and finally form two non-degenerate modes, HE−1 mode and HE+1 mode. Importantly, the non-degenerate HE−1 (HE +1) mode is inherently chiral, i.e., circularly polarized. The chiral SSPP mode on helically grooved metal wire can find important applications in terahertz near-field spectroscopy. For example, the helically grooved wire can be used as the probe to provide the strong chiral terahertz wave for the circular dichiroism detecting to the chiral biomolecules.

2 Terahertz Chiral SSPPs Mode

The sketch of helically grooved metal wire waveguide are shown in Fig. 1a. A metal wire of radius R is corrugated by a right-handed helix with pitch Λ and cross section radius r. To investigate the influence of helical groove on the waveguide, we used the FEM method (using RF module of COMSOL Multiphysics) to obtain the dispersion relation curves of different propagating modes, as shown in Fig. 1b. In calculation, R, r, Λ were set to 60, 25 and 100 μm, respectively, and the metal was modeled as perfect electric conductor.

From the dispersion curves, it is obviously seen that three curves of modes are at right of the light line, that indicate that they have a lager propagating wave number and these three modes are the slow surface wave. Besides, the primary degenerated HE1 mode split into two separated curves denoted by HE−1, HE+1 modes; and there emerges a degeneracy between HE−1 mode and fundamental TM0 mode at the band edge. Figure 2a show the mode distributions of TM, HE−1 and HE+1 modes on cross section at $k_z = \pi/\Lambda$. We can see that due to the degeneracy TM and HE−1 modes give an similar result with one maximum-field node, whereas the HE +1 mode have three nodes in circumferential direction.

In general, for the normal modes of a axial symmetry waveguide the number of azimuthal maximum nodes of modal field are always an even value because the nodes are the double of mode index (integral number). However, helical grooved

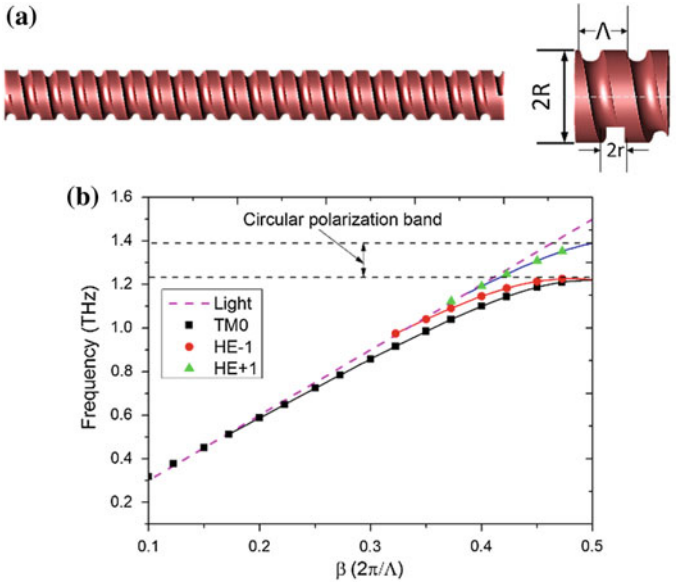


Fig. 1 **a** Sketch of helically grooved wire waveguide, and the dispersion relation curves of the SSPP modes supported by the right-handed helically grooved wire waveguide. **b** The horizontal dashed lines are to denote the band of chiral mode, HE+1

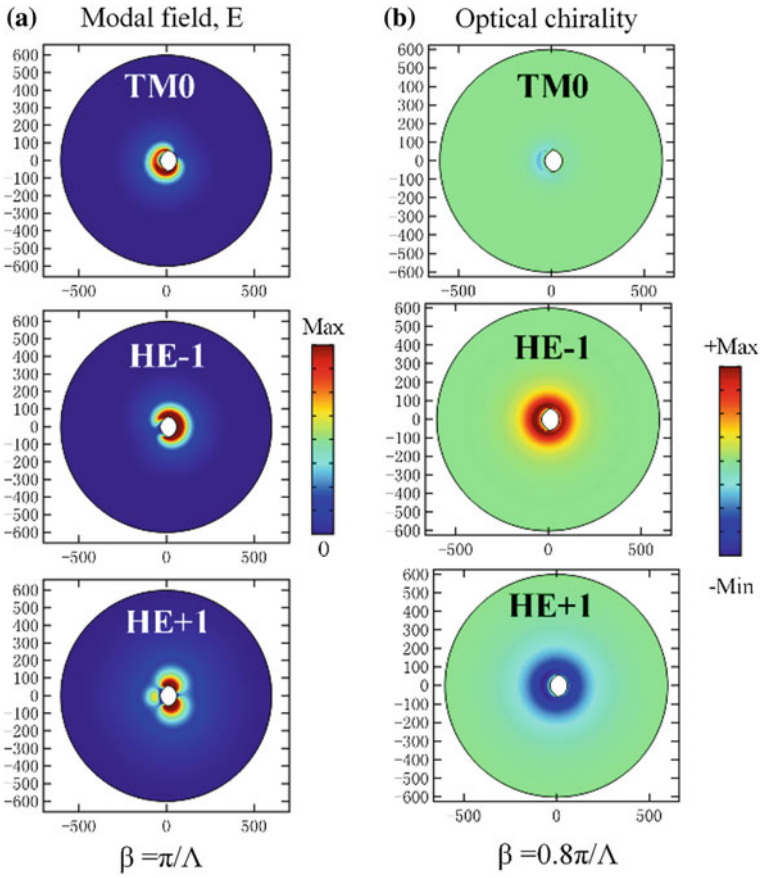


Fig. 2 **a** The modal field distribution of different SSPPs modes at band edge, $\beta = \pi/\Lambda$; **b** the "optical chirality" distribution of three SSPPs modes at $\beta = 0.8\pi/\Lambda$

wire waveguide has a different performance and the nodes of modal field are not rigid even number. Such the anomaly in azimuthal nodes is the consequence of the structural chirality. That can be explained by the characteristic of electromagnetic fields on the helical structure. All of field components in helical geometry can be expanded in terms of the eigen modes as [15]:

$$F_n(r, \phi, z) = A_n(r, z)\mathrm{e}^{i(\pm\tilde{n}\phi + \beta z - \omega t)} \tag{1}$$

$$\tilde{n} == n + \beta\frac{\Lambda}{2\pi} \tag{2}$$

where A_n is the modal amplitude,$\tilde{n}$ is the effective mode index, n is the integer denoting the mode order,β is the wave number of the propagating wave, and Λ is the period of the helical geometry. Equation (1) states that the electric field is of azimuth dependence, which is controlled by the factor $e^{i(\pm\tilde{n}\phi)}$. In Eq. (2), the effective mode index is a function of wave number, implying that the number of the azimuthal nodes will evolve from an even value to an odd value as the wave number increases from zero to the band edge, $\beta = \pi/\Lambda$. This is different from the case of the axial symmetry waveguide, whose electric field always appears as even nodes on its azimuthal distribution.

From the Eq. (1), we can see that for the high order mode with a mode index larger than zero, azimuthal angle θ in the phase will result in electric field rotation [16]. In fact, the HE+1 and HE−1 modes are the helical mode and the field rotation make them possess chirality. Commonly, in the case of azimuthal symmetry, the high order mode, HE1, is degenerate and it does not present any chirality. In contrast, for the helically grooved metal wire, the break of azimuthal symmetry lead to the degeneracy of HE1 modes broken. As a result, the two opposite chiral components of HE1 mode decompose, as seen in their dispersion curves. For the electromagnetic field, the "optical chirality" [17, 18] is usually used to evaluate the light chirality, which is defined as:

$$C = -\frac{\varepsilon_0 \omega}{2} \mathrm{Im}(E^* \cdot B) \tag{3}$$

where ε and ω are the permittivity of free space and angular frequency respectively, and E and B are the complex electric and magnetic fields. To demonstrate the difference in "optical chirality "of the HE+1 and HE−1 SSPP, here we mapped the distribution of the "optical chirality" on the cross section plane at $\beta = 0.8\pi/\Lambda$, showed in Fig. 2b. It can be seen that as expected the decomposed HE+1 mode has opposite optical chirality comparing to the HE−1 mode; besides, the primary TM0 mode of no chirality begin to has possess optical chirality due to the chirality of structure. The chiral SSPPs wave is mostly confined close to the surface of metal structure. Such the localized and enhanced chiral field is very valuable to enhance the interaction between terahertz and sample and thus, to improve the circular dichroism (CD) detecting for the chiral analyte. More importantly, the divergence between the dispersion curves of HE+1 and HE−1 modes just form a band, denoted by the horizontal dashed lines in Fig. 1b, within which the HE+1 mode could exists exclusively. This is to say, we can utilize the helically groove metal wire to obtain a broad band of sign-uniform chiral wave (HE+1 mode), that is important for future THz CD detecting. It is worth stressing that the desired frequency of chiral mode band can be controlled by adjusting the geometry size. Such the characteristic of selective frequency make this structure apply to various detecting sample through adjusting the waveguide geometry.

3 Conclusions

In summary, we have demonstrated a helically grooved metal wire for terahertz waveguide. It was found that due to the symmetry break introduced by the helical groove the primal degenerated HE1 SSPPs mode on metal wire breaks up into two opposite chiral modes, HE+1 and HE−1 mode. The diverge between dispersion curves of two modes gives rise to a band of chiral SSPP mode. This distinct advantages of supporting sign-uniform chiral SSPPs make such the helically grooved metal wire potential in terahertz circular dichroism detection for chiral bio-macromolecules.

Acknowledgements This work is supported by the Fujian Provincial Excellent Young Scientist Fund (2014J07007), the National Natural Science Foundation of China (51675103), the Training Program of Fujian Excellent Talents in Universities, the Specialized Research Fund for the Doctoral Program of Higher Education, the Ministry of Education, P. R. China (20133514110008) and the Ministry of Health, P.R. China (WKJ-FJ-27).

References

1. H. Raether, *Surface plasmons on smooth surfaces* (Springer, New York, 1988)
2. M. Tonouchi, Cutting-edge terahertz technology. Nat. Photonics **1**(2), 97–105 (2007)
3. J. Pendry, L. Martin-Moreno, F. Garcia-Vidal, Mimicking surface plasmons with structured surfaces. Science **305**(5685), 847–848 (2004)
4. F.J. Garcia-Vidal, L. Martín-Moreno, J.B. Pendry, Surfaces with holes in them: new plasmonic metamaterials. J. Opt. A: Pure Appl. Opt. **7**(2), S97–S101 (2005)
5. X. Shen, T.J. Cui, D. Martin-Cano, F.J. Garcia-Vidal, Conformal surface plasmons propagating on ultrathin and flexible films. Proc. Natl. Acad. Sci. **110**(1), 40–45 (2013)
6. X. Shen, T.J. Cui, Planar plasmonic metamaterial on a thin film with nearly zero thickness. Appl. Phys. Lett. **102**(21), 211909 (2013)
7. Z. Gao, L. Shen, J.-J. Wu, T.-J. Yang, X. Zheng, Terahertz surface plasmon polaritons in textured metal surfaces formed by square arrays of metallic pillars. Opt. Commun. **285**(8), 2076–2080 (2012)
8. A. Rusina, M. Durach, M.I. Stockman, Theory of spoof plasmons in real metals. Appl. Phys. A **100**(2), 375–378 (2010)
9. F.M. Zhu, Y.Y. Zhang, L.F. Shen, Z. Gao, Subwavelength Guiding of Terahertz Radiation by Shallowly Corrugated Metal Surfaces. J. Electromagn. Waves Appl. **26**(1), 120–129 (2012)
10. L. Shen, X. Chen, Y. Zhong, K. Agarwal, Effect of absorption on terahertz surface plasmon polaritons propagating along periodically corrugated metal wires. Phys. Rev. B **77**(7), 075408 (2008)
11. A. Fernández-Domínguez, L. Martín-Moreno, F. García-Vidal, S.R. Andrews, S. Maier, Spoof surface plasmon polariton modes propagating along periodically corrugated wires. IEEE J. Sel. Top. Quantum Electron. **14**(6), 1515–1521 (2008)
12. S.A. Maier, S.R. Andrews, L. Martin-Moreno, F. Garcia-Vidal, Terahertz surface plasmon-polariton propagation and focusing on periodically corrugated metal wires. Phys. Rev. Lett. **97**(17), 176805 (2006)
13. A.I. Fernandez-Dominguez, C.R. Williams, F.J. Garcia-Vidal, L. Martin-Moreno, S.R. Andrews, S.A. Maier, Terahertz surface plasmon polaritons on a helically grooved wire, Appl. Phys. Lett. **93**(14) (2008)

14. F. Rüting, A. Fernández-Domínguez, L. Martín-Moreno, F.J. García-Vidal, Subwavelength chiral surface plasmons that carry tuneable orbital angular momentum. Phys. Rev. B **86**(7), 075437 (2012)
15. P. Crepeau, P.R. McIsaac, Consequences of symmetry in periodic structures. Proc. IEEE **52** (1), 33–43 (1964)
16. H.-Z. Yao, S. Zhong, Wideband circularly polarized vortex surface modes on helically grooved metal wires. IEEE Photonics J. **7**(6), 1–7 (2015)
17. Y. Tang, A.E. Cohen, Enhanced enantioselectivity in excitation of chiral molecules by superchiral light. Science **332**(6027), 333–336 (2011)
18. Y. Tang, A.E. Cohen, Optical chirality and its interaction with matter. Phys. Rev. Lett. **104** (16), 163901 (2010)

Simulation and Experiment of Terahertz Wave Detection of Marine Protective Coating

Wanli Tu and Shuncong Zhong

Abstract Terahertz detection of multilayer marine protective coatings on a steel substrate using terahertz pulsed imaging (TPI) system in reflection mode was investigated. The multilayered anticorrosive/antifouling paint on steel substrate with different coating structures were produced as experimental samples for terahertz wave detection. The modeling and numerical simulation to calculate the propagation of THz radiation from marine protective coatings was studied by the finite difference time domain (FDTD) method. The reflected radiation properties in both simulation and experimentation were recorded for further analysis, such as the identification of the defect beneath coating. The results showed that it was able to provide much more detailed information on the position of the defect beneath the coating, such as detachment of large areas where adhesion has failed, or the blistering defect beneath the coating. Therefore, the developed model combined with the experimental analysis was effective and feasible for the detection and structure characterization of marine protective coatings and will provide timely and reliable information for ship maintenance work.

Keywords Terahertz pulsed imaging · Marine protective coatings

W. Tu
Marine Engineering Institute, Jimei University, Xiamen 361021, People's Republic of China

W. Tu · S. Zhong (✉)
Laboratory of Optics, Terahertz and Non-Destructive Testing, School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350108, People's Republic of China
e-mail: zhongshuncong@hotmail.com

S. Zhong
Shanghai University, Shanghai, P. R. China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_15

1 Introduction

For marine protective coatings, monitoring of the corrosion degradation of paint layers can provide timely and effective maintenance to avoid serious consequences due to coating failure [5]. Generally, the corrosion processes are usually monitored by using conventional electrochemical techniques, such as electrochemical impedance spectroscopy (EIS), electrochemical noise measurements (ENM), which allow both the evaluation of corrosion rates and the identification of corrosion mechanisms. However, for more localized corrosion, such as pitting, detachment and blistering, etc., EIS was difficult to give the local corrosion inspection and ENM data can be affected by extraneous signals which make it not so suitable for monitoring or quality control [1, 2]. Other commercially available thickness meters for protective coatings such as contact detection (ultrasonic testing [3] and eddy-current testing [4], etc.), optical probing method (infrared thermal wave detection [5], etc.) also have limitation in application. Ultrasonic testing requires smooth sample surface, eddy-current testing was easily influenced by the medium which makes it difficult to make reliable diagnosis for defect, and infrared thermal wave detection has low detective accuracy for micro defects [5]. Thus there is considerable need for developing new method for nondestructive inspection of localized defect because it is difficult to detect with the conventional thickness meter.

So far, TD-THz technology provides the capability for stand-off inspection of coatings that are otherwise opaque at visible and infrared wavelengths. The utility of time-domain terahertz technology for automobile paint thickness measurements has previously been demonstrated in the laboratory by T. Yasui et al. [6]. David J. Cook and coworkers were developing a standoff sensor for the real-time thickness measurement of wet (uncured) marine paints with antifouling paint layers + anticorrosive paint layers + steel substrate structure for the purpose of providing feedback to an automated system for painting ships in dry-dock. However, those studies were not cover the embedded defects detection and feature extraction for multilayered medium [7]. T. Kurabayashi et al. studied corrosion detection under paint films using a specific frequency of THz-wave in THz transmission imaging system. They found the mechanical defects and thru holes were visualized, but corrosion areas were not [8].

In this paper, different protective coating samples with various paint layer thickness were investigated by an experimental TPI system in reflective mode. And multilayered samples to mimic the occurrence of different kinds of defects (blister, detachment, etc.) beneath the coating were also detected and analyzed. Furthermore, various coating structures were modeled and computed using the FDTD method with the broad-band terahertz radiation covering a certain spectral range. Both the reflected radiation properties in experimentation and simulation were recorded to be further analyzed, such as extracting the impulse function, and evaluating the defect position.

2 Experimental Studay of Terahertz Wave Propagation in Marine Protective Coatings

Figure 1(a), (b) showed an experimental coating sample and the diagram of terahertz wave interacted with the marine protective coating, respectively. The samples consisted of a steel panel (10 mm × 10 mm) coated with layers of protective paint to mimic the marine protective coating below the waterline with antifouling paint layers (SeaMate) + link coat (5BP Safeguard Plus) + anticorrosive paint layers (Jotaprime 500) + steel substrate structure. Most paints are based on organic solvents. When a terahertz pulse is incident on the layers of paint over a reflective (i.e., metal) substrate, the reflections from the air/coating interface, internal interfaces existing at different paints and the coating/steel interface are indicted. The chemical and structural information in the medium would be encoded into the reflected THz waveform because it causes changes in the refractive index or the optical absorption coefficient of each coating layer. As a result, the structural information of a sample can be ultimately extracted by analysis of the recorded THz wave in the time domain.

For all TPI measurements, a reflect-mode TAS7500IM system (Advantest Corporation, Japan) was used. The terahertz radiation was broadband, covering a spectral range of 75–300 μm (0.1–4 THz). Figure 1(c) showed the reference signal and the measurement signal denoted reflection signals from the metal mirror and coated sample, respectively. The signal had been normalized to the maximum reference signal. Impulse function was often extracted for quantitative non-destructive evaluation of multilayer structure, such as determining the

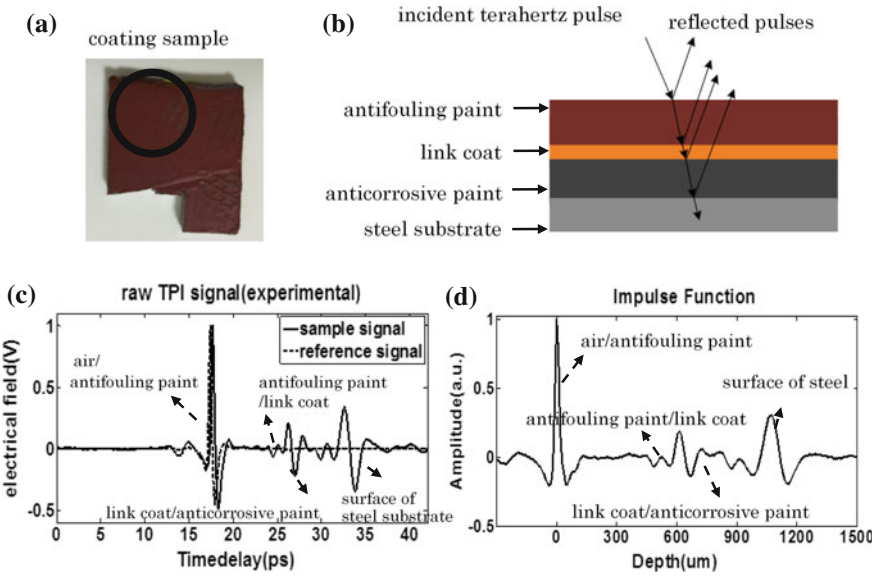


Fig. 1 TPI experimental test result of marine protective coating. **a** detected coating sample; **b** the diagram of terahertz wave interacted with the marine protective coating; **c** raw terahertz detected signal and reference signal; **d** impulse functions of detected sample

thickness of the coating layer. It can be obtained by deconvolution process, which divides the raw terahertz waveform reflected from a coated substrate by the reference signal in the frequency domain [9]. Figure 1(d) was the corresponding impulse function of the sample signal in Fig. 1c, where the horizontal axis represented the direction of the thickness of the sample and a double Gaussian filter was applied to remove high- and low-frequency components and also to produce a suitable time-domain response in the process of deconvolution.

From the figures, due to the differences in refractive indices of the multiple layers structure, reflections from air/coating surface and the coating/steel interface can be clearly resolved, along with subsequent reflections associated with internal interfaces at the air/coating boundary. The refractive index of antifouling paint can be obtained through the amplitude of the reference signal and sample signal, antifouling paint layers thickness and time delay [10]. It was calculated as 1.98. Besides, the value of reflection peak caused by the interface between two different kinds of material was associated with their difference in refractive index, thus it can be referred that the refractive index of link coat was little smaller than that of antifouling paint and much smaller than that anticorrosive paint.

During the corrosive deterioration process, the chemical and physical characteristics changes in the coating system and thus forms various defects, such as blistering, rust, detachment, cracking, etc. The blistering defect and detachment defect were generally caused by poor painting quality, contamination on the painting substrate, or deterioration by aging. Here, the coating samples with various defects beneath the coating were also detected by the experimental TPI system. Figure 2(a-1), (b-1) showed detected samples with blister defect and detachment defect within the coating, respectively; Fig. 2(a-2), (b-2) were the comparison of raw terahertz detected signals of the corresponding coating samples with and without defect; And Fig. 2(a-3), (b-3) were the corresponding impulse functions comparison of the TPI signals. The experiment results indicated that terahertz wave has a good carrying capacity of information. The defects would be distinguished and analyzed according to the time frequency analysis of detection signal and reference signal. As can be seen from the figures, the blistering defect and the detachment defect beneath the coating can be easily extinguished.

3 Simulation Study of Terahertz Wave Propagation in Marine Protective Coatings

For TPI in which a transient terahertz pulse is used, a time-domain method, such as the finite difference time domain (FDTD) method, is more applicable than a frequency-domain method because a differential equation is simpler to solve than an integral equation. The FDTD algorithm is a proven numerical method to model electromagnetic scattering problems, which solves the Maxwell's equations directly and obtains the solution of the electric field of electromagnetic waves. It analyses continuous electromagnetic problems by using finite difference and obtains the

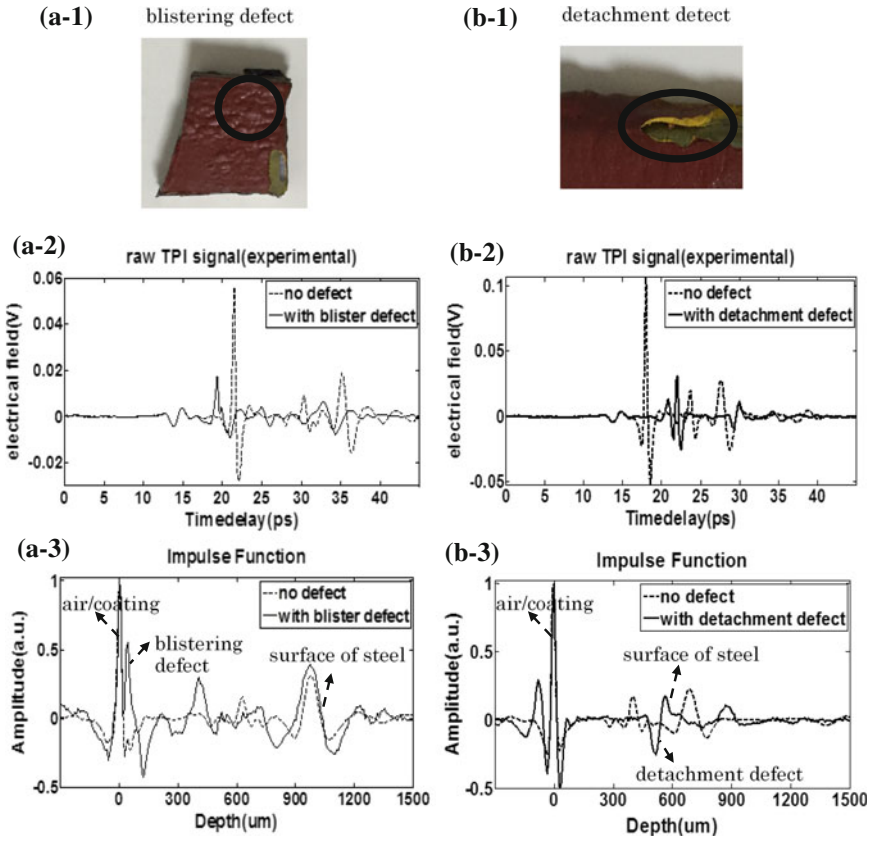


Fig. 2 TPI experimental test results of marine protective coating with defect. (**a-1**) ~ (**a-3**) detected sample with blister defect within the coating, raw terahertz detected signals of coating samples with and without blister defect, impulse functions of coating samples with and without blister defect; (**b-1**) ~ (**b-3**) detected sample with detachment defect beneath the coating, raw terahertz detected signals of coating samples with and without detachment defect, impulse functions of coating samples with and without detachment defect

electric field value at the sampling point. The FDTD method offers several advantages, such as robustness and the ability to study dispersive, nonlinear, or anisotropic materials [11, 12]. In our previous works, a theoretical model based on the FDTD method for marine protective coating with antifouling paint layers + anticorrosive paint layers + steel substrate structure was investigated [13, 14]. It has been demonstrated that FDTD model could be capable of calculating the propagation and reflection of THz radiation from a multilayered flat or curved structures.

All FDTD simulations here were performed using the Remcom XFDTD commercial software package and the model based on the experimental sample mentioned in Sect. 2. The incident terahertz plane wave was set perpendicular to the sample surface and a Gaussian profile of the amplitude distribution of the E-field

was initiated in the incident plane with peak amplitude of 1 V/m and with a spectral coverage range up to 1.5 THz. In simulation, a cylinder plate of 10 mm × 10 mm size with 0.2 mm thickness was taken as steel substrate, and the paint layers were successively coated on the steel substrate to mimic the coating samples. The refractive indices of paint layers were set as: $n_{\text{antifouling paint}} = 1.98$, $n_{\text{link coat}} = 1.8$, $n_{\text{anticorrosive paint}} = 2.2$. In our model, the protective coating was treated as a linear isotropic homogeneous layer without magnetic field loss. For simplicity, the dispersive properties and the specific conductivity of coatings are negligible as well [15]. The models were calculated and the simulated results were employed for the extraction of the impulse function. Figure 3 showed the FDTD simulation results of protective coatings structure with three layers of antifouling paint + one layer of link coat + two layers of anticorrosive paint + steel substrate.

Figure 3(a) was the simulated model; Fig. 3(b) was the FDTD calculated results of the corresponding coating model which were collected by the plannar sensor. From previous work using TPI for coating analysis, it is well known that scattering losses lead to significant distortion in the reflected terahertz pulse; hence, the waveforms close to sharp edges are routinely removed before quantitative coating analysis [16]. Figure 3(c) was the calculated TPI signal which was collected by the far zone sensor and represented the raw terahertz detected signal where Gaussian white noise was added in a SNR of 32 dB [17]; Fig. 3(d) was the corresponding impulse functions, where the double Gaussian filter was applied in deconvolution process. The dominant peak in simulated signal represented the reflection caused by coating/steel interface for its strong reflection, while the second largest peak referred to the signal reflected by air/coating interface. That was mainly because the

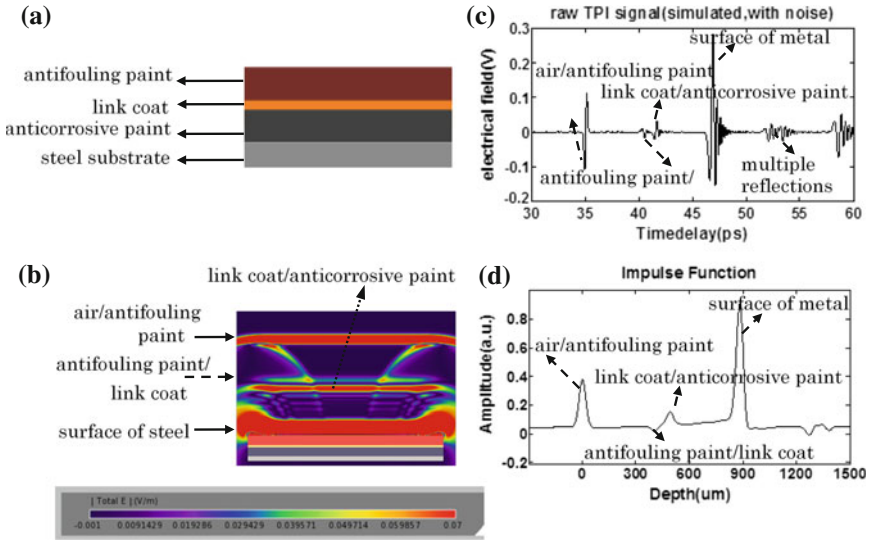


Fig. 3 FDTD calculated results of the raw terahertz detected signal of coated sample with three layers of antifouling paint/one layer of "link coat"/two layers of anticorrosive paint/steel substrate. **a** simulated coating sample; **b** simulated result collected by the plannar sensor; **c** TPI signal collected by the far zone sensor; **d** impulse functios of detected sample

dispersive properties and the specific conductivity of coatings were negligible in simulation, which lead to that the wave reflected by coating/steel interface has great energy reservations. However, it cannot be ignored for the actual material whose TPI testing result was shown in Fig. 1, which resulted in that the paint layers has reflected off lot of energy and the signal reflected by air/coating interface was much

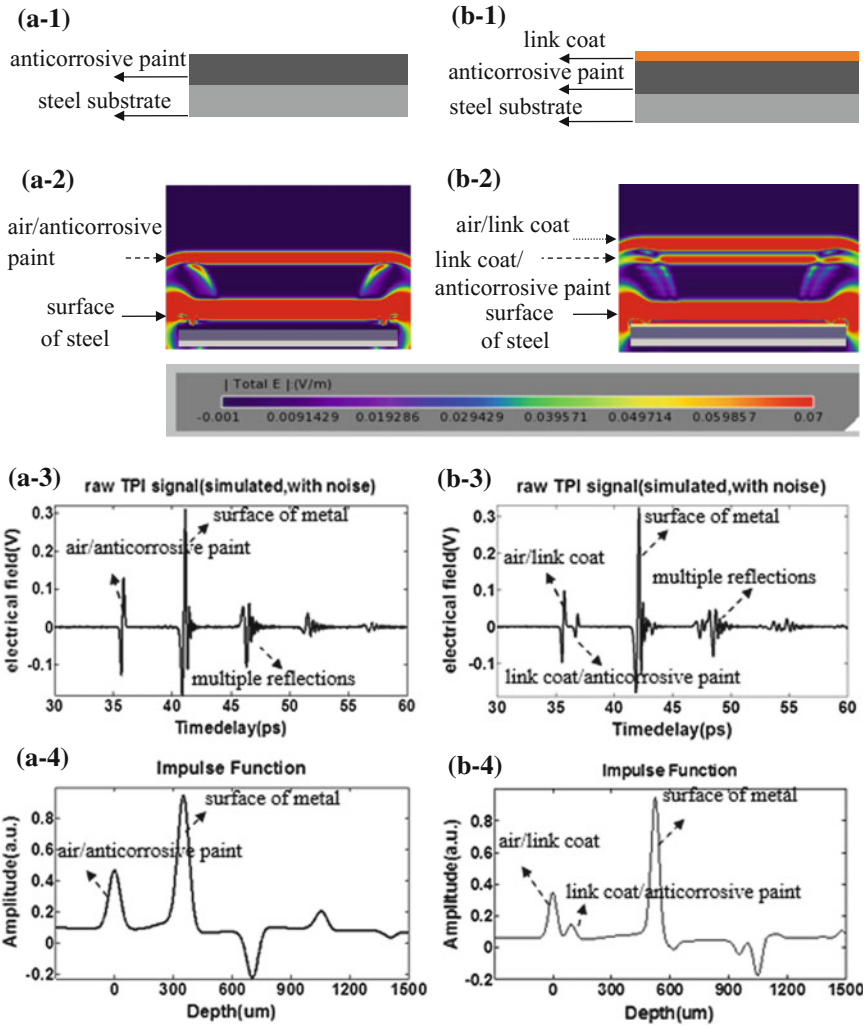


Fig. 4 FDTD calculated results of TPI propagation with various coating sample. (**a-1**) ~ (**a-4**) The simulated coating sample, simulated result collected by the plannar sensor, TPI signal collected by the far zone sensor and impulse function of detected sample of coating structure with two layers of anticorrosive paint + steel substrate; (**b-1**) ~ (**b-4**): The simulated coating sample, simulated result collected by the plannar sensor, TPI signal collected by the far zone sensor and impulse function of detected sample of coating structure with one layer of link coat + two layers of anticorrosive paint/steel substrate

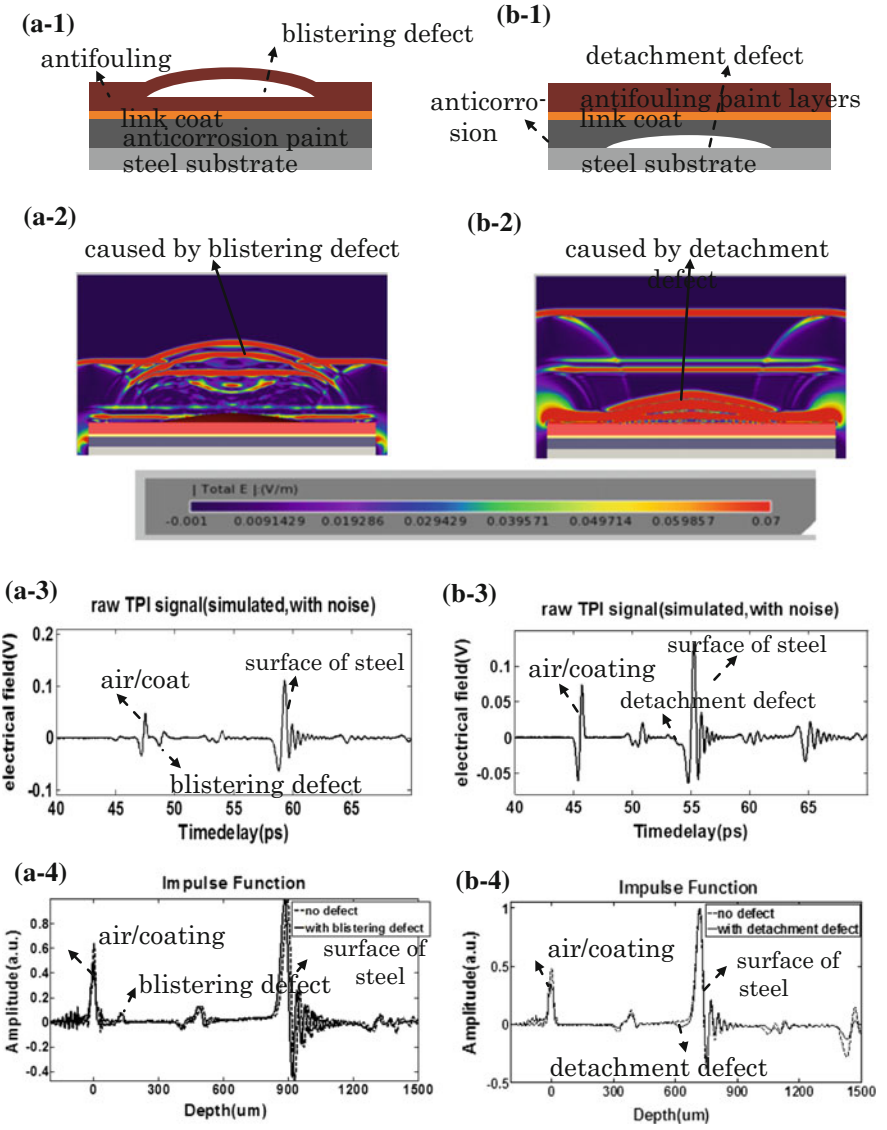


Fig. 5 FDTD calculated results of defective coating models. (**a-1**)～(**a-4**) The simulated coating sample, simulated result collected by the plannar sensor, TPI signal collected by the far zone sensor and impulse function of detected sample with blistering defect inside the coating; (**b-1**)～(**b-4**) The simulated coating sample, simulated result collected by the plannar sensor, TPI signal collected by the far zone sensor and impulse function of detected sample with detachment defect beneath the coating

smaller. Besides, taking into account the spray order in paint process, we also modeled and calculated the protective coating with other structures detected by terahertz wave, and the results were shown in Fig. 4.

Figure 4(a-1)∼(a-4) were the simulated coating sample, simulated result collected by the plannar sensor, TPI signal collected by the far zone sensor and impulse function of detected sample of coating structure with two layers of anticorrosive paint + steel substrate, respectively; Fig. 4(b-1)∼(b-4) were the simulated coating sample, simulated result collected by the plannar sensor, TPI signal collected by the far zone sensor and impulse function of detected sample of coating structure with one layer of link coat + two layers of anticorrosive paint + steel substrate, respectively.

Besides, we modelled the occurrence of defects in the marine protective coating, especially the defect beneath the paint surface. Figure 5(a-1), (b-1) were the simulated models of defective coating with different size of curved shape locating at different position to mimic the experimental samples, where the former was a blistering defect (with 2.5 mm radius and 350 μm thickness was embedded inside the antifouling layers with depth 100 μm from coating surface) and the later was a hidden detachment defect (with 3.5 mm radius and 150 μm thickness between the steel and coating). The FDTD calculated results which were collected by plannar sensor were shown in Fig. 5(a-2), (b-2), respectively. Figure 5(a-3), (b-3) were the calculated TPI signals which were collected by the far zone sensor, and their corresponding impulse functions were shown in Fig. 5(a-4), (b-4). As can be seen, the reflections caused by the paint layer interfaces and the defects can be identified in FDTD simulated results. The curved shape of the defect can be also easily distinguished and its location within the coating can be determined. It was believed that the defect feature would be easily extracted by the combination with signal processing technology. Therefore, it was able to provide much more detailed information on the position of the defect beneath the coating, such as detachment of large areas where adhesion has failed, and blistering defect beneath the coating. Compared to the expensive TPI measurement, FDTD simulation could be cheaper and more flexible and is of significance in the quality control of painting process.

4 Conclusions

The combined use of FDTD simulation and THz imaging measurements was carried out on the marine protective coating in the present work. The simulated results agreed well with the experimentally measured ones. The simulated and experimental investigation will provide a powerful method for TPI quantitative and non-destructive evaluation of protective coatings.

Acknowledgements We gratefully acknowledge support from the National Natural Science Foundation of China (51675103), the Fujian Provincial Excellent Young Scientist Fund (2014J07007), the Specialised Research Fund for the Doctoral Program of Higher Education, the Ministry of Education, P. R. China (20133514110008), the Ministry of Health, P. R. China (WKJ-FJ-27), the Fujian Provincial Natural Science Foundation (2016J01250, 2015J01234) and the Fujian Provincial Quality and Technical Supervision Bureau Project (FJQI2014008, FJQI2013024).

References

1. R. Moshrefi, M. Ghassem Mahjani, A. Ehsani, M. Jafarian, A study of the galvanic corrosion of titanium/L316 stainless steel in artificial seawater using electrochemical noise (EN) measurements and electrochemical impedance spectroscopy (EIS). Anti-Corros. Method Mater. **58**(5), 250–257 (2011)
2. S. Mabbutt, D.J. Mills, C.P. Woodcock, Developments of the electrochemical noise method (ENM) for more practical assessment of anti-corrosion coatings. Prog. Org. Coat. **59**, 192–196 (2007)
3. X. Zhang, F. Augereau, D. Laux, E. Le Clezio, N.A. Ismaili, M. Kuntz, G. Despaux, Non-destructive testing of paint coatings on steel plates by ultrasonic reflectometry. J. Nondestr. Eval. **33**(4), 504–514 (2014)
4. Y.Z. He, G.Y. Tian, H. Zhang, M. Alamin, A. Simm, P. Jackson, Steel corrosion characterization using pulsed eddy current systems. Sens. J., IEEE **12**(6), 2113–2120 (2012)
5. Y.Z. He, G.Y. Tian, M.C. Pan, D.X. Chen, H. Zhang, An investigation into eddy current pulsed thermography for detection of corrosion blister. Corros. Sci. **78**, 1–6 (2014)
6. T. Yasui, T. Yasuda, K.I. Sawanaka, T. Araki, Terahertz paintmeter for noncontact monitoring of thickness and drying progress in paint film. Appl. Opt. **44**, 6849–6856 (2005)
7. D.J. Cook, S. Lee, S.J. Sharpe, M.G. Allen, Accuracy and linearity of time-domain THz paint thickness measurements. Proc. of SPIE **6893**, 68930H (2008)
8. T. Kurabayashi, S. Yodokawa, S. Kosaka, in *Terahertz Imaging Through Paint Layers*, 37th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz), (2012), pp. 1–2
9. E. Pickwell-MacPherson, S. Huang, Y. Sun et al., in *Terahertz Image Processing Methods for Biomedical Applications*, Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (2008) pp. 3751–3754
10. Shen Yaochun, Philip F. Taday, Development and application of terahertz pulsed imaging for nondestructive inspection of pharmaceutical tablet. IEEE J. Sel. Top. Quantum Electron. **14** (2), 407–415 (2008)
11. P.A. Tirkas, C.A. Balanis, M.P. Purchine, G.C. Barber, Finite-difference time-domain method for electromagnetic radiation, interference, and interaction with complex structures. IEEE Trans. Electromag. Compat. **35**, 192–203 (1993)
12. C. Larsen, D.G. Cooke, P.U. Jepsen, Finite-difference time domain analysis of time-resolved terahertz spectroscopy experiments. J. Opt. Soc. Am. B **28**, 1308–1316 (2011)
13. W.L. Tu, S.C. Zhong, Y.C. Shen, Q. Zhou, L.G. Yao, FDTD-based quantitative analysis of terahertz wave detection for multilayered structures. J. Opt. Soc. Am. A **31**, 2285–2293 (2014)
14. W.L. Tu, S.C. Zhong, Y.C. Shen, A. Incecik, Nondestructive testing of marine protective coatings using terahertz waves with stationary wavelet transform. Ocean Eng. **111**, 582–592 (2016)
15. C.G. Soares, Y. Garbatov, A. Zayed, G. Wang, Influence of environmental factors on corrosion of ship structures in marine atmosphere. Corros. Sci. **51**, 2014–2026 (2009)
16. I.S. Russe, D. Brock, K. Knop, P. Kleinebudde, J.A. Zeitler, Validation of terahertz coating thickness measurements using X-ray microtomography. Mol. Pharm. **9**(12), 3551–3559 (2012)
17. Y. Chen, S. Huang, E. Pickwell-MacPherson, Frequency-wavelet domain deconvolution for terahertz reflection imaging and spectroscopy. Opt. Express **18**(2), 1177–1190 (2010)

The Development of in Situ Measurement Technique for Ship Plate Manufacturing

Yi-Fan Liao, Chi-Wen Huang, Chen-Min Fan, Shu-Sheng Lee, Pai-Chen Guan, Yu-Ling Chen and Yen-Hui Shi

Abstract Processing steel plate is the major step during the ship construction processes. In order to maintain more accurate control of the ship plate bending and reduce technical training time and cost, we developed a computer program for calculating the difference between the geometry of the manufactured curve plate with the design blue print. In the developed program, the coordinate rotation matrix and Euler angles (Deakin in Surv. Land Inf. Sys. 58:223–234, [1]; Lehmann in J. Geodesy 88:1117–1130, [2]; Shimizu in Evaluation of Three Dimensional Coordinate Measuring Methods for Production of Ship Hull Blocks. Kitakyushu, Japan, pp. 348–351 [3]) are applied to unify the coordinate system between the measured 3D data with the ship coordinate system. The program is written under the MATLAB coding environment. In our research, the most time consuming task is to determine which sets of rotation axes that can be used to transform the measuring coordinate to the design coordinate with the minimum error. Therefore, the least squares method is used to find the optimum Euler angle based on the minimization of error between the measured and design data. We have built an iterative scheme to search the most suitable set of rotation axes and the Euler angles, which spends about 30 s or less, and it has been tested for the in situ application. The program has the potential to speed up the plate processing time and aid the automated manufacturing capability of the ship yard.

Y.-F. Liao · C.-W. Huang · C.-M. Fan · S.-S. Lee (✉) · P.-C. Guan
Department of Systems Engineering and Naval Architecture, National Taiwan Ocean University, Keelung, Taiwan
e-mail: sslee@ntou.edu.tw

Y.-L. Chen · Y.-H. Shi
CSBC CORP, Kaohsiung, Taiwan

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_16

1 Introduction

Precision control of ship construction is very important for shipyards to reduce the cost and to produce high quality ships [4, 5]. Due to the reducing labor force, the automated production for the shipyard is becoming an urgent task. Introducing good measuring tool and procedures to most of the ship construction phases and to monitor the quality of the manufactured ship plates or the ship hull blocks is the first step for the automation. Therefore, the ability of measuring three-dimensional plate and blocks is the essential task in the past decade [3] in order to boost the manufacturing capability. The most popular device used in the ship yard is the Electronic Distance Measuring Device (EDM) and the measuring results are promising. However, there are few studies about the application of EDM for the curved steel plate manufacturing. In this work, we focused our effort on using EDM to obtain the steel plate 3D coordinate and to calculate the difference between the geometry of the manufactured curve plate with the design blue print. Our goal is to develop a complete algorithm and associated computer software for the searching of the coordinate transformation angle between the measuring and the design data and finish the comparison in 40 s. Hence, the newly employed shipbuilding technician can determine the state of the manufactured curve plate and decide the further process by the aid of this program without long training period. The labor requirement and cost can be reduced and can be used for the production automation of the shipyard.

2 Algorithm for Coordinate Transformation and Error Calculation

The Euler angle is commonly used to describe the orientation of rigid body in three dimensional space. Since the coordinate orientations of the measured plate geometry and the design blue print are usually not consistent, the coordinate transformation algorithm should first be developed to achieve the comparison steps and verify the accuracy of the manufacturing processes. To verify the rotating transformation matrix, we first choose certain reference points on the plate, such as corner points, center line, etc., to define lines of comparison between measured data and the accurate geometry on the blue print. Then we build the coordinate system locally on the plate and perform rigid body rotation to be consistent with the blue print data. The following section is the detail description of the algorithm.

2.1 Coordinate Transformation

In this research, we define the three basic rotating transformation matrices along the x-, y-, and z-axis in Cartesian coordinate system. The associated rotation matrices are

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{bmatrix},$$

$$\mathbf{R}_y = \begin{bmatrix} \cos\theta_y & 0 & -\sin\theta_y \\ 0 & 1 & 0 \\ \sin\theta_y & 0 & \cos\theta_y \end{bmatrix},$$

and

$$\mathbf{R}_z = \begin{bmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The angles θ_x, θ_y, and θ_z are the Euler angles referring to the x-, y-, and z-axis respectively. Any two consecutive rotations can be established if the rotation axes are not repeated. There are 12 different sequence to form the rotation matrices: xyx, xzx, yxy, yzy, zxz, zyz (the classic Euler angle), xyz, xzy, yxz, yzx, zxy, and zyx for three elemental rotations (Fig. 1).

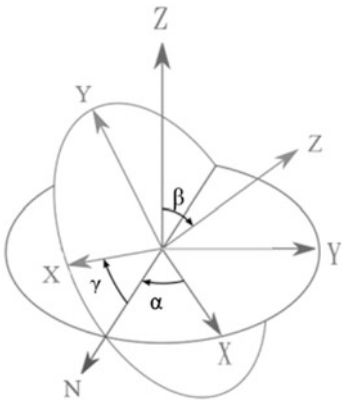


Fig. 1 The rotational angle along the axis in the coordinate system

2.2 Error Correlation Formula

To search the best orientation which can make the measured data coordinate match the design, we tested the rotation algorithm with all 12 different rotation sequence mentioned above. We choose several reference points and calculate their rotated 3D coordinate by the 12 rotation sequences. Then, we subtract the rotated data from the design blue print and calculate the sum of squared difference. Finally, find the smallest total error, it's associated rotation sequence and Euler angles θ_x, θ_y, and θ_z. Then take this configuration and build the major transformation matrix and rotation angles. If the summed squared difference value is larger than the criteria, we alter the Euler angles every 0.5° within ±5° and repeat the rotation calculation until the summed difference satisfies the criteria. The iterative scheme is shown in Fig. 2.

To minimize the error, we apply the error correction formula, and its basic concept is showed as the following.

$$\begin{bmatrix}\text{Average of the referring}\\ \text{design points 3D value}\end{bmatrix} = [\text{Rotation Matrix}]\begin{bmatrix}\text{Average of the referring}\\ \text{measured points 3D value}\end{bmatrix} + [\text{correction value}]$$

$$\begin{bmatrix}\text{Corrected rotated referring}\\ \text{measured points 3D value}\end{bmatrix} = \begin{bmatrix}\text{rotated referring measured}\\ \text{points 3D value}\end{bmatrix} + [\text{correction value}]$$

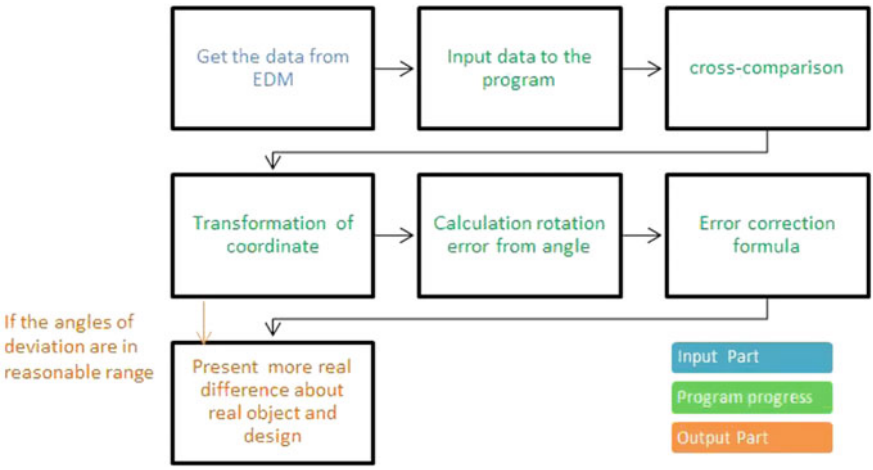


Fig. 2 The iterative scheme to search the most suitable set of rotation axes and the Euler angles

3 Data Processing Results

The program was then tested by performing two case studies. The first case is the in-situ measured curved steel plate. The second case is the assumed spherical surface.

3.1 In-Situ Curved Steel Plate 3D Data

This case is a real ship hull plate provided by CSBC CORP. The in-situ measured (of the curved steel plate) data is compared with the design data from design blue print. The input data are listed in Table 1. The result of adjusted reference points based on corrected coordinate are shown in Table 2. All the points in Tables 1–2 are plotted in the Fig. 3.

3.2 Partial Spherical Surface

In this example, we use a partial sphere plate as an example to test the developed program. Comparing to the real ship hull plate, more reference points and data are

Table 1 Reference points data from design blue print and in situ measurement

Reference points from design blue print			Reference points from in-situ measurement		
x (m)	y (m)	z (m)	x (m)	y (m)	z (m)
−1.6	0	0.022935	−1.600	0.015	0.028
−0.8	0	0.002970	−0.800	0.013	0.011
0.0	0	0.000000	0.000	0.000	0.000
0.8	0	0.023645	0.813	−0.005	0.008
1.6	0	0.048765	1.619	−0.008	0.048

Table 2 The rotated reference points on the steel plate and the Euler angles

x (m)	y (m)	z (m)
−1.6002	−0.0011	0.0207
−0.8001	0.0052	0.0078
0.0000	0.0000	0.0000
0.8130	0.0012	0.0119
1.6188	0.0016	0.0557
The square root of the average of the summed difference value: 0.0124		
θ_x	θ_y	θ_z
3°	−5°	−3.5°

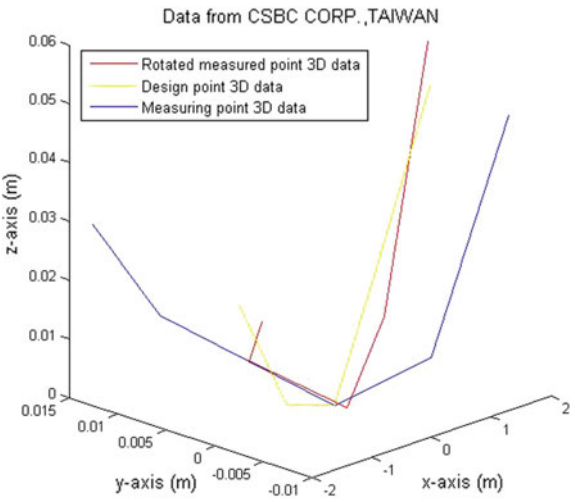


Fig. 3 The 3D plot of in-situ measure steel plate 3D data, and the design data and the rotated data

obtained to make the result more accurate. To simulate the real measured data, we introduce random error to the nodal coordinate. We introduce 10 × 10 points, including 10 points as the reference points listed in Table 3. The reference points are used to find the best transformation matrix and the Euler angles. The sequence rotation matrix for this case is yxz and the rotation angles were −0.5°, −4.5°, and −2.0°. The norm of error is 0.070943. The coordinate shifting for x, y, and z are −0.10258, −0.16060 and −0.11782 m. The calculation takes about 30 s. Table 4 shows the rotated and corrected reference points data. The simulated design, measuring, and rotated and corrected spherical surface plate data is shown in Fig. 4.

Table 3 The spherical plate data with and without errors

Spherical plate data without errors			Spherical plate data with errors		
x (m)	y (m)	z (m)	x (m)	y (m)	z (m)
37	53	51.24967	43.01529	50.63578	49.41597
37	54	50.96808	43.03472	51.47878	49.08641
37	55	50.67497	43.01003	52.54243	48.83126
37	56	50.37023	43.26704	53.61063	48.47020
37	57	50.05375	43.10213	54.53834	48.27734
37	58	49.72540	43.24319	55.41415	47.87985
37	59	49.38506	43.23313	56.40495	47.42539
37	60	49.03260	43.54460	57.45074	47.11221
37	61	48.66788	43.39556	58.46712	46.82141
37	62	48.29074	43.62301	59.49968	46.42510

Table 4 The rotated and corrected reference points data

x (m)	y (m)	z (m)
37.06187	53.11054	51.23608
37.02655	53.95621	50.91477
36.92737	55.01754	50.66817
37.11219	56.10662	50.32579
36.88182	57.02076	50.13526
36.96747	57.90941	49.75067
36.89552	58.90154	49.30478
37.13468	59.97198	49.01199
36.91657	60.97688	48.72493
37.07595	62.02853	48.34595

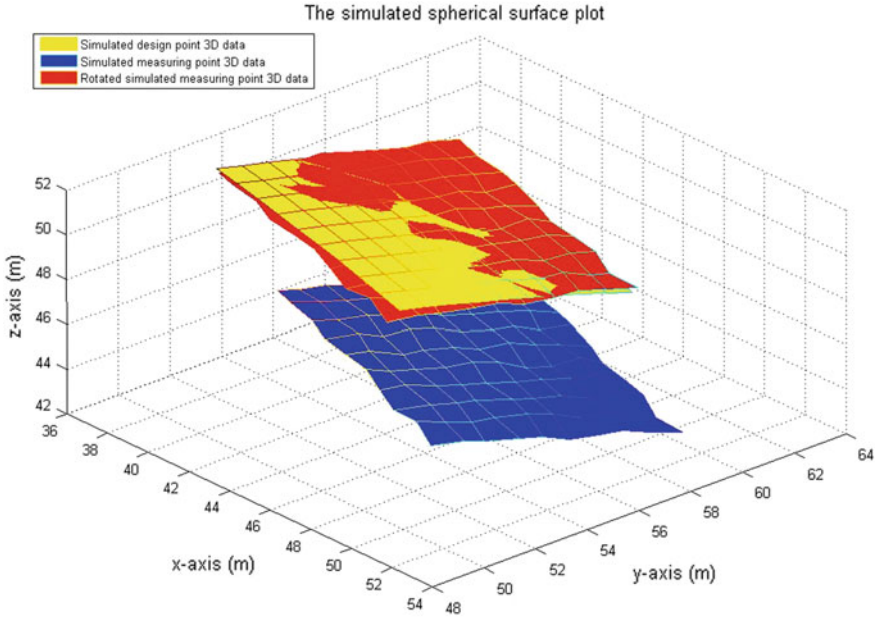


Fig. 4 The 3D display for the simulated design, measuring, and rotated and corrected spherical surface plate data

4 Conclusion

We have developed an iterative scheme to search the most suitable set of rotation axes and the Euler angles for in situ measured (of the curved steel plate) data rotating to the design blue print coordinate orientation. Then, the square root of the average of the summed difference value between measured data and the design blue print are used as the iteration criteria. The general calculation of our developed

program cost less than 30 s to complete the whole process. From the calculation result of the simulation and the real ship hill plate data, the developed process and program can successfully deal with the comparison between measured and design ship plate data. Therefore, the program has the potential to be used in situ ship plate bending check and helpful to the automation ship construction in the shipyard.

Acknowledgements This research work was financial support by the CSBC CORP., TAIWAN and the National Taiwan Ocean University.

References

1. R.E. Deakin, 3D coordinate transformations. Surv. Land Inf. Sys. **58**(4), 223–234 (1999)
2. R. Lehmann, Transformation model selection by multiple hypotheses testing. J. Geodesy **88** (12), 1117–1130 (2014)
3. H. Shimizu, in *Evaluation of Three Dimensional Coordinate Measuring Methods for Production of Ship Hull Blocks*, Proceedings of The Twelfth International Offshore and Polar Engineering Conference, Kitakyushu, Japan, (2002) pp. 348–351
4. F. Xiao-bin, Analysis and countermeasures on precision management of hull construction of small and medium sized shipyard. J. NanTong Vocat. Tech. Shipping College, **4**: 57–60 (2009), in Chinese
5. L.I.U. Yu-jun, L.I. Yan-jun, Calculation method of two dimensional size chain of ship production accuracy control. Shipbuild. China **45**(2), 81–87 (2004)

FEM Study of Graphene Based Tunable Terahertz Plasmonics Gaseous Sensor

Yi Huang, Shuncong Zhong and Haizi Yao

Abstract A suspended doped graphene monolayer based terahertz (THz) plasmonics gaseous sensor using Otto prism-coupling attenuated total reflection configuration was presented. The finite element method (FEM) software COMSOL MULTIPHYSICS is used to combine Maxwell's equation and Druse-Lorentz's dispersion model into a 2D simulations to investigate the performance of the sensor in terms of detection accuracy with different refractive index gaseous sample. The sensitivity was also acquired by using the dispersion relation analysis. Preliminary results showed good agreement between simulation and theoretical analysis, as well as supporting previously literature reported calculation result. This method also can be used to computationally explore and design graphene based plasmonics sensor with more complex geometries compared with current sensor.

Keywords Sensing · Graphene · Terahertz · Plasmonics

1 Introduction

Terahertz (THz) plasmonics sensor is an important aspect of surface plasmon resonance (SPR) applications [1, 2]. The collective oscillations of free charge carriers at the interface of metal-dielectric with permittivities of opposite sign gives

Y. Huang · S. Zhong · H. Yao
Laboratory of Optics, Terahertz and Non-Destructive Testing,
School of Mechanical Engineering and Automation, Fuzhou University,
Fuzhou 350108, People's Republic of China

S. Zhong (✉)
Fujian Key Laboratory of Medical Instrument and Pharmaceurical Technology,
No.2 Campus Road, Fuzhou 350108, People's Republic of China
e-mail: zhongshuncong@hotmail.com

S. Zhong
Shanghai University, Shanghai, P. R. China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_17

rise to surface plasmon (SPP) [3]. Such Surface plasmon polariton wave (SPPs) that propagating along interface [4]. They can enhance the intensity of electromagnetic field vastly in a confined region and confine the electromagnetic wave within subwavelength [5]. Furthermore, they are known to be dramatically sensitive to any changes of the dielectric environment close to the interface [6]. Because of different refractive index, different analyte that adhered on the metal surface make the dielectric environment change [7, 8]. Graphene is a 2D structure consisting of sp^2 bonded carbon atoms in a hexagonal lattice [9]. Recent studies have showed that the permittivities of graphene are similarly to metal at THz and midinfrared regime. It also can supported highly confined SPP with superiority of being highly tunable via electrical and chemical doping [10] and their 2D structure makes the electron transport through graphene highly sensitive to the adsorption of gas molecules [11]. Thus, it would be intriguing to incorporate graphene into the design of SPR sensor to enable dynamically tunable SPR for many different types of sensing applications. As a result, an effective and convenient method to investigate and design graphene based sensor, is of great practical value. Recently, an ultrasensitive THz-plasmonics gaseous sensor using doped graphene has been proposed [12] and Transfer Matrix Method (TMM) has been used to explore the performance of sensor. But TMM needs more complex computational process. At the same time, it can't visually present the results of simulation. In this paper, a suspended doped graphene monolayer based terahertz (THz) plasmonics gaseous sensor using Otto prism-coupling attenuated total reflection configuration is presented. The finite element method software COMSOL MULTIPHYSICS is utilized to combine Maxwell's equation and Druse-Lorentz's dispersion model into a 2D simulations to investigate the performance of the sensor in terms of detection accuracy with different refractive index gaseous sample.

2 Design of Sensor System

In this paper, we designed a suspended doped MLG based THz-plasmonics sensor using Otto prism-coupling attenuated total reflection configuration. The configuration of Otto-based THz- plasmonics sensor using a suspended doped MLG illustrates in Fig. 1. It consisted of a prism with high refractive index prism of $n_p = 4$ [13], suspended graphene between two media of refractive index n_s and n_a and graphene doping device which was made up of gold electrodes (Au), partly etched SiO_2 and degenerately doped silicon gate (Si). Here, we have considered $n_a = 1$ which gives us the suspended structure of graphene [14]. The suspended graphene was beneath the prism bottom and filled with gaseous analyte with a thickness of g between the prism base and graphene. The effective thickness of graphene $t_g = 0.34$ nm [15]. When the sensor system works, graphene is doped by adding the voltage to the poles of gold electrodes and doped silicon gate. After that, a paralleled THz beam was injected from the left facet of prism and then the evanescent wave created by total internal reflection at the bottom of prism with

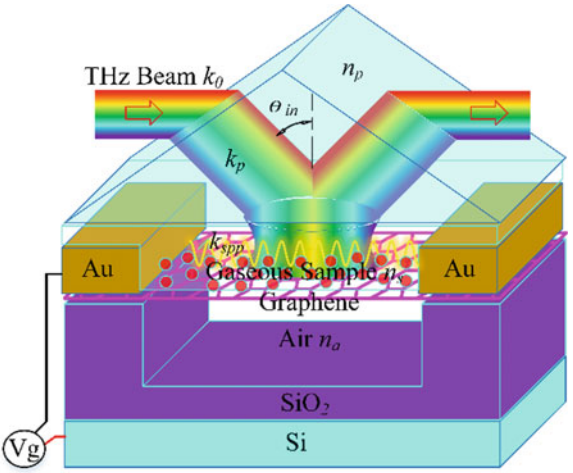


Fig. 1 Otto-based SPR sensing using doped MLG including high refractive index n_p, suspended doped MLG between two media of refractive index n_s and n_a and graphene doping device which was made up of gold electrodes (Au), partly etched SiO_2 and degenerately doped silicon gate (Si). k_0 is wavevector of vacuum and k_p is wavevector of prism with the internal incident angle θ_{in}. k_{spp} is the parallel components of SPP wave

the internal incident angle θ_{in}. After interacting with the gaseous sample, the modified THz radiation will reflect off prism base and exit to the other facet of prism where the modulated radiation would be detected. The incident THz beam has a wavevector k_p (where $k_p = n_p k_0$, k_0 is wavevector of vacuum;$k_p^2 = k_\parallel^2 + k_\perp^2$, $k_\parallel$ and $k_\perp$ are the parallel and perpendicular components of k_p). When the condition of momentum matching is fulfilled, i.e., $k_{spp} = k_\parallel$ (where k_{spp} is the parallel components of SPP wave), the corresponding frequency THz will couple into attenuated SPP wave propagating on the dielectric-graphene interface and different analyte corresponds to the frequency of coupling. As a result, the detected reflectivity spectrum will exhibit a sharp dip at coupled frequency. Thus, the dielectric parameters of medium between prism base and graphene can be detected by measuring the reflection spectrum.

3 Physical Mechanism

The current achievable levels of doping for graphene is, i.e., $E_F = 0.4 \sim 1.6$ eV [16], where E_F is the Fermi level energy of graphene. There are many methods for graphene doping, for example, chlorine plasma reaction [17], electro thermal reactions [18], chemical vapor deposition [19] and so on. It is noted here that, we have considered graphene conductivity derived within random phase approximation at zero temperature as [20]

$$\sigma(\omega) = \frac{i}{\pi}\frac{e^2}{\hbar}\left[\frac{\frac{E_F}{\hbar(\omega+i\tau^{-1})} - \frac{1}{4}\ln(2E_F + \hbar(\omega+i\tau^{-1}))}{2E_F - \hbar(\omega+i\tau^{-1})}\right] \quad (1)$$

The first component of the equation is the intraband contribution and the second component is the intraband contribution. At low THz frequencies, with high E_F of MLG ($\hbar\omega \ll E_F$), the graphene conductivity is dominated by the intraband contribution and can be approximated by a Drude model by [21]

$$\sigma(\omega) \approx i\frac{e^2 E_F}{\pi\hbar^2(\omega + i\tau^{-1})} \tag{2}$$

The permittivity of graphene can be given as

$$\varepsilon_g = 1 + \frac{i\sigma}{\omega\varepsilon_0 t_g} \tag{3}$$

where e is the charge of an electron, $\hbar$ is the reduced Planck constant, E_F is the Fermi level energy of graphene is determined by the carrier concentration $n_0 = (E_F/\hbar v_F)^2/\pi$ (v_F is Fermi velocity), τ is the phenomenological relaxation time, and ω is the angular frequency. ε_0 is the permittivity of vacuum, the variation of real and imaginary part of ε_g at THz frequency for three different value E_F = 0.4, 0.8, 1.2 eV were plotted in Fig. 2. It is shown that $|\Re\varepsilon_g| \gg |\Im\varepsilon_g|$ in the local limit ($\omega \gg \tau^{-1}$) for highly doped graphene. Thus, highly doped graphene would support highly confined SPP modes at THz frequency. What's more, the imaginary part of ε_g decreases (inverse of the real part) with frequency increases for the same E_F. By

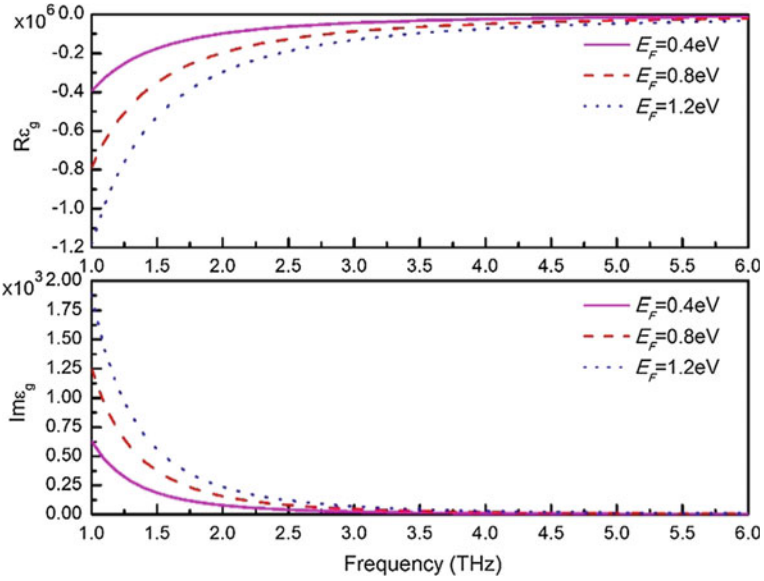


Fig. 2 The variation curve of real and imaginary of ε_g at THz frequency for different E_F

matching the boundary conditions for the n_s-MLG-n_a SPR system (Fig. 2), the dispersion relation of SPP can be derived as [22]

$$\frac{n_s^2}{k_{zs}} + \frac{n_a^2}{k_{za}} + \frac{i\sigma}{\omega\varepsilon_0} = 0 \tag{4}$$

here $k_{zi} = \sqrt{k_{spp}^2 - n_i^2 k_0^2}$ $(i = s, a)$ is the perpendicular components of wavenumber of SP.

To demonstrate the dispersion relations of SPP wave supported on the surface of MLG for different E_F, we plotted the dispersion curve of SPP wave by using Eq. (4) for three different value E_F = 0.4, 0.8 and 1.2 eV in Fig. 3. It is to be noted that two dielectrics above and below graphene was assumed to be vacuum with the refractive index of $n_s = n_a = 1$ and their height were $t_s = t_a = 40\,\mu\text{m}$ in the simulation. From the Fig. 3 we can see that different E_F of graphene have great influence on the slope of the dispersion curve, the greater value of E_F, the greater slope of the dispersion curve. The adjustability of dispersion relation by E_F of graphene illustrates that such SPP wave has a significant design flexibility. The effective SPR frequency can be designed across frequency of interest.

From the condition of momentum matching, $k_{spp} = k_{\|} = n_p \sin(\theta_{in}) k_0$, we can find that with the increase of θ_{in}, the Otto-based SPR sensing can support a larger $n_p \sin(\theta_{in})$, i.e., a larger internal incident angle has a broader detection range. What' more, with the change of θ_{in}, the $k_{\|}$ also will change accordingly, i.e., the effective

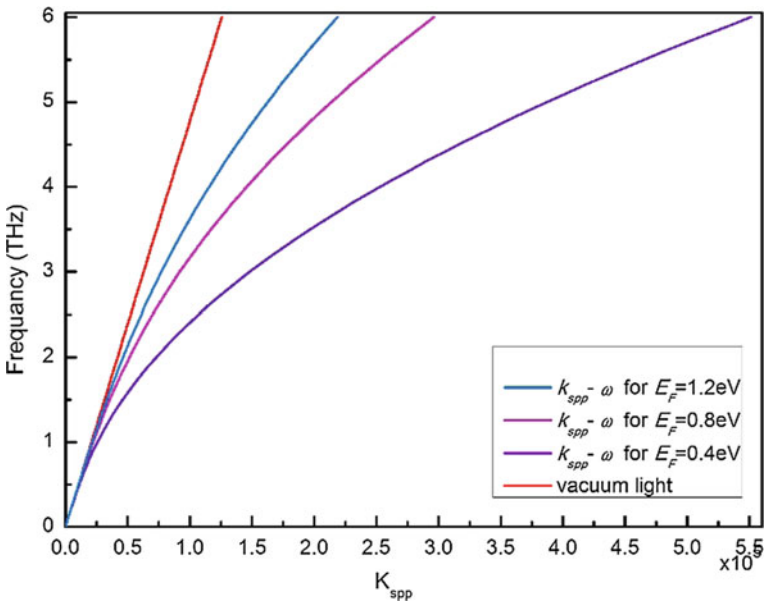


Fig. 3 Dispersion relation of different E_F of MLG

SPR frequency also can be adjust by the internal incident angle θ_{in}. The relationship between resonance frequencies and refractive index of gaseous sample n_s can be elicited by instead of the part of k_{spp} in Eq. (4) by $k_{\|}$. The sensitivity of SPR sensor is defined as the ratio of the change in resonance frequency (δf_{SPR}) with the change in refractive (δn_s):$S_n = \delta f_{SPR}/\delta n_s$ Thus, the theoretic sensitivities of SPR sensing for different value of E_F at different θ_{in} can be approximated by the slope of linear fit curve according. In our calculation, the theoretic sensitivities of SPR sensing are 2.22 THz/RIU, 4.43 THz/RIU and 6.65 THz/RIU for three different value of E_F = 0.4, 0.8 and 1.2 eV at θ_{in} = 30°, respectively. For the case of θ_{in} = 45°, they are 1.70 THz/RIU, 3.41 THz/RIU and 5.11 THz/RIU, respectively which are higher than the results reported previously [23, 24].

4 Simulation and Results

In order to further investigate the performance of MLG-based SPR sensing, we conducted 2D simulations to obtain the reflectivity values for various SPR resonance frequencies in all circumstances. It is to be noted that the computational domain to be considered in this work is the part of suspended graphene sandwich structure, i.e., gaseous sample-graphene-air as shown in the Fig. 1 for higher efficiency of calculation. The simulation model is the unit cell as shown in the Fig. 4. The two sides of model are Periodic boundary (BD). The periodic port was applied to inject TM-polarized THz wave and the bottom of model is perfectly matched layer (PML). Furthermore, the Maxwell's equation and Druse-Lorentz's dispersion model was applied to simulations. Thanks to the mesh size have great influence on the accuracy of FEM computation. Thus, in order to guarantee the precision of the simulation and satisfies the criterion of finite element method in COMSOL, we meshed the graphene into three layers by free triangular meshing with the minimum

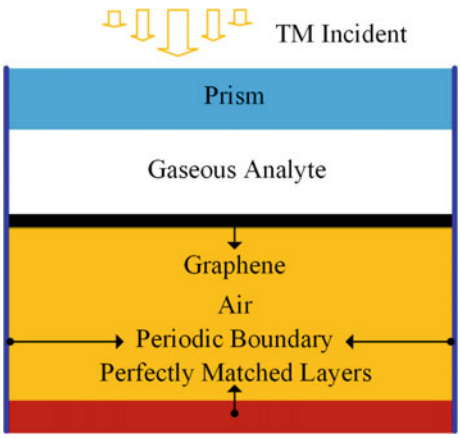


Fig. 4 Unit cell simulation model

element size of 0.1 nm; also the maximum element size outside of graphene layer was set to 1000 nm, which is within one fifth of wavelength of incident light. From the electric field distributions of simulation as shown in the Fig. 5, we can see that the intensity of electromagnetic field enhance vastly above and below MLG. It further illustrates that highly doped MLG could support highly confined SPP modes

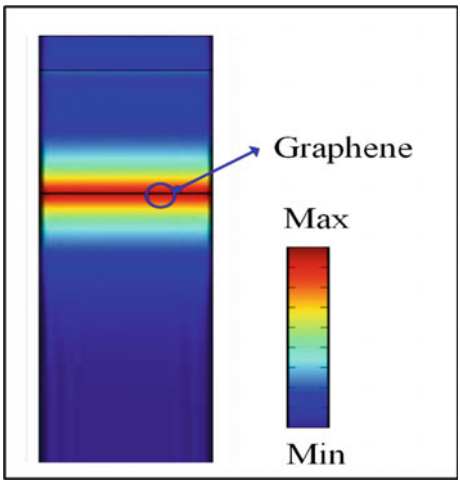


Fig. 5 The electric field distributions of simulation

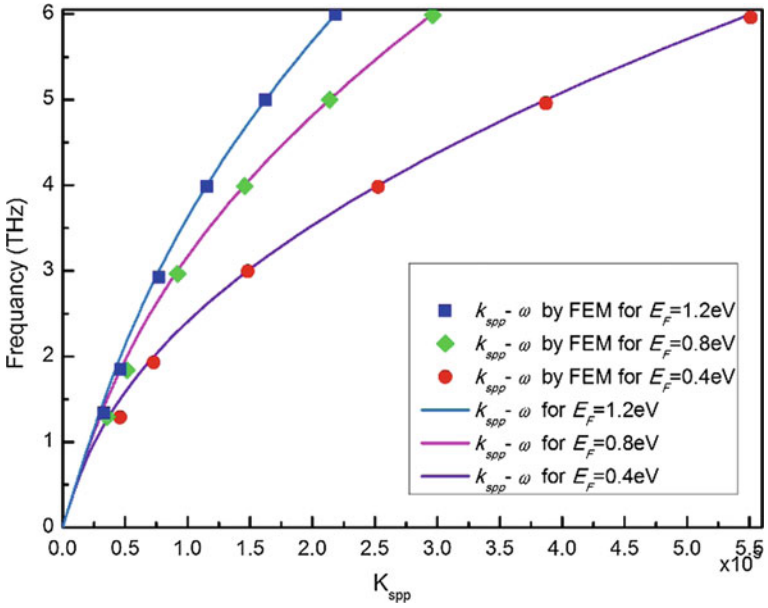


Fig. 6 The scatter data of the numerical results using COMSOL MULTIPHYSICS

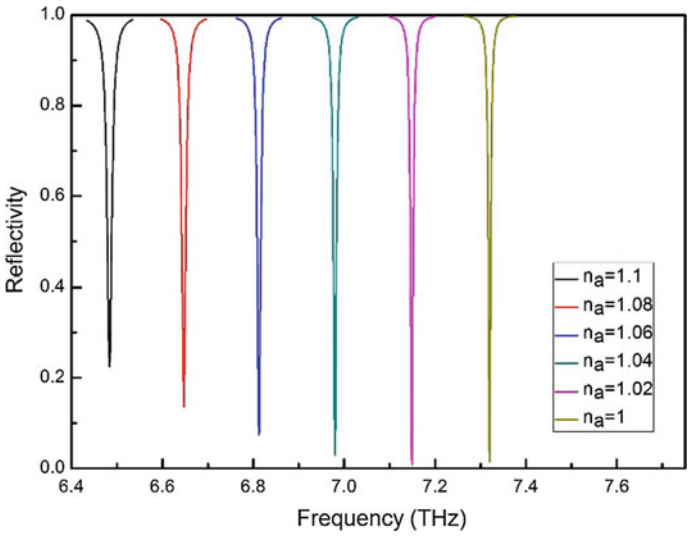


Fig. 7 The reflectivity spectra of various refractive index ns from 1.0 to 1.1

at THz frequency. As shown in the Fig. 6. The scatter data were obtained by finite element method by using COMSOL MULTIPHYSICS software. It is demonstrated that the numerical results are highly agreed with the dispersion curve of theoretical analysis. We also plotted the reflectivity spectra of various refractive index n_s from 1.0 to 1.1 for E_F = 1.2 eV and gap = 28 μm at θ_{in} = 30° as shown in the Fig. 7. Detection accuracy (DA) is defined as the reciprocal of full width at half maximum (FWHM), i.e. DA = 1/FWHM. It is found that the values of FWHM are so small that can greatly enhance detection accuracy. The DA values we calculated for ns = 1.00, 1.02, 1.04, 1.06, 1.08 and 1.10 is to be 178, 147, 133, 115, 98 and 83 THz^{-1}, respectively.

5 Conclusions

In summary, a suspended doped graphene monolayer based terahertz (THz) plasmonics gaseous sensor using Otto prism-coupling attenuated total reflection configuration is presented. We have demonstrated an analog Simulation of this tunable gaseous sensor by using the finite element method. Using COMSOL MULTIPHYSICS to combine Maxwell's equation and Druse-Lorentz's dispersion model into a 2D simulations to investigate the performance of the sensor. The numerical results are highly agreed with the dispersion curve of theoretical analysis.

This work provides reference value to a THz-Plasmonics gaseous sensor of refractive index develop and design.

Acknowledgments We gratefully acknowledge support from the Fujian Provincial Excellent Young Scientist Fund (2014J07007), the National Natural Science Foundation of China (51675103), the Training Program of Fujian Excellent Talents in Universities, the Specialised Research Fund for the Doctoral Program of Higher Education, the Ministry of Education, P.R. China (20133514110008) and the Ministry of Health, P.R. China (WKJ-FJ-27).

References

1. P. Pattnaik, Appl. Biochem. Biotechnol. **126**, 79 (2005)
2. J. Homola, S.S. Yee, G. Gauglitz, Sens. Actuators, B **54**, 3 (1999)
3. S.A. Maier, *Plasmonics: fundamentals and applications* (Springer Science & Business Media, 2007)
4. H. Raether, *Surface plasmons on smooth surfaces* (Springer, Heidelberg, 1988)
5. J. Pitarke, V. Silkin, E. Chulkov, P. Echenique, Rep. Prog. Phys. **70**, 1 (2006)
6. W.L. Barnes, A. Dereux, T.W. Ebbesen, Nature **424**, 824 (2003)
7. J.M. McDonnell, Curr. Opin. Chem. Biol. **5**, 572 (2001)
8. J. Homola, M. Piliarik, *Surface Plasmon Resonance based Sensors* (Springer, Heidelberg, 2006), p. 45
9. K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, Y. Zhang, S.V. Dubonos, I.V. Grigorieva, A.A. Firsov, Science **306**, 666 (2004)
10. C.H. Gan, H.S. Chu, E.P. Li, Phys. Rev. B **85**, 125431 (2012)
11. G. Lu, L.E. Ocola, J. Chen, Nanotechnology **20**, 445502 (2009)
12. A. Purkayastha, T. Srivastava, R. Jha, Sens. Actuators, B **227**, 291 (2016)
13. E.V. Loewenstein, D.R. Smith, R.L. Morgan, Appl. Opt. **12**, 398 (1973)
14. C.H. Gan, Appl. Phys. Lett. **101**, 111609 (2012)
15. D.K. Efetov, P. Kim, Phys. Rev. Lett. **105**, 256805 (2010)
16. C.-F. Chen, C.-H. Park, B.W. Boudouris, J. Horng, B. Geng, C. Girit, A. Zettl, M.F. Crommie, R.A. Segalman, S.G. Louie, Nature **471**, 617 (2011)
17. J. Wu, L. Xie, Y. Li, H. Wang, Y. Ouyang, J. Guo, H. Dai, J. Amer. Chem. Soc. **133**, 19668 (2011)
18. X. Wang, X. Li, L. Zhang, Y. Yoon, P.K. Weber, H. Wang, J. Guo, H. Dai, Science **324**, 768 (2009)
19. H. Liu, Y. Liu, D. Zhu, J. Mater. Chem. **21**, 3335 (2011)
20. B. Wunsch, T. Stauber, F. Sols, F. Guinea, New J. Phys. **8**, 318 (2006)
21. F.H. Koppens, D.E. Chang, F.J. Garcia de Abajo, Nano Lett. **11**, 3370 (2011)
22. G.W. Hanson, J. Appl. Phys. **103**, 064302 (2008)
23. B. Reinhard, K.M. Schmitt, V. Wollrab, J. Neu, R. Beigang, M. Rahm, Appl. Phys. Lett. **100**, 221101 (2012)
24. B. Ng, S. Hanham, V. Giannini, Z. Chen, M. Tang, Y. Liew, N. Klein, M. Hong, S. Maier, Opt. Express **19**, 14653 (2011)

Part III
Intelligent Systems

Advanced Manufacturing Systems—An Introduction to Holonic Manufacturing System

Nobuhiro Sugimura

Abstract Much emphasis has been given to research and development of manufacturing systems to produce machine products, such as automotive and machine tools, with high accuracy and quality since the first industrial revolution in 18th Centuries. Industry 4.0 initiated by Germany also deals with the advanced manufacturing systems, which have the capability called "Plug and Produce", by applying such recent information and communication technologies as CPS (Cyber Physical Systems) and IoTs (Internet of Things) . The paper deals with the new manufacturing systems called HMS (Holonic Manufacturing Systems), which was proposed by an international collaborations for the advanced manufacturing systems under the recent ICT (Information and Communication Technologies) environment.

1 Introduction

New system architectures have been proposed aiming at realizing more flexible management and control structures for the manufacturing systems in the last two decades. Autonomous distributed control architectures named HMS (Holonic Manufacturing System) has been proposed and applied to the manufacturing system, to cope with the dynamic changes in the volume and the variety of the products, and also with the unscheduled disruptions such as malfunction of manufacturing resources and interruption by high priority jobs. The objective of the research works in the HMS project is to develop a manufacturing system which satisfies the following requirements.

(1) Agility: The system carries out the manufacturing processes for small lot productions.

N. Sugimura (✉)
Graduate School of Humanities and Sustainable System Science,
Osaka Prefecture University, Sakai, Japan
e-mail: sugimura@me.osakafu-u.ac.jp

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_18

(2) Flexibility: The system can easily adopt to the changes in the volume and the variety of the products.
(3) Reconfigurability: Configurations of all the system components are easily changed to cope with the new production environment.

2 Overview of HMS Project in Japan

The HMS project was established in 1995 as an international cooperative project under the umbrella of the IMS (Intelligent Manufacturing System) program proposed by MITI (Ministry of International Trade and Industry) in Japan, and continued until 2006. The international HMS consortium includes about 30 organizations from Australia, Canada, EU (Germany and Belgium), Japan and USA aimed at developing new advanced and intelligent manufacturing systems looking toward the 21st century. Figure 1 shows the Japanese partners of the HMS consortium.

The HMS is a kind of autonomous distributed manufacturing systems and composed of a set of autonomous and co-operative elements called "holons", such as HMUs (Holonic Machining Units), HOLOFIXs (Holonic Fixturing Units), and holonic assembly units. The HMS forms a hierarchical architecture called holarchy as shown in Fig. 2. In the holarchy, higher level holons, such as companies, are composed of a set of lower level holons, such as factories.

The basic concept of the HMS is schematically illustrated in Fig. 2, and the holons and the holarchy are defined as follows.

(1) Holon:
An autonomous and cooperative building block of a manufacturing system. The holon consists of an information processing part and often a physical processing part. A holon can be a part of another holon.
(2) Autonomy
The capability of an entity to create and to control the execution of its own plans and strategies.

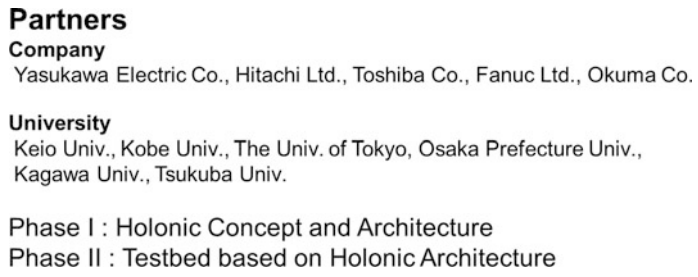


Fig. 1 Japanese partners for holonic manufacturing system (HMS) projects

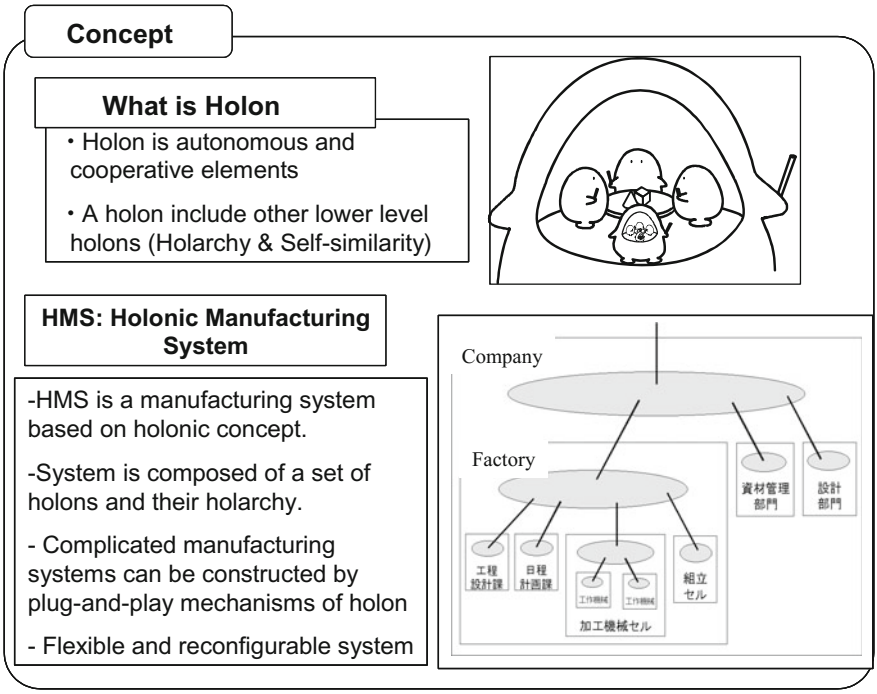


Fig. 2 Concept of HMS [1]

(3) Cooperation
A process whereby a set of entities develops mutually acceptable plans and executes these plans.

(4) Holarchy
A system of holons that can cooperate to achieve a goal and/or objective. The holarchy defines the basic rules for cooperation of the holons and thereby limits their autonomy.

Japanese HMS consortium has established six research teams named work-packeges shown in Fig. 3 to carry out the research and development separately and effectively. The R&D topics of the WPs are summarized in the followings.

(1) WP1 & 2: System Architecture and System Engineering & System Operation
The objectives of the WP1&2 are to investigate the functional requirements of the HMS, to establish the holon models, HMS models and information models, and to provide the standardized tools for implementation of the HMS. The following items are investigated.

Object-oriented modeling technique is applied to the development of the simulation model of the Japanese testbed of the induction motor manufacturing system. In particular, a model is proposed to represent the process planning system and the scheduling system in the testbed.

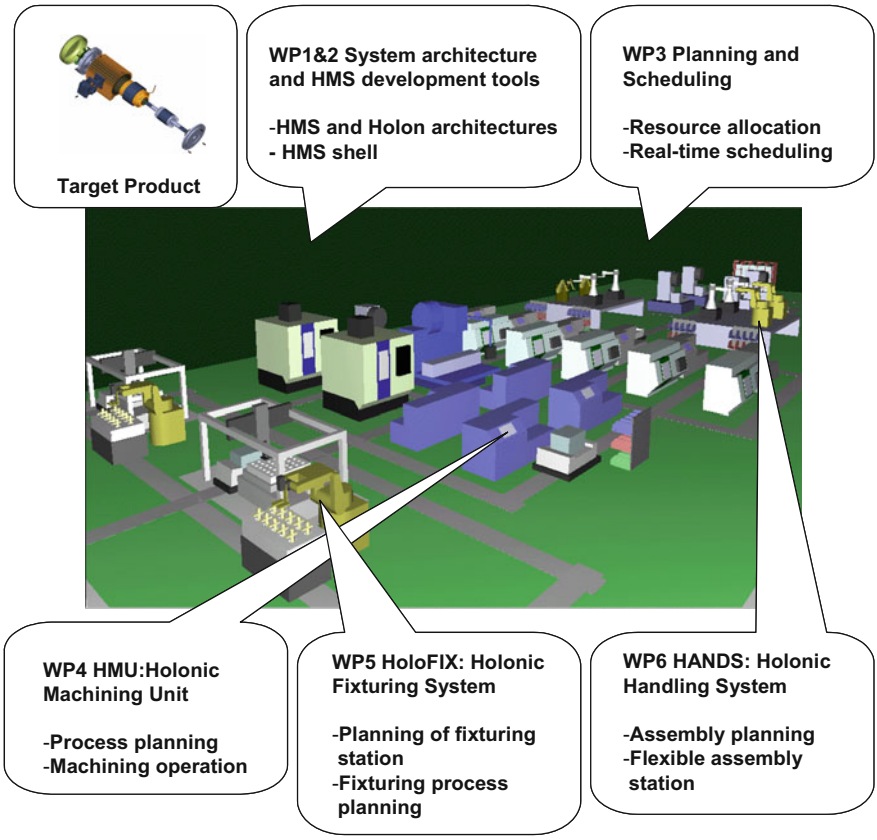


Fig. 3 Work packages in HMS projects [2]

A system development and execution environment for HMS named HMS shell is developed, based on the investigations to the functions and the information required for the HMS implementation. The environment includes the messaging mechanisms, the cooperation protocol libraries, the supporting tools for the holon and holarchy definition, the holon libraries and the simulation system.

(2) WP3: Holonic Resource Allocation

The objectives of WP3 are to establish common technologies that enable effective and efficient management and control system of HMS, based on the results proposed by WP1 & 2, and to develop the testbed systems in order to verify the feasibility and the capability of HMS. The HMS management and control technologies relating to cooperative planning and scheduling are discussed in the WP.

Contract-net protocol with CSP (Constraint Satisfaction Problem) is adopted as a basic method for the cooperation mechanisms to solve the problems.

The main research issues are the development of the efficient CSP solving algorithms. The object-oriented model is proposed to represent the holonic elements, such as orders, workpieces and equipment, which are the holonic elements needed in the process planning and production scheduling problems. The main research topics are the development the of simulation models in which the individual holons cooperate with each other by exchanging their autonomous decisions, and the development of the holonic real-time scheduling systems based on the models. [3]

(3) WP4: Holonic Machining Unit (HMU)

The objective of WP4 is to develop HMUs (Holonic Machining Units) which are basic system elements of a holonic manufacturing system. The participants of HMU individually develop the autonomous components for the machining units, such as process planning systems, machine tools, finishing robots, inspection pallets and sensors, and cooperatively develop the cooperation functions in order to construct the holonic machining system which integrates all of the machining units developed.

The manufacturing processes of the induction motors are set up as a common task among the participants and the concept of process primitives corresponding to the machining features are proposed for the holonic machining units in order to enable all the manufacturing processes in a holonic way.

The process planning tasks of the job holons representing the workpieces to be machined are considered in the research [4]. A process planning system is developed to carry out the following tasks: they are, extraction of the machining features to be machined, selection of the suitable machining equipment, and determination of the suitable machining sequence and the machining equipment. Some case studies are carried out to verify the effectiveness of the system for the products of the induction motors.

This work package also includes the holonic machine tool, the holonic finishing system and the holonic inspection system, as the basic holonic components for the machining of the products, the fixturing of the blank materials on the fixtures, and for the finishing of the final products.

(4) WP5 HOLOFIX (Holonic Fixturing Station)

The objective of WP5 is to develop a fixturing station named HOLOFIX as a component of HMS. The workpiece is fixed on the base plates and/or on the palettes autonomously at the HOLOFIX Station.

(5) WP6: Holonic Assembly

The objective of WP6 is to develop a holonic assembly system for the target products, in particular, for the induction motors. In the robot based assembly systems, the reconfiguration of the robotic system is in general very time and cost consuming tasks. If the reconfiguration can be carried out easily, the system can cope with the change of the production volumes and the occurrence of robot breakdowns rapidly, by adding or removing some robots in the system. Therefore, we consider the assurance of easy reconfiguration to be necessary

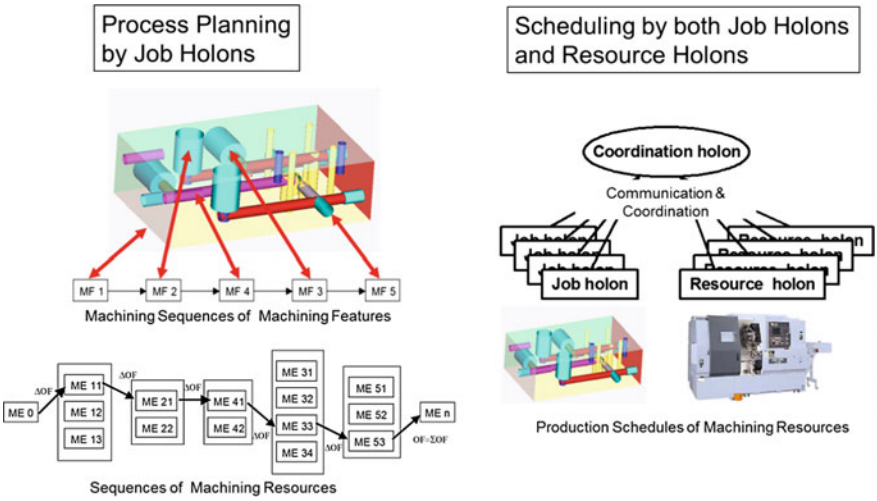


Fig. 4 Distributed process planning and production scheduling [3]

for the manufacturing assembly in the future, and call this function as "Plug & Produce," in comparison to the "Plug & Play" on personal computers.

Figure 4 summarizes the basic concepts discussed in WPs 3 and 4 for the process planning tasks for machining processes to be carried out by the job holons and the scheduling tasks to be carries out by both the job holons and the resource holons. These tasks have been mainly carried out by OPU.

3 Simulation System for Process Planning and Scheduling

3.1 Simulation System

The OPU research team has developed an object oriented simulation system for verifying the holonic process planning and production scheduling systems.

The simulation system consists of the following five parts.

(1) Manufacturing system model
The manufacturing system model consists of a set of objects representing the holons in the HMS. The model includes such models of the holons representing the machine tools, AGVs, storages, and jobs to be manufactured.

(2) Timing control system
The timing control system monitors all the events that occur in the manufacturing system model and controls the simulation time of all the holons and the objects in the simulation system. The timing control system has the event table, which describes all the events occurred in the manufacturing systems (Fig. 5).

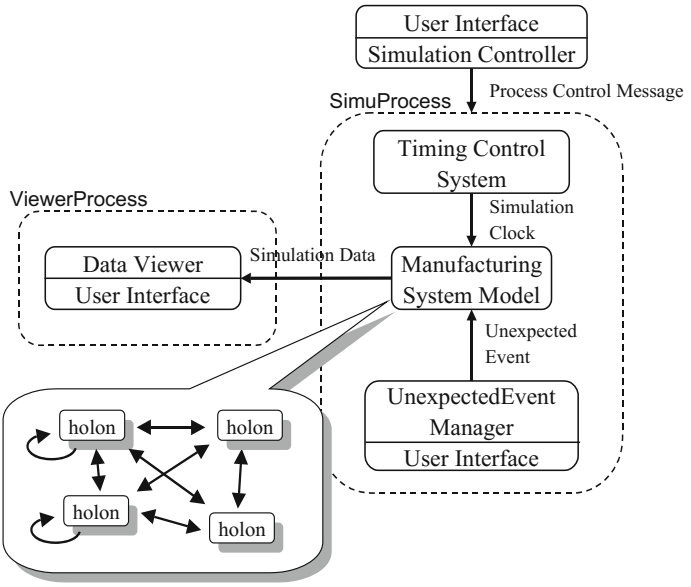


Fig. 5 Simulation system for process planning and production scheduling

(3) Simulation controller
The simulation controller is a user-interface for controlling the simulation system.
(4) Data viewer
The data viewer is a user-interface for reviewing the data representing the status of the individual holons.
(5) Unexpected event manager
The unexpected event manager is a user-interface for inputting unexpected disturbances directly to the simulation system, such as breakdown of resources and input of high priority jobs.

The simulation system has been implemented by using the Smalltalk language, and applied to the simulation of the real-time scheduling process of the holonic manufacturing testbed for the induction motor parts. In this application, the controller holons are developed for carrying out the decision-making process of the real-time scheduling.

3.2 Real-Time Scheduling Simulation

An object-oriented simulation system has been developed and applied to simulation of real-time scheduling in the holonic machining systems discussed in WP3 and 4.

The machining system has been designed for the induction motor parts manufacturing test-bed in the HMS consortium. The system consists of the following resources that are designed based on the machining processes of all the induction motor parts considered. They are, 12 turning centers, 3 machining centers, 3 grinding centers, one buffer, one input/output buffer and 7 AGVs. All the parts needed for assembling 3 types of motors are machined in the system.

The real-time scheduling considered here means that the production schedule is determined dynamically only when the status of the manufacturing system is changed due to the progress of the manufacturing processes and the unforeseen events. Therefore, the scheduling system only determines the schedules of the jobs and the resources for a very short time period. The scheduling procedure is activated when one of the following events occurs.

(1) A machining operation of a job is finished,
(2) A job is inputted to the input/output buffer,
(3) A resource is broken down, or is recovered, and
(4) A status of a job is changed from the normal one to the high priority one.

The scheduling system consists of a set of holons of the jobs and the resources. The production schedule is determined based on the decision-makings of the individual holons and their communications. A rule-based approach is adopted for determining the schedule. The scheduling rules are given to the job controllers who select the suitable resources for their machining operations.

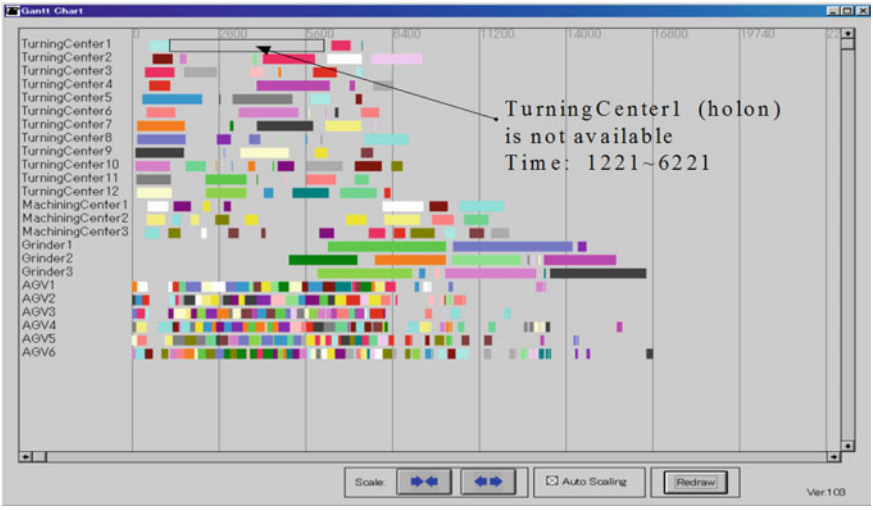
The machining process information about 30 jobs has been provided, and the real-time scheduling processes of the jobs have been simulated for the testbed. Figure 6a, b give the simulation results given by the Gantt chart which represents the production schedules of the resources obtained by applying the heuristic rules. The figures show that the proposed process planning and production scheduling systems are able to determine suitable process plans and production schedules for both the cases of the normal situation and the situations with unforeseen malfunctions of the resources.

4 Summary

The paper give a brief summary about the international collaborative project named "Holonic Manufacturing Systems" aimed at developing an autonomous distributed manufacturing systems looking toward the 21st Century.

(1) An autonomous distributed architecture was discussed and the basic concept about the holonic manufacturing systems was established, which is applicable to agile and flexible manufacturing processes for the machine products.
(2) A set of research projects have been carried out to develop a set of basic components for the holonic manufacturing systems. The topics considered here

(a) Without Machine Failure



(b) With Machine Failure

Fig. 6 Simulation results for process planning and scheduling. **a** Without machine failure. **b** With machine failure

include system architectures, resource allocations, process planning, production scheduling, machining, fixturing, and assemblying.

(3) A computer simulation system was developed to verify the holon based process planning and production scheduling methods proposed here and applied to the test bed problems. The simulation results show that the proposed methods generate suitable process plans and production schedules for the holonic manufacturing systems.

References

1. Manufacturing Science and Technology Center, *Annual report on Holonic Manufacturing Systems* (in Japanese) (2000)
2. N. Sugimura, HMS Japanese Member, Holonic Manufacturing System(HMS), in *Proceedings of 2000 Pacific Conference on Manufacturing*, vol. 2, (2000), pp. 989–996
3. N. Sugimura et al., Modeling of holonic manufacturing system and its application to real-time scheduling. Manuf. Syst. **25**(4), 1–8 (1996)
4. N. Sugimura et. al., A study on an integrated process planning and scheduling system based on shape generation functions, in *Proceedings of the 31st CIRP International Seminar on Manufacturing Systems* (1998), pp. 170–175

Development of Ultra–low Emission Multi–fuel Boiler System Using Plasma–Chemical Hybrid Clean Technology

Tomoyuki Kuroki, Hidekatsu Fujishima, Atsushi Tanaka, Keiichi Otsuka and Masaaki Okubo

Abstract The number of small boilers using city natural gas, heavy oil (HO), and waste oils has been increasing annually in Japan, and more stringent regulations for nitrogen oxides (NO_x) emission are being anticipated to reduce environmental NO_x concentration. Taking this into consideration, it is envisioned that a suitable flue gas treatment system for small boilers will be required. The author proposed a plasma–chemical hybrid clean technology, consisting of an indirect nonthermal plasma process and a wet-chemical treatment. The tested flue tube boiler has an original rotary burner for gas and/or oil and is operated using city natural gas, biomass oil, or HO. The boiler has a steam generation rate of 2.5 ton/h, and for the clean technology, two sets of silent discharge-type plasma ozonizers are employed to generate ozone. Using the combination of ozone injection and chemical scrubber for flue gas, NO_x can be effectively decomposed to nitrogen and oxygen, thus purifying the resulting effluent. The amount of nitrogen monoxide (NO) removed is almost the same as the amount of the corresponding ozone required to oxidize NO to nitrogen dioxide (NO_2) (1:1 stoichiometric ratio). Previously reported experimental data are also discussed in the paper. A NO_x removal efficiency of more than 85% was achieved over an operating time of 23 h using city natural gas as fuel. Based on the concept of carbon neutrality, $\sim$80% carbon dioxide (CO_2) reduction or fuel saving is also possible using waste vegetable oil (WVO)/HO mixed fuels. This low-emission boiler system can be used in industry.

Keywords Boiler emission · Plasma · Pollution control · Biofuel · NO_x

T. Kuroki · H. Fujishima · M. Okubo (✉)
Osaka Prefecture University, 1-1 Gakuen-Cho, Naka-Ku, Sakai 599-8531, Japan
e-mail: mokubo@me.osakafu-u.ac.jp

A. Tanaka · K. Otsuka
Takao Iron Works Co., Ltd, 1-29-3 Shimae-Cho, Toyonaka 561-0826, Osaka, Japan

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_19

1 Introduction

The use of small boilers employing city natural gas, heavy oil (HO), and waste oils as fuel is steadily increasing in Japan. Therefore, stringent regulations on nitrogen oxides (NO_x) emission are being anticipated to reduce environmental NO_x. It is therefore projected that a suitable flue gas treatment system for small boilers will be required soon. Several laboratory-scale studies have been conducted concerning nonthermal (NTP) plasma–chemical hybrid processes used for NO_x removal from gases emitted from various stationary sources. The authors propose a plasma–chemical hybrid process consisting of an indirect nonthermal plasma process and a wet-chemical treatment.

The principle of NO_x removal is as follows:

$$NO + O_3 \rightarrow NO_2 + O_2 \tag{1}$$

$$2NO_2 + 4Na_2SO_3 \rightarrow N_2 + 4Na_2SO_4 \tag{2}$$

In this study, the pilot-scale investigation of a plasma–chemical hybrid NO_x removal process was carried out using a low-emission boiler system comprising a multi-fuel boiler and a chemical scrubber [1–6]. Previously reported experimental data are discussed in this paper. In addition, the relationship between NO removal and the injected ozone, the combustion characteristics and NO_x removal performance when firing a mixture of HO and waste vegetable oil (WVO), and successive operation for four days are explained.

2 Experimental Setup

Figure 1 shows a schematic of the plant. The flue tube boiler (Takao Iron Works Co., Ltd.) has an original rotary burner for gas and/or oil and is operated by using city natural gas (13A) at 157 Nm^3/h and A-type HO at 171 L/h and waste oils. The boiler has a steam generation rate of 2.5 t/h. Two sets of silent-discharge-type ozonizers (EW-90Z, Ebara Jitsugyo Co., Ltd.), equipped with pressure swing adsorption (PSA) oxygen generators, are employed to generate O_3.

Figure 2 shows the structure of the discharge section of the plasma reactor in the ozonizer. A coaxial NTP reactor or silent-discharge-type reactor is used. O_3 gas is generated after O_2 gas produced by the PSA oxygen generators passes through the 30 plasma reactors inside the ozonizer and is then injected into a flue gas duct for NO oxidation. Table 1 shows the specifications of the ozonizer. The O_3 gas flow rate is constant at 0.9 Nm^3/h. When the discharge power is 1.5 kW and the PSA power is 1.6 kW, maximum O_3 generated is 90 g/h with a concentration of 4.7% and the energy efficiency of 29 g/kWh. High-concentration O_3 gas is injected into the duct and then diluted in order to oxidize almost all the NO to NO_2 according to

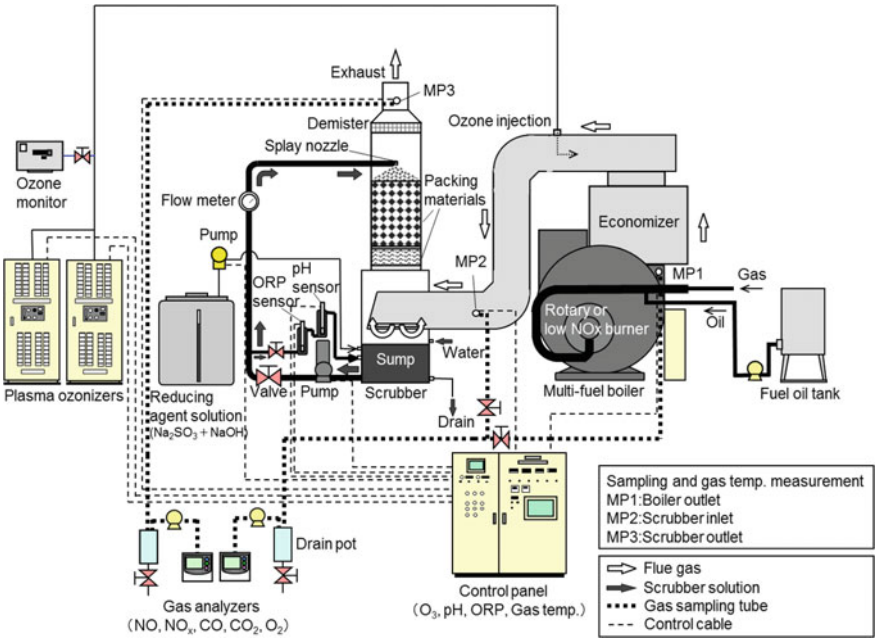


Fig. 1 Schematic diagram of the low-emission boiler system

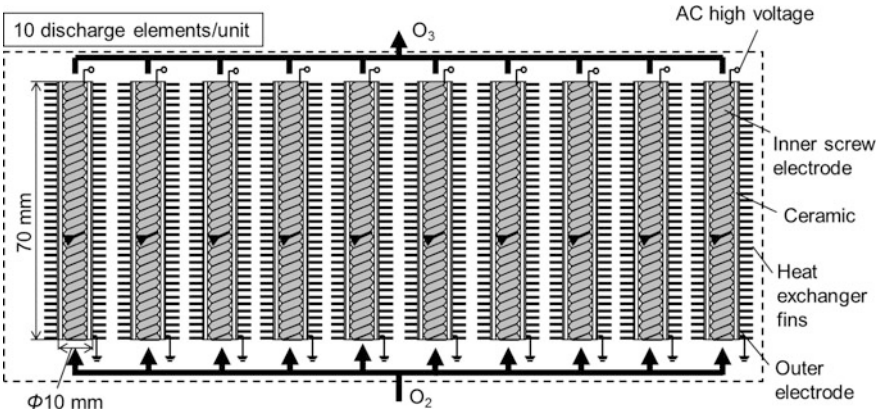


Fig. 2 Structure of the discharge elements in the ozonizer

reaction (1). The amount of nitrogen monoxide (NO) removed is almost the same as the amount of the corresponding ozone required to oxidize NO to nitrogen dioxide (NO_2) (1:1 stoichiometric ratio) [2, 4].The flue gas is then introduced into the scrubber to reduce NO_2 to N_2 according to reaction (2). An aqueous solution of Na_2SO_3 and NaOH (1.0 and 0.15 mol/L, respectively) is continuously added into

Table 1 Specifications of the ozonizer

Type	EW-90Z
O_2 generation	PSA (O_2 more than 90%)
O_3 generation	Air-cooled silent discharge plasma
O_3 mass flow rate	0–90 g/h
O_3 concentration	0–100 g/Nm3 = 0–4.7%
O_3 volume flow rate	0.9 Nm^3/h
Power consumption	1.6 kW (PSA), 1.5 kW (Discharge)
Dimension	740 mm W × 925 mm D × 1840 mm H

the sump from the solution tank at a flow rate of 0.04–0.12 m^3/h. The initial Na_2SO_3 concentration of the scrubbing solution in the sump is set to more than 0.13 mol/L. The scrubbing solution is pumped at a flow rate of 3 or 5 m^3/h from the sump to the top of the scrubber and sprayed through a nozzle over the packing material layer. It is known from reaction (2) that maintaining proper Na_2SO_3 concentration is important for continuous stable NO_2 reduction. Therefore, the oxidation reduction potential (ORP) and pH are monitored because ORP increases and pH decreases when Na_2SO_3 concentration decreases. The temperature of the flue gas and the concentrations of its constituents are measured at the following three sampling points: MP1 (boiler outlet), MP2 (scrubber inlet), and MP3 (scrubber outlet). Throughout this article, the NO_x concentration is defined as the converted value at 5 or 4% O_2 concentration in the case of the fuel gas or oil, respectively.

3 Results and Discussion

Experiments in a mixed oil of WVO and HO were performed to confirm the combustion characteristics and NO_x removal performance when firing the mixed oil. Figure 3 shows the time-dependent NO_x concentrations at MP1 and MP3 for HO and the mixed oils [1]. The volume mixing ratios of WVO and HO (WVO:HO) are 4:6 and 8:2. The flow rates of the flue gas are 890, 980, and 1030 Nm^3/h, respectively. O_3 is injected at a rate of 146–157 g/h and the average injection rate of Na_2SO_3 solution is 86.4–100 mol/h with the total flow rate of scrubbing solution of 3 m^3/h. The NO_x concentrations at MP1 for HO, mixed oils with 4:6 and 8:2 are approximately 102, 111, and 116 ppm, respectively. The combustion air should be increased for stable combustion when firing the mixed oils; therefore, NO_x concentration when firing the mixed oils is a little higher than that when firing HO. The NO_x concentrations at MP3 fluctuated within 150 min, but became stable thereafter, ~20 ppm for all cases. The average NO_x removal efficiencies were 85–86%. It proved that the WVO/HO mixed oil can be combusted satisfactorily up to the mixing ratio of WVO:HO = 8:2. Based on the concept of carbon neutrality, ~80% carbon dioxide (CO_2) reduction or fuel saving is possible using the mixed oil.

Figure 4 shows NO and NO_x concentrations at MP1 and MP3 as a function of successive operating time [7]. The experiment is carried out with daily start and stop operation. The average flow rate of city natural gas is 52 Nm^3/h and that of flue gas is 718 Nm^3/h at an O_2 concentration of 3.3–4.0%. The average rates of O_3 injection and Na_2SO_3 supply are 86 g/h and 63 mol/h, respectively. The total flow rate of scrubbing solution is 5 m^3/h. The NO_x concentrations at MP1 are 45–50 ppm. This small difference in NO_x each day is due to the fluctuations in the O_2 concentrations at MP1 between 3.3 and 4.0%. The NO_x concentrations at MP3 are 5–8 ppm, equivalent to a NO_x removal efficiency of more than 85% over an operating time of 23 h. The NO_x removal performance achieved is satisfactory and stable. The specific gravity of the scrubbing solution increases gradually and is stable at $\sim$1.08, which is an allowable value as a control index.

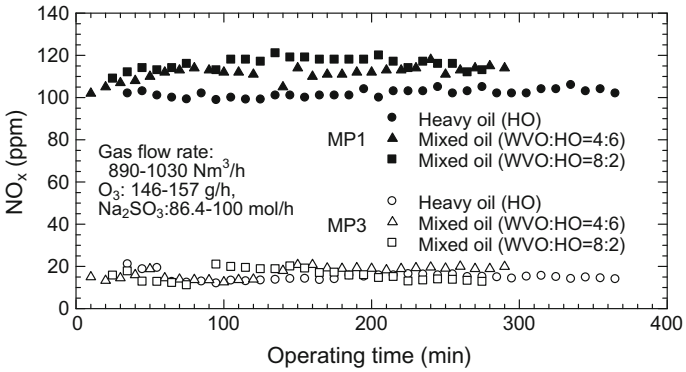


Fig. 3 NO_x concentrations at MP1 and MP3 as a function of operating time when firing HO and WVO mixed oils (MP1: boiler outlet and MP3: scrubber outlet in Fig. 1)

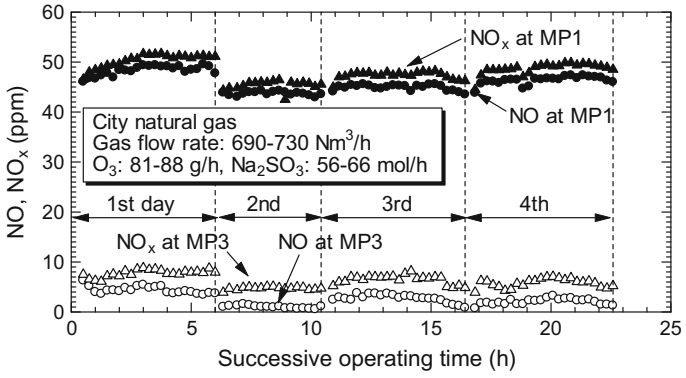


Fig. 4 Time-dependent NO and NO_x concentrations at MP1 and MP3 during a four successive day operation in firing city natural gas (MP1: boiler outlet and MP3: scrubber outlet in Fig. 1)

4 Conclusions

The pilot-scale investigation of a plasma–chemical hybrid NO_x removal process is carried out using a low-emission boiler system comprising a multi-fuel boiler and a chemical scrubber. The amount of NO removed is almost the same as the amount of the corresponding O_3 concentration needed to oxidize NO to NO_2 (1:1 stoichiometric ratio), regardless of the flue gas flow rate. The WVO/HO mixed oil could be combusted satisfactorily for up to 80% mixture ratio and average NO_x removal efficiencies of 85–86% are obtained. Based on the concept of carbon neutrality, ~80% CO_2 reduction or fuel saving is possible using the mixed oil. NO_x removal efficiency of more than 85% is achieved over an operating time of 23 h when firing city natural gas. It is thus confirmed that the low-emission multi-fuel boiler system with plasma–chemical hybrid NO_x aftertreatment is applicable for industrial use.

Acknowledgements The authors thank Mr. K. Takekoshi, Mr. Y. Nakagishi, Mr. Y. Fukumoto and Mr. H. Mochizuki, who were all graduate students at the Osaka Prefecture University, for assistance with carrying out the experiments. Some expressions and content are reused from my previous published open access paper [7] for this article. This study was supported by the Regional Research and Development Resources Utilization Program conducted in the Japan Society Technology Agency (JST).

References

1. H. Fujishima, K. Takekoshi, T. Kuroki, A. Tanaka, K. Otsuka, M. Okubo, Appl. Energy **111**, 394–400 (2013)
2. H. Fujishima, Y. Fukumoto, T. Kuroki, A. Tanaka, K. Otsuka, M. Okubo, International Journal of Plasma. Environ. Sci. Technol. **8**(1), 45–51 (2014)
3. H. Fujishima, Y. Yoshioka, T. Kuroki, A. Tanaka, K. Otsuka, M. Okubo, I.E.E.E. Trans, Ind. Appl. **47**(5), 2210–2217 (2011)
4. H. Fujishima, T. Kuroki, T. Ito, K. Otsuka, T. Yamamoto, K. Yoshida, M. Okubo, I.E.E.E. Trans, Ind. Appl. **46**(5), 1707–1714 (2010)
5. H. Fujishima, A. Tatsumi, T. Kuroki, A. Tanaka, K. Otsuka, T. Yamamoto, M. Okubo, I.E.E.E. Trans, Ind. Appl. **46**(5), 1722–1729 (2010)
6. T. Yamamoto, H. Fujishima, M. Okubo, T. Kuroki, I.E.E.E. Trans, Ind. Appl. **46**(1), 29–37 (2010)
7. M. Okubo, Development of Ultra-Low Emission Multi-Fuel Boiler System Using Plasma-Chemical Hybrid Clean Technology. J. Phys. Conf. Series **646**, 012051 (2010). doi:10.1088/1742-6596/646/1/012051

Design and Analysis on a Novel Electric Vehicle Powertrain

Jie Yu, Ligang Yao, Chengcheng Ren and Xiaolei Yan

Abstract In order to obtain synthetically advantages from the independent multiple-motors-driven system and the centralized single-motor-driven system, a novel centralized dual-motor-coupling powertrain (DMCP) system with a single row double pinions gear train for electric vehicle (EV) is proposed. By means of changing the drive modes, the proposed powertrain can satisfy the demands of different running conditions with high energy utilization efficiency. In this paper, the analyses on the mechanism structure and operating principles of the proposed powertrain are carried out firstly. Then, for further improving the operating efficiency and providing the effective ways for optimal matching design of the novel DMCP, the parameters sensitivity analysis is studied under the certain tested driving cycle based on a systematic simulation model. The simulation result shows that the energy consumption is increased by 3% as compared to the result of automobile test with the single motor powertrain.

Keywords Dual-motor · Electric vehicle · Powertrain · Simulation

1 Introduction

Since the powertrain system design has an important influence on dynamic performance and fuel economy for electric vehicle (EV) [1], the centralized dual-motor-driven coupling powertrain (DMCP) is proposed recently except the independent multiple-motors-driven system and the centralized single-motor-driven system [2]. Although the in-wheel electric motor powertrain can simplify the structure

J. Yu · L. Yao (✉) · C. Ren
School of Mechanical Engineering and Automation,
Fuzhou University, Fuzhou, P. R. China
e-mail: ylgyao@fzu.edu.cn

J. Yu · X. Yan
School of Mechanical and Automobile Engineering,
Fujian University of Technology, Fuzhou, China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_20

of the drivetrain, it is still difficult to overcome the shortcomings of high control complexity and manufacture level requirement [3]. On the other hand, the centralized single-motor-driven system has been applied on most modern EVs because of its reliability. However, it claims the need of ratio changing mechanism to improve motor's speed-torque characteristics, which involves more cost and shift control issues [4]. Hence, some new drivetrain with wider torque/speed characteristic and efficient operating region as well as high reliability are desirable for EV. For this reason, new powertrains based on centralized dual-motor-driven are proposed [2, 5–8]. However, some of the existing structures can only realize speed coupling function of two motors, and the others with speed and torque coupling function have complex structure. Based on the preliminary research, a novel DMCP with simple mechanical structure is designed in this paper to expend the speed-torque characteristics of motors and to reduce energy consumption. First, the mechanism structure and operating principles of the powertrain are introduced and analyzed in the remainder of this paper. Then, the parameters sensitivity is studied under certain tested driving cycle based on a systematic simulation model, which aims at offers effective ways for optimal matching design of the novel DMCP. And Multi-Island Genetic Algorithm (MIGA) is used to obtain the optimal value for parameters.

2 Structure and Operating Principles for the DMCP

For the understanding of novel powertrain system, the structure of DMCP is illustrated in Fig. 1, which is composed of two motors (M1 and M2), a single row double pinions gear train (DEG), two clutches (CL1 and CL2) and a brake (B). In this powertrain, the main-drive motor (M1) is connected to the sun gear of the DPG and could connect to the carrier when CL1 and CL2 both are engaged, while the auxiliary-drive motor (M2) could connect to the sun gear or the carrier based on the position of two clutches. As the three rotational motions (Sun, Ring, Carrier) of DPG been controlled by switching the position of CL1, CL2 and B, the DMCP can work in different modes, which have been shown in Table 1. The tractive force form two motors is coupled by the DPG unit and then exported from the ring to the final drive. As shown in Table 1, the DMCP can operate in five driving modes: low-speed single motor drive (SM), low-speed dual-motor torque coupling drive (LTC), Moderate-speed single motor drive (MSM), Moderate-speed dual-motor torque coupling drive (MTC), and high-speed dual-motor speed coupling drive (HSC). By means of changing drive modes, the powertrain can satisfy the demands of different running conditions with high energy utilization efficiency.

It is worth mentioning that the model shift for DMCP is ease to realize. As the vehicle acceleration process (From low-speed to high-speed), only one of the CL1, CL2 and B change status each time to realize a new model. This means that the control for DMCP is easy and flexible to design.

To reveal the operating principles of DMCP, the kinematics equation of DPG is proposed as follows by ignoring the energy loss:

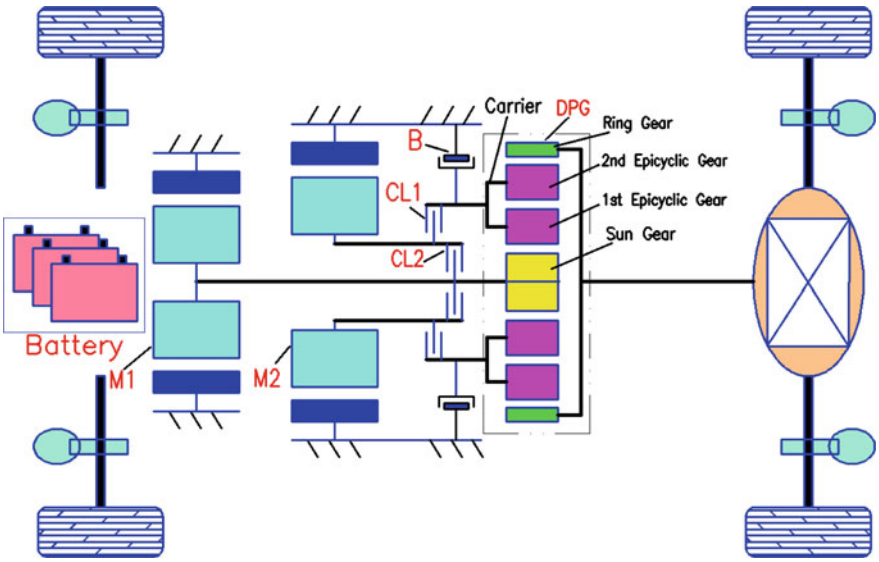


Fig. 1 Schematic of the DMCP

Table 1 System operation schedule

Modes	M1	M2	CL1	CL2	B	Driving state
LSM	●	○	○	○	●	Low Speed Light Load
LTC	●	●	○	●	●	Low Speed Heavy Load
MSM	●	○	●	●	○	Middle Speed Light Load
MTC	●	●	●	●	○	Middle Speed Heavy Load
HSC	●	●	●	○	○	High Speed

Note "●" Indicates that the motor is on and CL/B is engaged
"○" Indicates that the motor is off and CL/B is disengaged

$$\begin{cases} \omega_s + (k-1)\omega_c = k\omega_r \\ T_s = \frac{1}{k-1} T_c = -\frac{1}{k} T_r \end{cases} \tag{1}$$

where ω_s denotes the sun gear speed, ω_c is the carrier speed, ω_r is the ring gear speed, k denotes the characteristic parameter of DPG, its value is Z_r/Z_s.

Then, based on the Eq. (1), the operating principles of the DMCP were briefly analyzed as follows:

$$\text{LSM mode} \quad \begin{cases} \sum I_w \dot{\omega}_w + I_{m1} \dot{\omega}_{m1} k i_0 = T_{m1} k i_0 \eta_t - \sum T_{req} \\ \omega_{m1} = \omega_s = k i_0 \omega_w \\ \omega_{m2} = 0 \end{cases} \tag{2}$$

$$\text{LTC mode} \quad \begin{cases} \sum I_w \dot{\omega}_w + (I_{m1} + I_{m2}) \dot{\omega}_{m1} k i_0 = (T_{m1} + T_{m2}) k i_0 \eta_t - \sum T_{req} \\ \omega_{m1} = \omega_{m2} = \omega_s = k i_0 \omega_w \end{cases} \tag{3}$$

$$\text{MSM mode} \quad \begin{cases} \sum I_w \dot{\omega}_w + I_{m1} \dot{\omega}_{m1} i_0 = T_{m1} i_0 \eta_t - \sum T_{req} \\ \omega_{m1} = \omega_{m2} = \omega_s = \omega_r = i_0 \omega_w \end{cases} \tag{4}$$

$$\text{MTC mode} \quad \begin{cases} \sum I_w \dot{\omega}_w + (I_{m1} + I_{m2}) \dot{\omega}_{m1} i_0 = (T_{m1} + T_{m2}) i_0 \eta_t - \sum T_{req} \\ \omega_{m1} = \omega_{m2} = \omega_s = \omega_r = \omega_c = k i_0 \omega_w \end{cases} \tag{5}$$

$$\text{LTC mode} \quad \begin{cases} \sum I_w \dot{\omega}_w + \left(I_{m1} \dot{\omega}_{m1} + \frac{1}{k-1} I_{m2} \dot{\omega}_{m2} \right) k i_0 = min\left(T_{m1}, \frac{1}{k-1} T_{m2} \right) k i_0 \eta_t - \sum T_{req} \\ \omega_{m1} + (k-1) \omega_{m2} = k i_0 \omega_w \end{cases} \tag{6}$$

where $\sum I_w$ is the total wheel moment of inertia, I_{m1} and I_{m2} denote the moment of M1 and M2, ω_{m1} and ω_{m2} denote the speed of M1 and M2, T_{m1} and T_{m2} denote the torque of M1and M2, ω_w is the wheel speed, i_0 is the final drive gear ratio, $\sum T_{req}$ denotes the total required torque at the output axle, η_t represents the total drivetrain efficiency.

It should be noted that the torque of the ring is $k \cdot min\left(T_{m1}, \frac{1}{k-1} T_{m2}\right)$ in the Eq. (6) because of the limit of DPG kinematic equation.

3 Optimal Matching for the DMCP

In order to reveal the characteristics of the novel DMCP in dynamic and efficiency performance, an electric bus is selected as the target vehicle, which parameters are shown in Table 2. And as shown in Fig. 2, the real driving cycle and the required power is test on a bus. Based on the test data analysis, the commonly used speed range of city bus is about 0–10 km/h and 20–35 km/h. And the peak speed almost at 43 km/h with required power about 150 kW. Moreover, the kilowatt consumption per 1 km is about 8.3 with the single motor powertrain system.

The parameters needed to design for the DMCP are: the motor parameters, the characteristic parameter of DPG and the final drive gear ratio. And the powertrain parameters matching EV mainly have two principles: meeting the dynamic performance requirements (as is shown in Table 3) and high efficiency optimal matching design.

Table 2 Vehicle parameters

Components	Parameters	Values
	Curb Mass (Kg)	11,400
	Gross Mass (Kg)	18,000
Vehicle body	Frontal Area (m^2)	7.8
	Drag Coefficient	0.6
	Rolling Radius (m)	0.436
	Final drive gear ratio	6.166
Battery pack	Rated Capacity (A · H)	412 AH
	Rated Voltage (V)	432 V

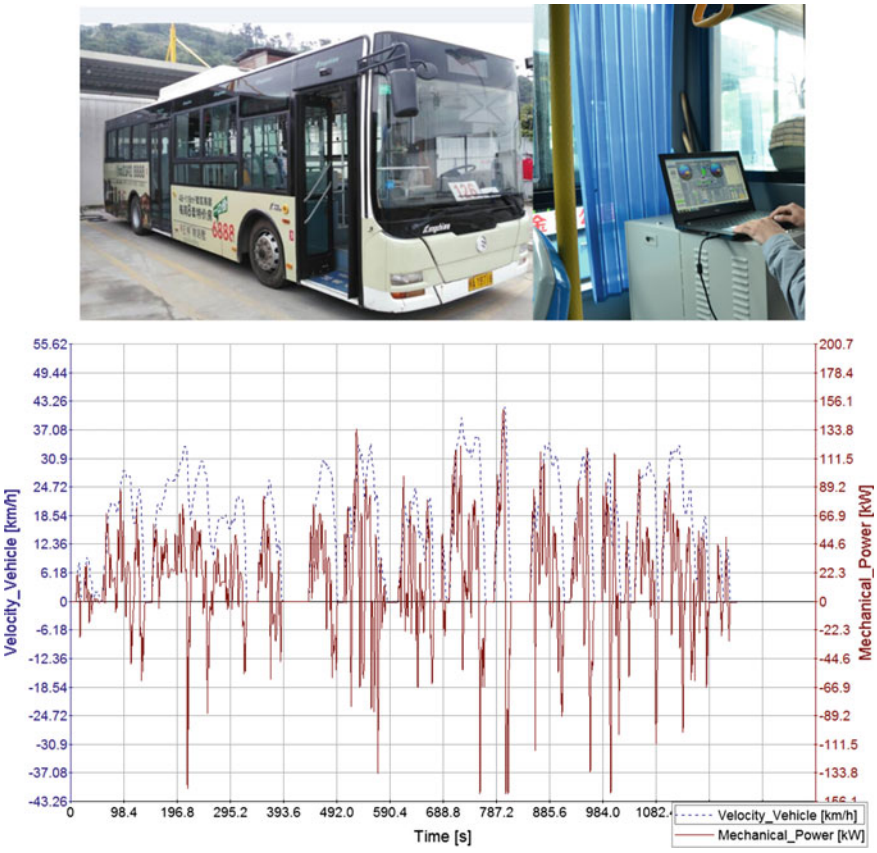


Fig. 2 Vehicle drive cycle and power required for a city bus

Table 3 The dynamic performance requirements

Name	Value
Maximum speed (km/h)	≥ 65
Maximum slope at 10km/h vehicle speed (%)	≥ 30
0–50 km/h acceleration time t_{50} (s)	≤ 15

For the first principle, the total power must satisfy the following equation:

$$P_{1max} + P_{2max} \geq max\left(P_v, P_i, P_j\right) \tag{7}$$

$$max\left(T_{di}, P_{dj}\right) \leq T_{1max} + T_{2max} \leq F_\varphi \cdot r/ki_0 \tag{8}$$

$$max\left(T_{di}, P_{dj}\right) \leq T_{1max} + T_{2max} \leq F_\varphi \cdot r/ki_0 \tag{9}$$

$$T_{1max} = 9549P_{1max}/\omega_{1max} \tag{10}$$

$$T_{2max} = 9549P_{2max}/\omega_{2max} \tag{11}$$

where P_{1max} is the peak power of M1, P_{2max} is the peak power of M2, P_v is the required power at maximum speed, P_i is the required power at maximum slope, P_j is the required power for $t_{50} = 15$, T_{1max} is the maximum torque of M1, T_{2max} is the maximum torque of M2, T_{di} is the required torque for maximum slope, T_{dj} is the required torque for acceleration, F_φ is maximum force under the road adhesion condition, r is the wheel radius, u_{max} is peak speed for vehicle, ω_{1max} is the peak speed of M1, ω_{2max} is the peak speed of M2.

For the second principle, the rated power P_{1e} and speed ω_{1e} should be fit to the commonly used speed range, from which the rated speed and power of M1 can be calculated under certain driving cycle. Then, rated torque of M1 T_{1e} can be calculated as follow:

$$T_{1e} = 9549P_{1e}/\omega_{1e} \tag{12}$$

From the Eqs. (7)–(12), the two motors' parameters of DMCP can be obtained under the test cycle (as shown in Table 4), and the values of k and i_0 are calculated through optimal matching method, which Multi-Island Genetic Algorithm (MIGA) is used for the multi-objective optimization problem. The weight method is used to obtain dynamic performance optimal objective function from the peak speed, 0–50 km/h acceleration time and peak slope, meanwhile the kilowatt consumption per 1 km is taken as the economic optimal objective function. The optimal calculating flow is shown in Fig. 3, and the optimal result is shown in Fig. 4. When the value of k is 2.3 and the value of i_0 is 6.21, the kilowatt consumption per 1 km is about 8.0 and the dynamic performance satisfies the design requirements. And the Fig. 5 shows that

Table 4 The parameters of DMCP

Components	Name	Value
	Rated/peak power (kW)	70/141
M1	Rated/peak speed (rpm)	1500/3000
	Rated/peak torque (N · m)	30/60
	Rated/peak power (kW)	45/90
M2	Rated/peak speed (rpm)	1800/3500
	Rated/peak torque (N · m)	160/320

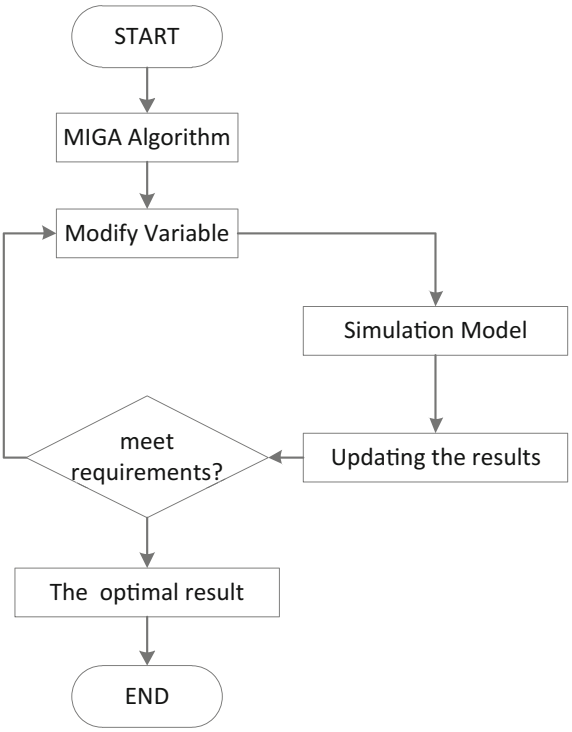


Fig. 3 The flow of MIGA optimal calculation

the current velocity can follow the desired velocity well. Besides, with the analysis on the parameters sensitivity of DMCP, it shows that the characteristic parameter k has more influence on system efficiency and peak slope, and the final drive gear ratio i_0 has more influence on the 0–50 km/h acceleration time and peak speed.

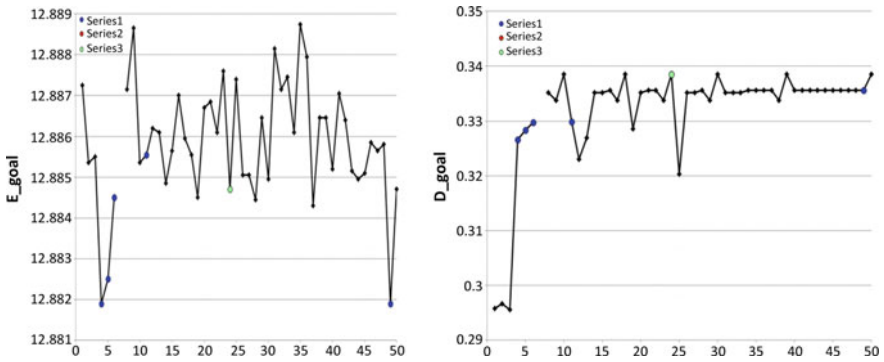


Fig. 4 The value of parameters and optimal goals with the number of iterations

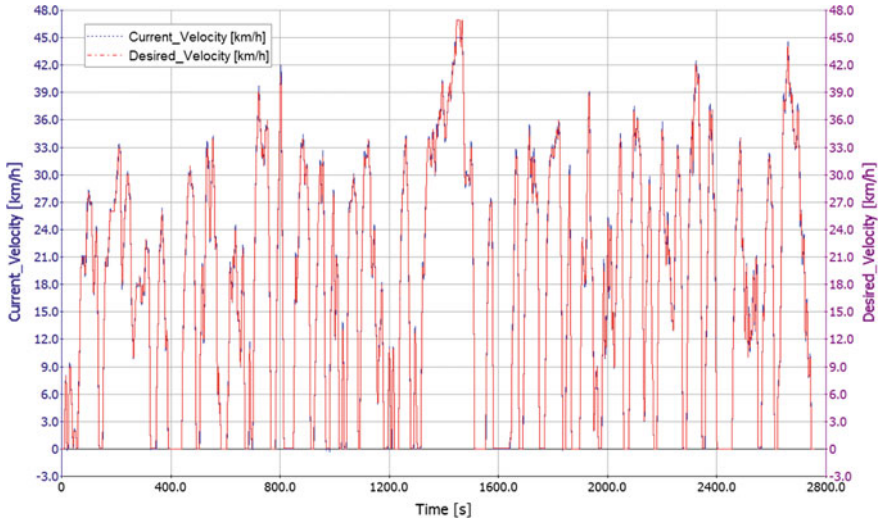


Fig. 5 The current velocity and desired velocity in simulation

4 Conclusions

In order to improve the performance on the all-electric vehicle, a dual-motor coupling powertrain system is provided with two motors, DPG, two clutches and a brake. Based on the structure and operating principles analyzed for the DMCP, the parameters are designed to follow two principles about meeting the dynamic performance requirements and high efficiency optimal matching design. The simulation results show that the energy consumption is increased by 3%, which proves the application value of the novel powertrain.

References

1. M. Ehsani, Y. Gao, A. Emadi, *Modern electric, hybrid electric, and fuel cell vehicles-fundamentals, theory, and design*, 2nd edn. (China Machine Press, Beijing, 2012), pp. 89–104
2. M. Hu, J. Zeng, S. Xu, C. Fu, D. Qin, Efficiency study of a dual-motor coupling EV Powertrain. IEEE Trans. Veh. Technol. **64**(6), 2252–2260 (2015)
3. H. Alipour, M. Sabahi, M. Bagher, Lateral stabilization of a four wheel independent drive electric vehicle on slippery roads. Original Res. Article Mechatron. **30**, 275–285 (2015)
4. Q. Ren, D.A. Crolla, A. Morris, *Effect of geared transmissions on electric vehicle drivetrains* (IDETC/CIE. Conference, San Diego, CA, USA, 2009), pp. 959–966
5. S. Zhang, R. Xiong, C. Zhang, Pontryagin's minimum principle-based power management of a dual-motor-driven electric bus. Appl. Energy **159**, 370–380 (2015)
6. L. Zhang, L. Li, B. Qi, C. Wang, Y. Zhang, J. Song, Parameters optimum matching of pure electric vehicle dual-mode coupling drive system. Technol. Sci./Appl. Energy. **57**(11), 2265–2277 (2014)
7. G. Han, C. Zhang, S. Zhang, X. Wu, Control strategy of the dual motors coupling propulsion for battery electric buses. Appl. Mech. Mater. **528**, 364–370 (2014)
8. X. Wu, S. Han, H. Sun. Efficiency analysis of planetary coupling drive system with dual motors on electric bus. in *Proceedings of 2015 IEEE International Conference on Mechatronics and Automation*, 2547–2552

Hybrid Proton Exchange Membrane Fuel Cell/Lithium-ion Battery System Power Management Strategy Design for Lifetime Extending of the Main Power Source

Fei-Fei Qin and Ya-Xiong Wang

Abstract Proton exchange membrane fuel cell (PEMFC) is regarded as one of the most promising power solutions for driving electric vehicle (EV) depending on its high efficiency, high energy density, low operation temperature, as well as zero emission. The limitations of wide application and commercialization of PEMFC in EV are low power density, unidirectional energy flow that cannot meet EV transient operating condition as well as store regenerating energy. Moreover, the transient load of EV may cause PEMFC lifetime dramatically degrade. Lithium-ion battery (LIB) is considered used in this study to hybridize with PEMFC to supply transient power demand as well as to save extra energy. To distribute power and to extend PEMFC lifetime, power management strategy based on the PEMFC operation point fixing is designed and implemented. The proposed hybrid PEMFC/LIB power system comprised by PEMFC, LIB, boost converter as well as bidirectional converter is modeled in Matlab/Simulink. Subsequently, the coordinated current-voltage control and cascade current-voltage control power splitting methods are developed to regulate DC bus voltage and PEMFC stack current. The fixed PEMFC stack current can lead to PEMFC operate in a suitable condition even the external load variations. The simulation and experimental results demonstrate that the designed power management strategy can balance the load demand and protect PEMFC.

Keywords Hybrid PEMFC/LIB power system · Power management · DC bus voltage and PEMFC current regulation · PEMFC lifetime extending

F.-F. Qin
School of Information Science and Engineering, Wenzhou Business College, Chashan Higher Education Park, Wenzhou 325035, Zhejiang, China

Y.-X. Wang (✉)
School of Mechanical Engineering and Automation, Fuzhou University, 2 Xueyuan Road, Minhou, Fuzhou 350116, Fujian, China
e-mail: yxwang@fzu.edu.cn

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_21

1 Introduction

The clean alternative fuel vehicles have widely been paid attention by academy and industrial communities. In particular, electric vehicles (EVs), including battery electric vehicle (BEV), hybrid electric vehicle (HEV) as well as fuel cell vehicle (FCV), have demonstrated many advantages in terms of high efficiency, low or zero emission, low noise, simple structure, and convenient maintenance [1–3]. Compared to BEV and HEV, proton exchange membrane fuel cell (PEMFC)-based FCV has a promise of high energy density that can lead to a long driving range. However, PEMFC is insufficient of feeding transient power to meet sudden load variation, due to its relatively slow electrochemical reaction process. To address the low power density limitation as well as unidirectional power flow, the power solutions of FCV are usually consist of PEMFC (as the main power source) and energy storage device (battery and/or supercapacitor as the additional power source). Since lithium-ion battery (LIB) has exhibited acceptable power density as well as good energy density to ensure both dynamic performance and cruising ability, LIB is the preferred additional power supply for configuring the hybrid power train of FCV [4].

Moreover, power converters are usually used to regulate the voltage and current to meet the motor operation requirements under different driving conditions. To obtain good speed and torque governing performance as well as overload capacity, AC motors are always utilized in EVs. Thus, DC-AC inverter is required for transforming the hybrid power system generated/saved electricity. The polarization characteristics of PEMFC suggest that the stack current change makes output voltage varied significantly [5]. The unregulated voltage of PEMFC may result in DC bus voltage unstable, which is not good for subsequent DC-AC inverter operating. The DC bus voltage is also higher than PEMFC output voltage, since using higher DC bus voltage can lead to high efficiency of the motor. Therefore, DC-DC boost converter is applied by connecting PEMFC output to regulate and to step up PEMFC voltage to satisfy request of DC bus voltage [6]. Furthermore, the bidirectional DC-DC converter is used for governing LIB charge and discharge as well [7]. Consequently, the dual-converter based hybrid PEMFC/LIB power system is composed of PEMFC, LIB, DC-DC boost converter and bidirectional DC-DC converter.

In order to efficiently implement the presented hybrid power system, the delicate power management is required. Many power management methods have been reported in the literatures. A maximum fuel economy-oriented power management approach based on PEMFC average power and energy storage device SOC control was developed to meet the load demand as well as to achieve the high efficiency of the hybrid power train [8]. An optimal control scheme according to the Minimum Principle was designed to minimize fuel consumption and to optimize the power distribution [9]. Similarly, Pontryagin's Minimal Principle was applied to address real-time operation cost optimization of a plug in PEMFC city bus by comprehensively analyzing battery SOC and the working points of the PEMFC system

[10]. Predictive control has been recently proposed to fulfill the energy management of the hybrid PEMFC power system [11–13]. The reference currents of PEMFC and LIB were obtained through the predictive controller, and PI two loops control was subsequently used to stabilize DC bus voltage as well as to track the setpoints of currents [11]. A model predictive control method that optimally distributed power between PEMFC and supercapacitor was presented to respond to rapid variations in load by DC bus voltage and the currents of the power sources regulating [12]. To protect PEMFC and battery by keeping fuel cell's and batteries' current slope in reference to certain values, as well as to attain a stable DC bus voltage, model predictive control-based strategy was studied to generate reference current for individual power supply, and tracked using hysteresis control in the simulation and experiment [13]. Rule-based fuzzy logic scheme has been widely discussed in power management of the hybrid power system [14–16]. An energy management system based on fuzzy logic control was demonstrated to obtain the high efficiency operation region of each energy carrier and to maintain the DC output voltage stable as well as to possess fuel cell and battery currents to follow the load demand within a limited slope [14]. To contribute to a longer lifetime of PEMFC and maintain a suitable battery SOC, a fuzzy logic-based decision-making strategy was employed by considering shutdown and overcurrent conditions [15]. For the sake of meeting the power requirement under the unknown driving cycles and balancing the power distribution among different power supplies, a real-time fuzzy logic controller was adapted, and PI inner current control was used for regulating currents of PEMFC and supercapacitor [16]. Additionally, the combinational control methods for the PEMFC-based hybrid power system, for example, flatness control with fuzzy logic control [17], wavelet transform [18, 19], and optimization approach based on stochastic dynamic programming [20], have been increasingly produced to solve the complicated control problems such as nonlinearity, uncertainty, and multi-objective.

In this study, a lifetime extending control scheme for PEMFC based on the coordinated current-voltage control and cascade current-voltage control of the hybrid PEMFC/LIB power system is proposed. In particular, the presented method is designed for stabilizing DC bus voltage as well as current of PEMFC while the load changes with driving cycle. The rest of this paper is organized as follows. Section 2 describes the model of the hybrid PEMFC/LIB power system. Lifetime extending control strategy is presented in Sect. 3. Section 4 provides the simulation and experiment results of the developed approach. The conclusion is drawn in the last section.

2 Hybrid PEMFC/LIB Power System Modeling

The presented hybrid PEMFC/LIB power system driving AC motor used for propelling FCV can be depicted in Fig. 1. The components of the DC power system and their functions are described as following: PEMFC is severing as the main

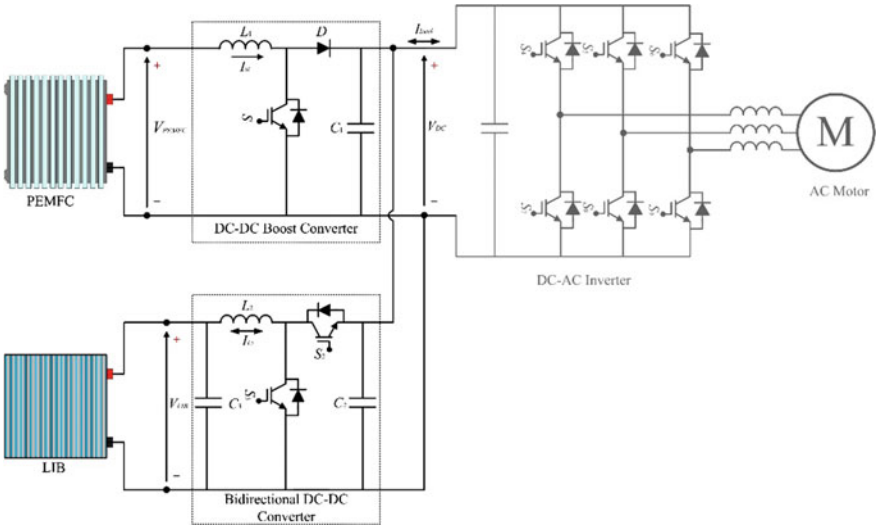


Fig. 1 Hybrid PEMFC/LIB power system driving AC motor schematic diagram

power source with boost converter regulating output voltage, and LIB is acting as the energy storage device to supply and save extra power as well as feeding transient load demand working together with bidirectional converter. Then, the hybrid PEMFC/LIB power system modeling can be comprised by PEMFC subsystem modeling, LIB subsystem modeling, boost converter and bidirectional converter modeling. The modeling method of the power sources is depended on equivalent electrical circuit to represent the relations between input currents and output voltages.

For PEMFC modeling

The polarization curve of PEMFC reveals the output voltage, V_{PEMFC}, changing with the stack current I_{st}, and the relation can be denoted by [21]

$$V_{PEMFC} = E_{ov} - nA\ln\left(\frac{I_{st}}{I_0}\right) \times \left(\frac{1}{sT_d/3+1}\right) - R_{ohm}\, I_{st} \tag{1}$$

where E_{ov} is the open-circuit voltage determined by Nernst voltage, n is the number of cells, A is the Tafel slope, I_0 is the exchange current, and the detail calculation of above parameters can be extracted in the reference [21]. The transient phenomenon of PEMFC is represented by using a first-order transfer function being equivalent of a parallel RC branch with the stack settling time, T_d. The last term in above equation describes the ohmic losses of PEMFC, due to electrodes and electrolyte resistances [21].

For LIB modeling

Since LIB supplies power when it discharges, and saves energy while it is charging, there are two functions provided to define the model of LIB [22]

$$V_{LIB} = E_0 - K\frac{Q_M}{Q_M - q}i^* - K\frac{Q_M}{Q_M - q}q + A\exp(-B \cdot q) \tag{2}$$

$$V_{LIB} = E_0 - K\frac{Q_M}{q + 0.1Q_M}i^* - K\frac{Q_M}{Q_M - q}q + A\exp(-B \cdot q) \tag{3}$$

in which, V_{LIB} is the output voltage of LIB, E_0 is the constant voltage, the polarization constant is denoted by K, Q_M represents the maximum capacity of LIB, the extracted capacitor and the low frequency current dynamics are presented as q and i^*, respectively, A and B are the exponential voltage and capacity [22].

For DC-DC converters modeling

In this study, there are two types of DC-DC converters, namely, boost converter and bidirectional converter. The boost converter is applied for stepping up PEMFC voltage as well as fixing PEMFC stack current. The model of the boost converter can be derived from basic electrical circuit laws- Kirchhoff's voltage law and Kirchhoff's current law. The governing equations are given by

$$\frac{dI_{st}}{dt} = -(1 - u_1)\frac{V_{DC}}{L_1} + \frac{V_{PEMFC}}{R_1} \tag{4}$$

$$\frac{dV_{DC}}{dt} = (1 - u_1)\frac{I_{st}}{C_1} - \frac{V_{DC}}{R_1 C_1} \tag{5}$$

where L_1 and C_1 are the inductance and capacitance used in the boost converter, respectively, R_1 denotes the resistance parallel connected with the capacitor C_1 branch, u_1 is the control input representing the function of the switch S_1 and taking a value in the set of [0, 1] and V_{DC} is the DC bus voltage.

The bidirectional converter is regarded as two operation modes: boost mode and buck mode. The boost mode is similar to boost converter, and the model can be drawn from Eqs. (4) and (5). The buck mode is modeled as

$$\frac{dI_{L_2}}{dt} = -\frac{V_{LIB}}{L_2} + u_3\frac{V_{DC}}{L_2} \tag{6}$$

$$\frac{dV_{LIB}}{dt} = \frac{I_{L_2}}{C_3} - \frac{V_{DC}}{R_3 C_3} \tag{7}$$

In above equations, L_2 and C_3 are the inductance and capacitance used for the buck mode of the bidirectional converter, respectively, R_3 represents the resistance parallel connected with the capacitor C_3 branch, I_{L_2} is the current of inductor L_2, u_3 is the control input of the switch S_3. The model of the hybrid PEMFC/LIB power system is then constructed in Matlab/Simulink platform.

3 Power Management Strategy for PEMFC Lifetime Extending

To efficiently and safely implement the hybrid PEMFC/LIB power system, a power management strategy is developed for stabilizing DC bus voltage as well as regulating PEMFC current. The control strategy is depending on two proposed power splitting structures: coordinated current-voltage control and cascade current-voltage control. For the coordinated current-voltage control, PEMFC and LIB act as the power supplies for feeding high load demands. In this case, the boost converter connecting with PEMFC is used for regulating PEMFC current, and LIB-fed bidirectional converter is working on the boost mode and applied for stabilizing DC bus voltage. The coordinated current-voltage control structure is shown in Fig. 2. For the cascade current-voltage control, PEMFC is supplying power for LIB and load. LIB serves as the energy storage device to save the extra power, and the bidirectional converter operates on the buck mode. The PEMFC-fed boost converter has two control loops, the inner loop regulating PEMFC current, and the outer loop fixing DC bus voltage as a constant value. The cascade current-voltage control is displayed in Fig. 3. The proposed power splitting methods can reduce the PEMFC current fluctuate in the load variation conditions. Moreover, the PEMFC operation point is fixed at the suitable range to ensure the high efficiency as well as to avoid high current slope which may result in PEMFC starvation. The lifetime of PEMFC can be extended.

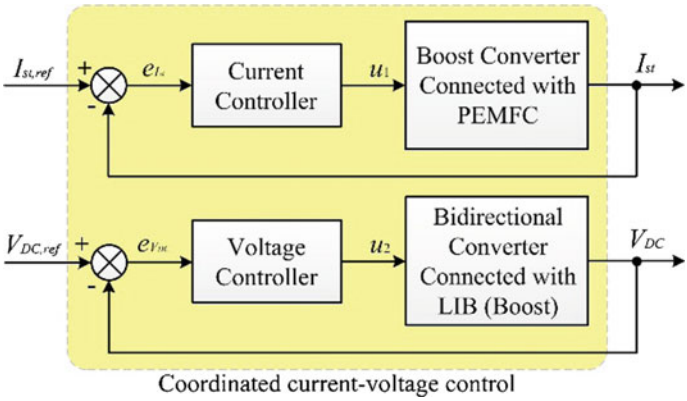


Fig. 2 Coordinated current-voltage control power splitting structure

4 Simulation and Experimental Validation

To validate the efficacy of the proposed power management strategy for the hybrid PEMFC/LIB power system through simulation and experiment, a downsized system configuration is applied. The parameters of the hybrid power system are given in Table 1.

First, a varied step load current profile (3.8 A → 5 A → 6.35 A → 3.8 A) was provided to test the coordinated current-voltage control power splitting method. The simulation results were shown in Fig. 4.

The stack current of PEMFC was tracking the reference value of 4.5 A. The DC bus voltage was also stabilizing at 24 V. Although the load was stepping up or down, PEMFC was working at the same point, and LIB was compensating the rest of the load power. The experimental validation results illustrated in Fig. 5, kept consistent with the simulation. The PEMFC constant voltage indicated that the

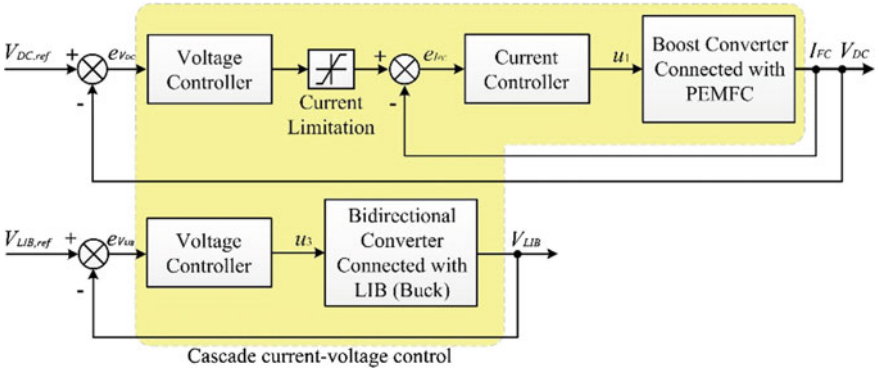


Fig. 3 Cascade current-voltage control power splitting structure

Table 1 Hybrid PEMFC/LIB power system parameters

Parameters	Description of parameters	Value
n	Number of cells	20
$P_{st,rate}$	Rate power of PEMFC	100 W
$V_{DC,ref}$	DC bus voltage reference	24 V
$I_{st,ref}$	PEMFC stack current reference	4.5 A
V_{nom}	Nominal voltage of LIB	14.8 V
Q_M	Maximum battery capacity	2400 mAh
L_1	DC-DC boost converter inductance	500 µH
C_1	DC-DC boost converter capacitance	200 µF
L_2	Bidirectional converter inductance	500 µH
C_2	Boost mode used capacitance for bidirectional converter	200 µF
C_3	Buck mode used capacitance for bidirectional converter	500 µF

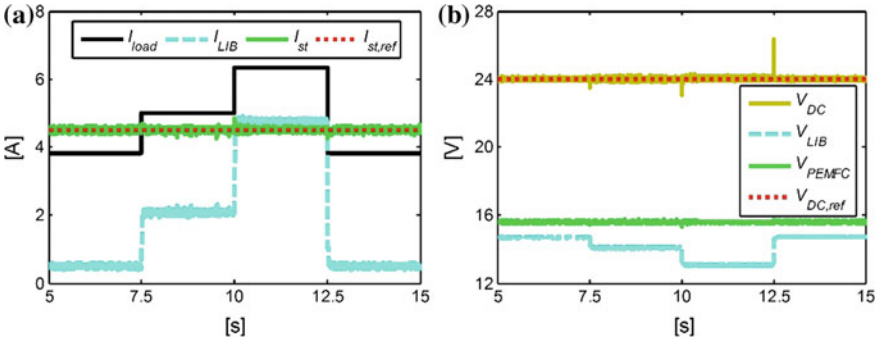
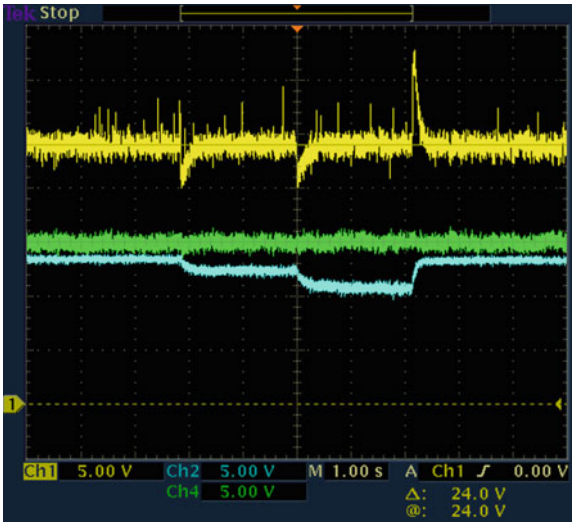


Fig. 4 Coordinated current-voltage control simulation: **a** current results; **b** voltage results

Fig. 5 Coordinated current-voltage control experimental validation



current of PEMFC was well tracking the current reference. Furthermore, the ECE-15 driving cycle was applied to the downsized hybrid PEMFC/LIB power system. According to the speed profile, the load required power was calculated based on the downsized system configuration. The simulation results were exhibited in Fig. 6b, where the PEMFC operated at a fixed point under the load required power varying conditions, and LIB provided the rest power at high load case, and LIB saved the extra power when the load was low or regenerative braking was happening. The driving cycle simulation results demonstrate that the proposed method is prospective for reducing FCV current fluctuation to extend PEMFC lifetime.

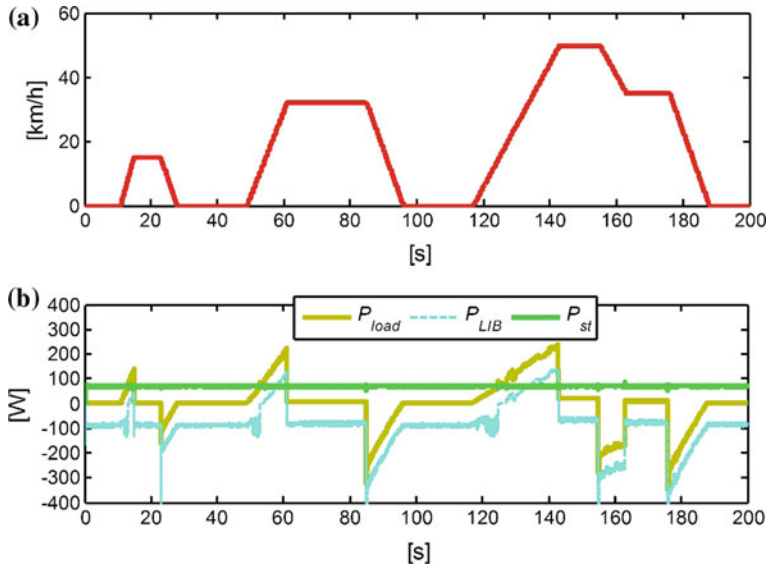


Fig. 6 ECE-15 driving cycle on a downsized hybrid PEMFC/LIB power system simulation results: **a** speed profile; **b** power distribution

5 Conclusion

In this study, a lifetime extending-oriented power management strategy of the hybrid power system is introduced. To design the power management system, the hybrid PEMFC/LIB power system model is firstly constructed in Matlab/Simulink. The hybrid system model can be divided into three parts: PEMFC, LIB, and DC-DC converters. Then, the power management strategy is developed by using coordinated current-voltage control and cascade current-voltage control to stabilize DC bus voltage and to regulate PEMFC stack current. Moreover, the simulation and experimental implementation are carried out to validate the feasibility of the proposed approach. The results indicated that the main power source was fixed at the operation point even though the load demand sudden changed. The fluctuation of the PEMFC can be ignored, which avoids the starvation and high current stress. The presented method provides a way to protect PEMFC and extend its lifetime.

Acknowledgements The authors would like to thank the support from Qishan Scholar Program in Fuzhou University (No. XRC-1643), and Fujian Provincial Collaborative Innovation Center for High-end Equipment Manufacturing (No. 0020-50006103).

References

1. N. Sulaiman, M.A. Hannan, A. Mohamed, E.H. Majlan, W.R.W. Daud, Renew. Sustain. Energy Rev. **52**, 802–814 (2015)
2. Q. Ning, D. Xuan, Y. Kim, Int. J. Automot. Technol. **11**(2), 229–238 (2010)
3. A. Tani, M.B. Camara, B. Dakyo, IEEE Trans. Veh. Technol. **61**(8), 3375–3386 (2012)
4. Y.-X. Wang, K. Ou, Y.-B. Kim, Int. J. Hydrogen Energy **40**, 11713–11721 (2015)
5. S.N. Motapon, L.-A. Dessaint, K. Al-Haddad, IEEE Trans. Ind. Electron. **61**, 1320–1334 (2014)
6. Y.-X. Wang, D.-H. Yu, S.-A. Chen, Y.-B. Kim, J. Power Sources **261**, 292–305 (2014)
7. W. Jiang, B. Fahimi, I.E.E.E. Trans, Ind. Electron. **57**(2), 752–761 (2010)
8. L. Wang, H. Li, I.E.E.E. Trans, Ind. Appl. **46**(3), 1011–1020 (2010)
9. C.H. Zheng, N.W. Kim, S.W. Cha, Int. J. Hydrogen Energy **37**(1), 655–663 (2012)
10. L. Xu, M. Ouyang, J. Li, F. Yang, L. Lu, J. Hua, Int. J. Hydrogen Energy **38**(24), 10104–10115 (2013)
11. J.P. Torreglosa, P. Garcia, L.M. Fernandez, F. Jurado, IEEE Trans. Ind. Inf. **10**(1), 276–285 (2014)
12. W. Greenwell, A. Vahidi, IEEE Trans. Ind. Electron. **57**(6), 1954–1963 (2010)
13. R.T. Amin, A.S. Bambang, C.J. Rohman, R. Dronkers, A. Ortega, Sasongko. IEEE Trans. Ind. Inf. **10**(4), 276–285 (2014)
14. A.A. Ferreira, J.A. Pomilio, G. Spiazzi, L.D.A. Silva, IEEE Trans. Power Electron. **23**(1), 107–115 (2008)
15. M.G. Simoes, B. Blunier, A. Miraoui, I.E.E.E. Ind, Appl. Mag. **20**(4), 41–49 (2014)
16. H. Hemi, J. Ghouili, A. Cheriti, Energy Convers. Manage. **80**, 63–70 (2014)
17. M. Zandi, A. Payman, J.-P. Martin, S. Pierfederici, B. Davat, F. Meibody-Tabar, IEEE Trans. Veh. Technol. **60**(2), 433–443 (2011)
18. X. Zhang, C.C. Mi, A. Masrur, D. Daniszewski, J. Power Sources **185**, 1533–1543 (2008)
19. O. Erdinc, B. Vural, M. Uzunoglu, J. Power Sources **194**, 369–380 (2009)
20. T. Fletcher, R. Thring, M. Watkinson, Int. J. Hydrogen Energy **41**(46), 21503–21515 (2016)
21. S.N. Motapon, O. Tremblay, L.-A. Dessaint, in *Proceedings IEEE Power & Energy Society General Meeting*, (Quebec, 2009), pp. 1–8
22. O. Tremblay, L.-A. Dessaint, A.-I. Dekkiche, in *Proceedings IEEE Vehicle Power & Propulsion Conference*, (Arlington, 2007), pp. 284–289

Vision-Based Measurement System for Structural Vibration Monitoring and Damage Detection

Jianfeng Zhong and Shuncong Zhong

Abstract A vision-based measurement system for vibration monitoring was proposed by using a non-projection fringe pattern. The designed artificial fringe pattern was similar to the interferogram of 2D-OCVT system, which was named as quasi-interferogram fringe pattern (QIFP) and pasted on the surface of a vibrating structure. A high-speed CMOS camera worked as a detector was used to capture the image sequence of the fringe pattern during the structural vibration. The period density of the imaged QIFP changes due to the structural vibration, from which the vibration information of the structure could be obtained. The change of the dynamics parameters of a cracked structure was analyzed by Finite Element Method (FEM), traditional accelerometer-based method and the proposed method using a roving auxiliary mass, from which the frequency shift curves can be obtained. The crack position information can be achieved confidently from the discontinuity of the frequency shift curves owing to the auxiliary mass effect when the mass was located at the crack position. The results demonstrated that the proposed method was an effective and accurate technique to measure structural vibration without introducing extra mass on the tested structure. Significant advantages of the proposed method making the measurement system suitable for vibration monitoring of engineering structures and damage detection of beam structures.

Keywords Vision-based · Vibration measurement · Damage detection · Auxiliary mass · Frequency shift curve

J. Zhong · S. Zhong (✉)
Laboratory of Optics, Terahertz and Non-Destructive Testing,
School of Mechanical Engineering and Automation, Fuzhou University,
Fuzhou 350108, People's Republic of China
e-mail: zhongshuncong@hotmail.com

S. Zhong
Shanghai University, Shanghai, P. R. China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_22

1 Introduction

One or more of the dynamic properties will change when a structure suffers from damages. Conversely, the damage could be identified using the changes in the structural dynamic response characteristics. Reviews on these issues have been presented by Salawu [1] and Jassim et al. [2]. The most useful damage detection and localization methods are probably those using changes in natural frequencies because frequency measurements can be conveniently conducted, less contaminated by experimental noise and often reliable. Zhong et al. [3–5] proposed an approach based on auxiliary mass spatial probing to provide a method for crack detection in beam-like structure. The key of the method is the precise positioning of the auxiliary mass and accurate identification of the natural frequencies of the test structures from the vibration signals.

Techniques available for vibration measurement can be generally classified as contact and noncontact types. Contact-type sensors usually have limitations in practical applications, such as the physically connection to the structure and extra mass loading on the structure. In recent years, there were many vision-based techniques for non-contact measurement of dynamic deformation or vibration in different industrial areas, in which various physical quantities need to be measured based on two-dimensional sensors like CCD or CMOS camera. Projected-fringe measurement method is a vision-based non-contact whole field optical methodology allowing measurement of surface contour or low-frequency vibration [6, 7]. Since it is based on several phase-shift images to determine the magnitude and phase of the vibrating object, it is suitable for very low-frequency vibration measurement. Motion detection systems [8, 9] based on a COMS image sensor (CIS) were introduced by many researchers, where all or part pixels of the successive images were used for comparison using a complex algorithm.

Recently, the authors proposed 1D- and 2D-optical coherence vibration tomography (OCVT) systems [10–12] offering the possibility of performing high resolution and non-intrusive vibration measurements. The vibration information was obtained from the interferogram that vary with the change of the distance between the surface of a reference mirror and the tested object in the 2D-OCVT system [11]. A non-projection QIFP vibration measurement method similar to the principle of OCVT was proposed in this study, which can realize vibration measurements and damage detection by using QIFP as a sensor and a high-speed CMOS image sensor (CIS) camera as a detector.

2 Theoretical Background

2.1 Principle of the Proposed Method

Figure 1a is an interferogram from the 2D-OCVT system, from which the displacement of the object at the sample arm can be obtained from the change of the fringe density. Taking the concept from the 2D-OCVT system, a quasi-interferogram fringe pattern (QIFP) that looks similar to the real interferogram had been designed, as shown in Fig. 1b. The artificial QIFP was printed by a normal printer. As can be seen from the photograph of experimental device shown in Fig. 1c, the printed QIFP was attached to the surface of the vibrating structure.

Figure 1b is the schematic layout of imaging principle for the QIFP. The high-speed camera was used to record the image sequence of the QIFP during the structure vibrating. Note that the QIFP was imaged in the middle of the sensor and only the fringe scope marked by the blue box was recorded to decrease the image size and save the memory. Subsequently, the video images of the QIFP were digitized in 8-bit grey scales and were streamed into the computer for further processing. The width of the imaged QIFP varies with the change of the object distance between the QIFP and the lens. Consequently, different vibration amplitudes bring about different period densities of the imaged QIFP. The fringe intensities in the middle line of each imaged QIFP were used for the calculation of

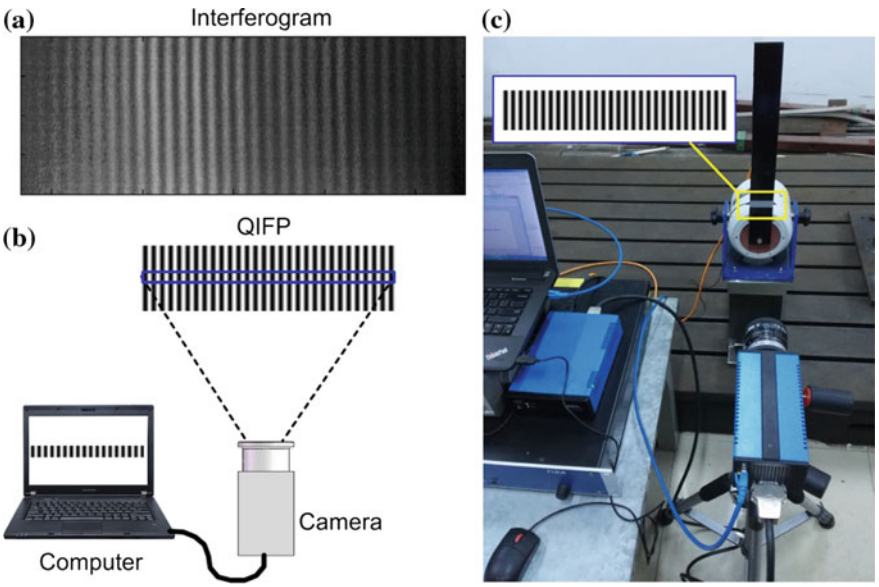


Fig. 1 **a** An interferogram from a 2D-OCVT system; **b** schematic diagram of the vision- and QIFP-based vibration measurement system; and **c** the photograph of the proposed system for dynamic measurement of a beam

the dynamic displacement by applying Fast Fourier Transform (FFT) to them. Subsequently, the real-time displacement can be obtained by combing the main peak of each QIFP-FFT waveform. It should be noted here that the calculation of the period density of the imaged QIFP was corrected using a spectral center correction method (SCCM) [3] to decrease the signal leakage effect of the Fast Fourier Transform (FFT).

According to the imaging principle shown in Fig. 1c, the displacement in time domain can be expressed as

$$\Delta Z(t) = f(1 + L/C_0)[d(t) - d_0]/d_0 \tag{1.1}$$

where f is the focal length of the lens, L is the physical length of the QIFP, C_0 is the length of imaged QIFP as reference that could be calculated from $C_0 = aN_{\mathrm{i}}$, a is the width of the pixel, whilst N_{i} is the pixels number covered by the QIFP. d_0 and $d(t)$ are the calculated fringe period density of the reference imaged QIFP and the QIFP at time t. Note that the detectable displacement should be smaller than the depth of field of the lens to prevent blurry images.

2.2 *Performance Evaluation of the Vision-Based Quasi-OCVT System*

In the vision-based Quasi-OCVT, the resolution of the camera CMOS image sensor (CIS) is 1024 × 1280 pixels and the pixel size is 12 × 12 μm. During the acquisition, the random distribution of dark current, temporal noise and fixed pattern noise of CIS are the major troublesome issues that will introduce noise on the obtained fringe intensity images and these noise sources will be affected by the illumination condition and the exposure time. The exposure time of the CIS will introduce different level of noise on the imaged fringe intensity, which has significant impact on the performance of the proposed system. In order to estimate the influence of the noise sources on the performance of the Quasi-OCVT and to characterize the measurement accuracy of the proposed system, some simulations were carried out. Normally, the noise of the CIS increases with the decrease of the exposure time. In the simulation, therefore, different levels of noise were added into the fringe intensities of the simulated QIFP series to simulate the noise under different exposure time. Two QIFPs with fringe period density of 10 period/cm (Fig. 2a) and 20 period/cm (Fig. 2b) were considered. The simulated movement was a 20 Hz sine vibration with amplitude of 50 μm. The pixel number for sampling the intensity was 800.

As shown in Fig. 2, the displacements can be correctly obtained if the intensity of the two QIFPs without noise. However, the displacements in Fig. 2a have more noise than that in Fig. 2b when the signal-to-noise ratios of the intensities of the QIFP were 20 and 30 dB. Therefore, the increase of the fringe period density of the QIFP can improve the measurement accuracy when the noise level of the CIS was

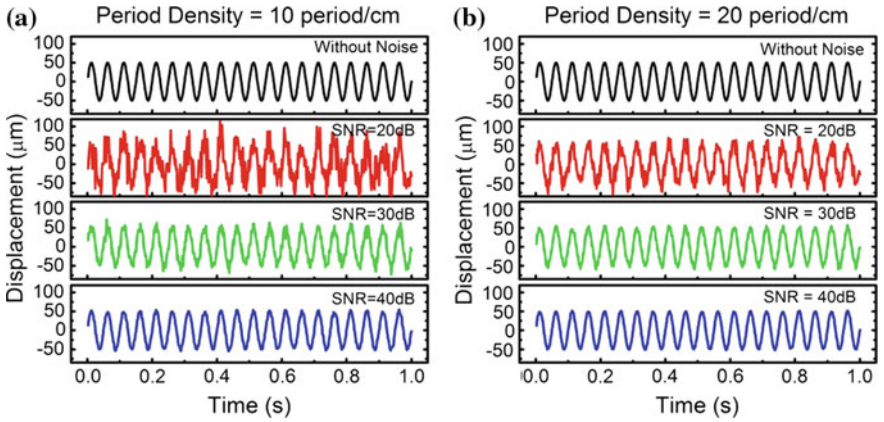


Fig. 2 The displacements obtained from two kinds of QIFP with different fringe period densities under different level of noise

high. That is, we can use a QIFP with higher fringe period density when the exposure time of the CIS is short or the illumination condition is bad for the performance improvement of the method. Because the increase of the fringe period density increased the signal-to-noise ratio of the imaged QIFP intensity on the condition that the noise energy is constant, which will improve the correction accuracy of the SCCM. However, the noise of the CIS decreased significantly if the exposure time was long enough, and the fringe period density had a decreasing effect on the measurement accuracy, which could be seen from the displacements in Fig. 2a, b when the SNR of the intensities were equal to 40 dB.

In addition, the resolution of the image sensor also influences the measurement accuracy. The measurement accuracy increases with the increasing of the sampling point of the QIFP. The sampling point has relationship with the object distance and the physical size of the pixel. Generally, the pixel size is constant for a camera. Therefore, there are two methods that can improve the measurement accuracy by increasing the sampling points for a QIFP with constant size and period density: one is shortening the object distance for more pixels involved to sample the QIFP; the other one is using the camera image sensor with smaller pixel size.

2.3 Damage Detection for a Cantilevered Beam

When the cantilevered beam is subjected to a crack, the dynamic behavior of the structure will be changed. This change could be amplified by an auxiliary concentrated mass. Therefore, in the damage detection experiment, an auxiliary mass was employed since it can influence the natural frequency of the beam structure. During the detection process, the auxiliary mass traversed from one end to the other

end with a constant spatial distance. At each position, the proposed vision-based vibration system was used to measure the dynamic signals to characterize each order of frequency of the cantilevered beam. Noted that the QIFP was pasted on the free end of the beam to avoid interfere the movement of the auxiliary mass. The frequency shift curves could be obtained by the combination of the same order of natural frequencies. The frequency shift curves will be smooth for an intact beam when the auxiliary mass moves along the length direction from one end to the other end with a fixed spatial interval. However, a discontinuous point or sharp variation point occurs in the FSC when the mass located close to the crack position [4] because the presence of a crack causes the local stiffness of the beam to decrease.

3 Experiments and Discussion

In the experiment, two kinds of cantilevered bean were employed to verify the performance of the vision-based Quasi-OCVT system. One is a plastic cantilevered beam excited by an exciter to verify the performance of the system in dynamic monitoring with a large amplitude vibration; the other one is a cracked steel cantilevered beam excited by a hammer to test the performance of the system in dynamic parameters characterization for damage detection and crack localization.

3.1 Vision-Based Quasi-OCVT Technique for Dynamic Monitoring of a Plastic Beam

In the experiment, a plastic beam-like structure and a QIFP with a period density of 20 period/cm were used. The QIFP was pasted on the surface of the beam at a distance of 190 mm from the fixed end. The cantilever beam (420 × 50 × 4 mm) was excited by the exciter. The excitation signal was generated by a function generator and was amplified by a power amplifier. Initially, the camera was placed perpendicular to the beam surface and the QIFP pasted on the beam surface was imaged in the middle of the CIS. The optical lens of the high-speed camera can be manually tuned for better imaging quality. Then the swept exciting signal with frequency from 0–400 Hz during a period of 8 s was inputted and the structure vibrated and deformed with the excitation. The camera sampling frequency was set to 1869 Hz. Figure 3 shows the obtained displacement of the plastic beam and its FFT result, from which we can know that the structural vibration amplitudes became larger when the frequency of the exciting signal close to the natural frequencies. The calculated first six natural frequencies are 5.25, 33.90, 95.73, 176.53, 198.92, and 318.61 Hz respectively, which can be obtained from the inset figure s5. For clarity, the displacement curves in the red boxes are shown in the inset figures named s1, s2, s3, and s4, respectively. The results obtained by the proposed

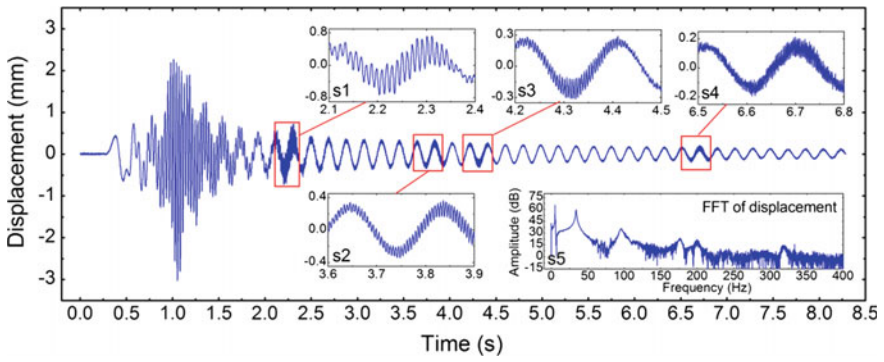


Fig. 3 The obtained displacement and its FFT of a plastic cantilever beam under a swept excitation

quasi-OCVT method demonstrated its high precision and stability for the dynamic monitoring for a vibration with amplitude up to 3 mm.

3.2 *Vision-Based Quasi-OCVT Technique for Damage Detection of a Steel Cantilevered Beam*

A cantilevered beam with a transverse crack was used to further illustrate the performance of the quasi-OCVT technique for damage detection. The size of the steel beam was 700 × 40 × 4 mm. One transverse saw cut with depth of 2 mm and width of 1.5 mm was located at 430 mm from the fixed end of the beam. Both the precise positioning of the auxiliary mass and the exact determination of the natural frequencies were the key points for the success of the frequency shift curve based damage detection method [3]. Grids were drawn on the surface of the cantilevered beam to accurately positioning the auxiliary mass. During the experiment, a magnet with weight of 80 g as the auxiliary mass was traversed along the beam from the fixed end to the free end. The spatial probing distance was 10 mm, which results in a total number of 69 measuring points.

Cantilevered beam model with same dimension, defect and auxiliary mass was also studied using Finite Element Method (FEM). In the simulation, the steel beam model have the following material properties: Young's modulus of elasticity $E = 2.1 \times 10^{11} \mathrm{N/m^2}$, mass-density $\rho = 7860 \mathrm{kg/m^3}$, Poisson ratio $\nu = 0.3$. Sixty-nine cases were simulated when the magnetic auxiliary mass located at different positions. After each simulation, the first five natural frequencies of bending mode shapes were extracted and plotted against the axial location of the auxiliary mass to generate the frequency shift curves. Traditional accelerometer-based vibration measurement system was also employed to obtain the frequency shift curves for further comparison.

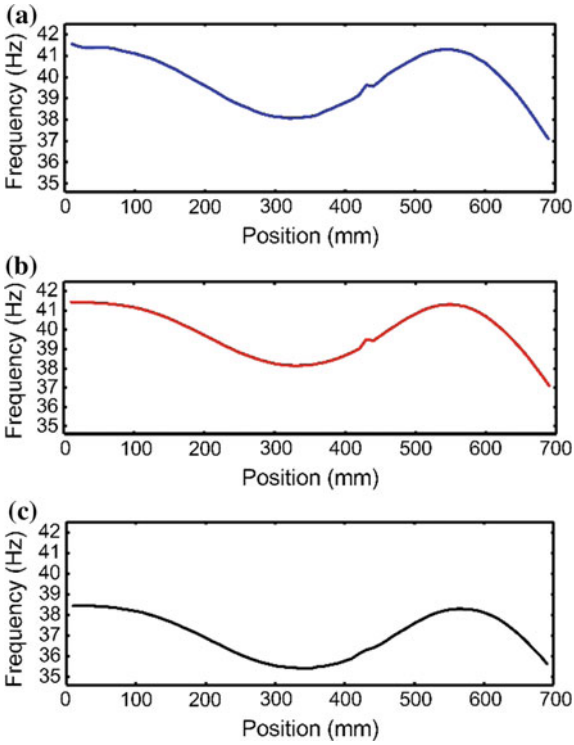


Fig. 4 The 2nd frequency shift curves obtained by Quasi-OCVT technique (**a**), simulation (**b**) accelerometer-based method (**c**)

The obtained experimental and simulated frequency shift curves are shown in Fig. 4. The blue solid line (a), the red solid line (b) and the black solid line (c) are corresponding to the result obtained by quasi-OCVT technique, simulation and traditional accelerometer-based method, respectively. From the blue and red solid lines, a step change or discontinuity in the curves occur when the auxiliary mass was exactly at the crack location. Hence, it can directly detect the crack in the beam from these curves. The good agreement of the Quasi-OCVT based result and numerical result illustrates the validity, efficiency and accuracy of the proposed method. However, the frequency shift curve extracted from the acceleration data is smaller than both the simulation and quasi-OCVT-based results owning to the mass effect of the accelerometer, as shown by the black solid line in Fig. 4c.

4 Conclusions

This paper proposes a vision-based vibration measurement approach using an artificial QIFP and a high-speed camera. Both the systemic factors that have influence on the measurement accuracy and the performance of the system were discussed and illustrated by simulation. The good agreement of the results between

the simulation and Quasi-OCVT laboratory tests verify the efficiency, practicability and high accuracy of the proposed Quasi-OCVT method. In addition, the proposed vision-based vibration measurement system is simple and non-contact in nature. Compared with the traditional contact methods, complicated setup is not needed and mass-effect is not introduced, making it attractive for the application in the characterization of structural modal parameters and damage detection of light structures.

Acknowledgements This work is supported by the Fujian Provincial Excellent Young Scientist Fund (2014J07007), the National Natural Science Foundation of China (51675103), the Training Program of Fujian Excellent Talents in Universities, the Specialized Research Fund for the Doctoral Program of Higher Education, the Ministry of Education, P. R. China (20133514110008) and the Ministry of Health, P.R. China (WKJ-FJ-27).

References

1. O.S. Salawu, Detection of structural damage through changes in frequency: a review. Eng. Struct. **19**(9), 718–723 (1997)
2. Z.A. Jassim et al., A review on the vibration analysis for a damage occurrence of a cantilever beam. Eng. Fail. Anal. **31**, 442–461 (2013)
3. S. Zhong, S.O. Oyadiji, K. Ding, Response-only method for damage detection of beam-like structures using high accuracy frequencies with auxiliary mass spatial probing. J. Sound Vib. **311**(3–5), 1075–1099 (2008)
4. S. Zhong, S.O. Oyadiji, Identification of cracks in beams with auxiliary mass spatial probing by stationary wavelet transform. J. Vib. Acoust. **130**(4), 041001 (2008)
5. S. Zhong, S.O. Oyadiji, Analytical predictions of natural frequencies of cracked simply supported beams with a stationary roving mass. J. Sound Vib. **311**(1–2), 328–352 (2008)
6. Y. Fu, Low-frequency vibration measurement by temporal analysis of projected fringe patterns. Opt. Lasers Eng. **48**(2), 226–234 (2010)
7. M.H. Ortiz, E.A. Patterson, Location and shape measurement using a portable fringe projection system. Exp. Mech. **45**(3), 197–204 (2005)
8. Y. Fukuda et al., Vision-based displacement sensor for monitoring dynamic response using robust object search algorithm. Sens. J., IEEE **13**(12), 4725–4732 (2013)
9. D. Ribeiro et al., Non-contact measurement of the dynamic displacement of railway bridges using an advanced video-based system. Eng. Struct. **75**, 164–180 (2014)
10. J. Zhong, S. Zhong, Q. Zhang, Two-dimensional optical coherence vibration tomography for low-frequency vibration measurement and response-only modal analysis. Mech. Syst. Signal Process. **79**, 65–71 (2016)
11. J. Zhong et al., Two-dimensional optical coherence tomography for real-time structural dynamical characterization. Opt. Lasers Eng. **66**, 74–79 (2015)
12. S. Zhong, H. Shen, Y. Shen, Real-time monitoring of structural vibration using spectral-domain optical coherence tomography. Opt. Lasers Eng. **49**(1), 127–131 (2011)

Application of the LM-HLP Neural Network to Automatic Smartphone Test System

Wei-Ting Hsu, Chia-Chi Lu and Jih-Gau Juang

Abstract Industry 4.0 has become an inexorable trend. To increase the efficiency of testing screen on smartphone, we design a system which includes a 5 degrees-of-freedom robotic arm to smartphone automatic test. The coordinate conversion is important for the operation between robot and camera. We use an algorithm named improved back propagation (BP) neural network on the robot arm to solve the coordinate conversion problem. First, the neural network is used to fix the error of coordinate difference between robot and camera. Then the improved BP algorithm will train the network through the error. For the machine vision, we use a video camera to catch the patterns on the screen of tested smartphone. The control scheme calculates angle of motor through image processing and fuzzy control. When robot arm cannot press buttons correctly, fuzzy system will fix the error. Experimental results show that the proposed control scheme is capable of drive the robotic arm by DH model to press the desired positions on the tested smartphone.

Keywords DH model · Pattern match · Improved BP neural network · Fuzzy control

1 Introduction

In recent years, intelligent robots become more and more common in our life, it provide service such as health care, entertainment, life safety etc. Also it contains many professional fields like path planning, visual image processing technology, body positioning, obstacle avoidance techniques, and arm control [1–3]. To make the human-robot interaction more harmonious, many researches are designed to

W.-T. Hsu · C.-C. Lu · J.-G. Juang (✉)
Department of Communications, Navigation and Control Engineering,
National Taiwan Ocean University, Keelung, Taiwan
e-mail: jgjuang@mail.ntou.edu.tw

C.-C. Lu
e-mail: luchingrose@gmail.com

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_23

convert expertise into the robot system. Hence, the main purpose of this paper is to control the robotic arm to replace hands-on smartphone testing process. The proposed method tends to test a smartphone by using a robotic arm as precisely as possible. With neural network and fuzzy system, complex mathematical formulas are not required [4] that can simplify computation in control process.

In this study, back propagation neural network is improved by the use of Le-venberg Marquardt-Hidden Layer Partition (LM-HLP). And the DH model which we utilized for coordinate conversion is able to drive robot arm to target positions [1]. Then image processing and recognition methods will identify the position of the target button. In the image processing, image will be converted to binary image through the HSL color space.

2 System Description

2.1 Hardware Description

In this paper, the embedded computer is used as the main controller. The Dynamixel MX-28 servo motor, which is produced by ROBOTIS Company, is the main motive force of the robot arm, as shown in Fig. 1. The motor's weight is 72 g. Dimension is 35.6 mm × 50.6 mm × 35.5 mm. Resolution is 0.088°. Running Degree is 0–360°. Gear Reduction Ratio is 193:1. Stall Torque is 24 kgf cm

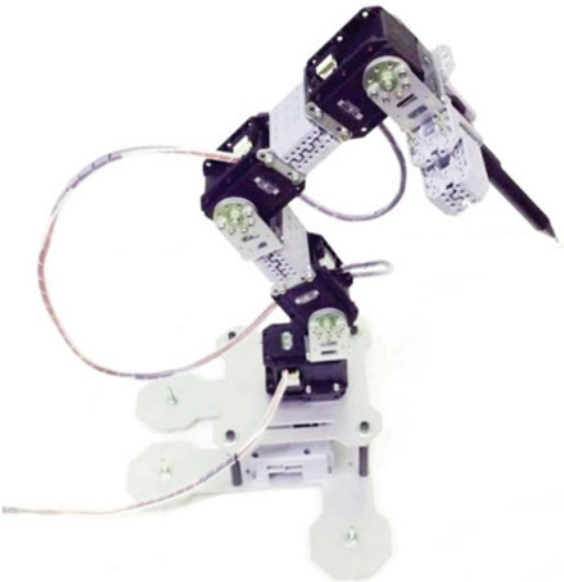


Fig. 1 Robot arm with Dynamixel MX-28 servo motor

(at 12 V, 1.5 A). No load speed is 54 rpm (at 12 V). In this paper, we use the network camera Microsoft LifeCam Studio 1080P Full-HD for image process.

3 Control Scheme

3.1 Pattern Match

Machine vision applications require a combination of hardware and software to ensure success. While choosing the right hardware is important, the actual visual inspection software forms the core of any machine vision system. We acquire image by camera, and configure and benchmark a sequence of visual inspection steps as well as deploy the visual inspection system for automated inspection. For the regions of inspection in relation to the number keys, we need to set a coordinate system that relatives to a significant and original feature of the number keys. A feature is chosen such that it is always in the camera field of view despite the different locations in which the number keys may appear from image to image. Also, we need to make sure the feature is not affected by major defects that could drastically modify the visual appearance of the feature. To configure the Match Pattern step, as shown in Fig. 2, draw a region of interest around the area of the image containing the center point. This region becomes the pattern matching template. We look for and locate the specified template inside the green region of interest in the image, as shown in Fig. 3. The location of the match is overlaid on the inspection image with a red rectangle.

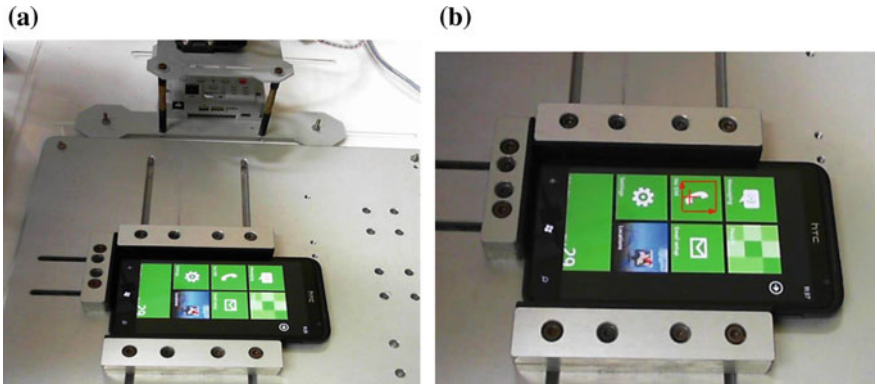


Fig. 2 **a** Match Pattern, **b** get coordinate of target

Fig. 3 Pattern of target

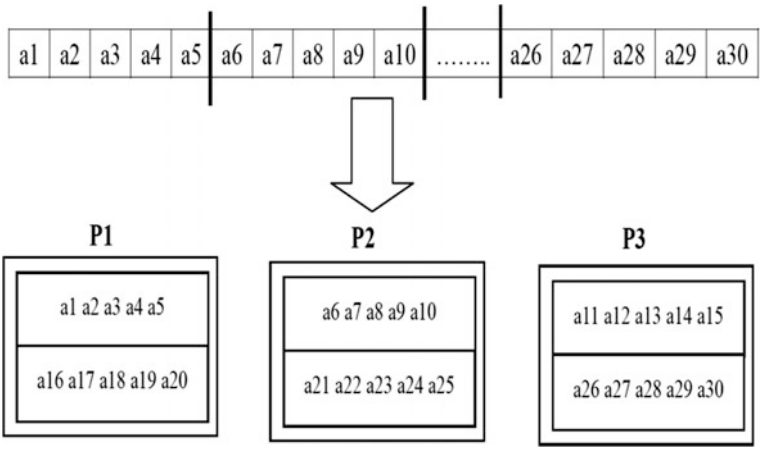


Fig. 4 Block-cyclic distribution

3.2 *Improved BP Neural Network*

There are few disadvantages of BP neural network. For example, it usually falls to local optimum and costs more time in training. An improved BP neural network with Levenberg Marquardt-Hidden Layer Partition (LM-HLP) [5] is used to solve the problem of coordinate conversion between webcam and robotic arm. The LM-HLP applies Levenberg Marquardt to replace steepest descent method of BP in (1), the Block-Cyclic distribution, shown in Fig. 4, is equally distributed huge data to each processor. The LM-HLP scheme is shown in Fig. 5, and flowchart of LM-HLP is shown in Fig. 6.

$$\Delta \mathrm{w} = -[J^T(w_k)\mathrm{J}(w_k) + \mu(k)]^{-1} \times J^T(w_k) \times e_{k,q} \tag{1}$$

where $J = s_q^{m+1} \times a^{m-1}$ is Jacobian Matrix, s is sensitivity, a is output, μ is learning rate, w is weight, and e is error. In Figs. 4 and 5, "a" is data to be identified, "p" is group number.

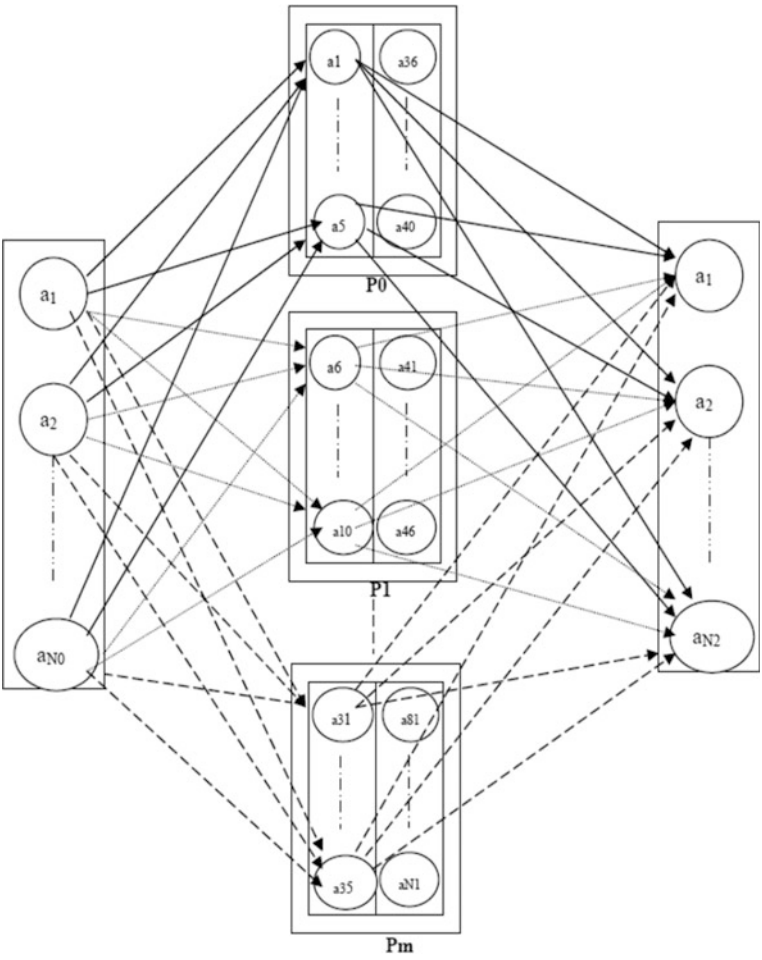


Fig. 5 LM-HLP scheme

3.3 Fuzzy Control

Hierarchy fuzzy controller and traditional fuzzy controller have different input and output architectures. The difference affects the number of fuzzy rules based structure. Traditional fuzzy control rules grow exponentially with the increase of the input variable number. Using hierarchical fuzzy theory to design the controller can improve this disadvantage and effectively reduce the establishment of fuzzy rules. Figure 7 shows the hierarchical fuzzy structure [6]. The fuzzy inputs are coordinates from webcam, outputs are degrees of robot arm, and feedbacks are the distance with standard point, these are shown in Fig. 8.

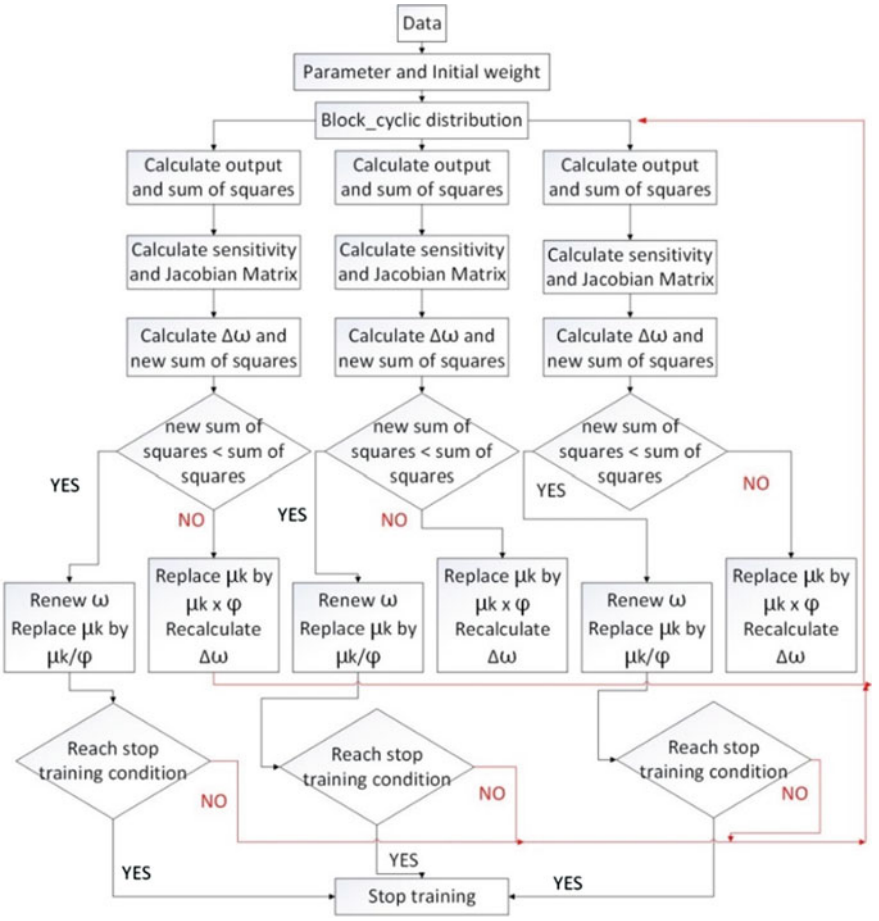


Fig. 6 LM-HLP processing flowchart

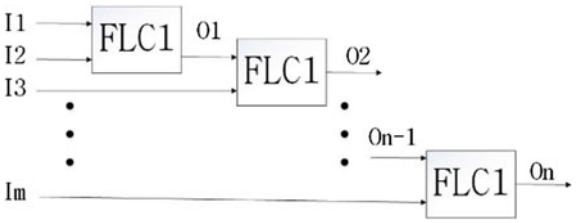


Fig. 7 Hierarchical fuzzy structure

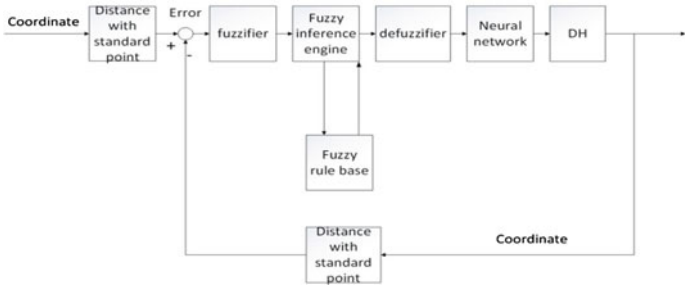


Fig. 8 Fuzzy control scheme

Fig. 9 The relative position of camera and robot arm

4 Experiment Results

Figure 9 shows the relative position of camera and robot arm. Results of position control are shown in Fig. 10. Experimental results show that the proposed control scheme can drive the robot arm to press the desired button of the tested smartphone successfully.

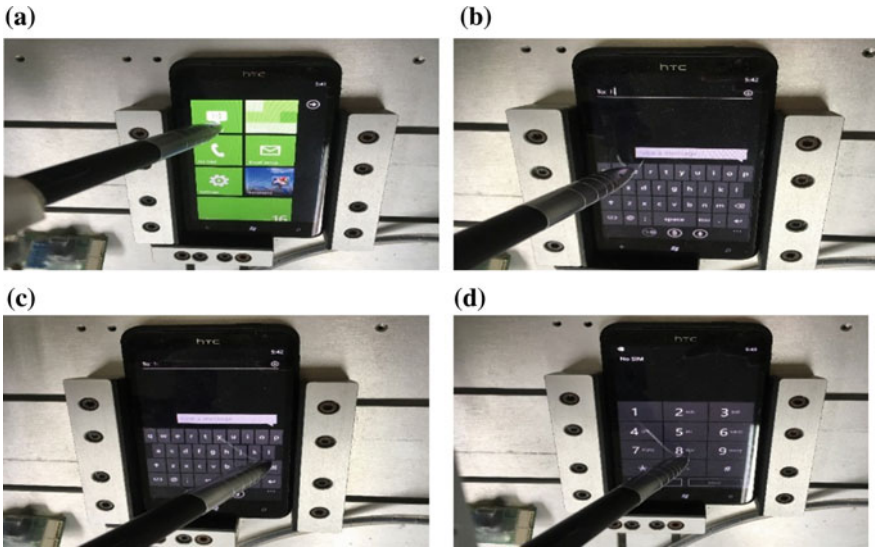


Fig. 10 **a** Move the arm to touch the desired button "message", **b** move the arm to touch the desired button "a", **c** move the arm to touch the desired button "Delete", **d** move the arm to touch the desired button "8"

5 Conclusions

In this study, intelligent system is applied to control a robot arm and drive the robot arm to press the desired buttons of the tested smartphone. An improved BP neural network is utilized in this study to identify pattern locations. The robot arm performs position control in real time by fuzzy control. With pattern matching process, the system is able to identify image correctly. Experimental results show that the proposed intelligent scheme can control the robot arm to press the desired buttons of the tested smartphone.

References

1. Y.H. Weng, Automatic robot assembly with eye-in-hand binocular visual servoing and structured lighting, Master Thesis, Department of Electrical Engineering, National Taipei University of Technology, 2011
2. C.C. Liu, Commanded control of a robotic manipulator employing kinect and binocular vision sensors, Master Thesis, Department of Electrical Engineering, National Taipei University of Technology, 2011
3. C.L. Yu, Application of real-time image recognition and path planning to wheeled mobile robot for taking elevator, Master Thesis, Communications, Navigation and Control Engineering, Nation Taiwan Ocean University, 2012

4. T. Wei, Design and realization of autonomous mobile robot tracking system by multi-switching fuzzy sliding mode control, Master Thesis, Department of Electrical Engineering, National Cheng Kung University, 1998
5. S.F. Liang, *A Parallel Back-Propagation Algorithm with the Levenberg Marquardt Method* (Dept. of Computer Science & Information Management, Soochow University, 2006)
6. J.S. Huang, Servo motion control of robot arm, Master Thesis, Department of Power Vehicle and System Engineering, National Defense University, 2012

Two-Vectors-Based Model-Free Predictive Current Control of a Voltage Source Inverter

Cheng-Kai Lin, Jen-te Yu, Yen-Shin Lai, Hsing-Cheng Yu, Jyun-Wei Hu and Dong-Yue Wu

Abstract This paper presents a two-vectors-based model-free predictive current control (TVB-MFPCC) for a voltage source inverter (VSI). In a single-vector-based model predictive current control (SVB-MPCC), the VSI can generate one out of the eight voltage vectors during each fixed sampling period. To increase the number of the available voltage vectors up to fourteen, two voltage vectors are synthesized as new ones for current prediction. Moreover, unlike the SVB-MPCC that requires the system model and load parameters, the proposed TVB-MFPCC is based on measuring the load currents, such that any system model or system parameter is not required. Both the TVB-MFPCC and the SVB-MPCC are realized on a 16-bit microcontroller, TMS320F28335, and they are applied to a VSI for testing and comparing their current-tracking performances.

1 Introduction

A single-vector-based model predictive current control (SVB-MPCC) was successfully applied to a voltage source inverter (VSI) in 2007 [1]. In each sampling period of the SVB-MPCC, the used VSI can generate six active-voltage vectors

C.-K. Lin (✉) · J.-W. Hu · D.-Y. Wu
Department of Electrical Engineering, National Taiwan Ocean University, Keelung, Taiwan, Republic of China
e-mail: cklin@mail.ntou.edu.tw

J. Yu
Department of Electrical Engineering, National Taiwan University, Taipei Taiwan, Republic of China

Y.-S. Lai
Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan, Republic of China

H.-C. Yu
Department of Systems Engineering Naval Architecture, National Taiwan Ocean University, Keelung, Taiwan, Republic of China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_24

(namely V_1–V_6) and two zero-voltage vectors (namely V_0 and V_7) as shown in Fig. 2a. During each sampling period, the SVB-MPCC chooses only one voltage vector to control the VSI directly. However, it is difficult to meet the requirement of low current ripple with this method as it is limited by the number of the available voltage vectors. Moreover, another drawback is that the SVB-MPCC needs to know the system model and the load parameters (resistance and inductance) in order to predict the future load currents. To solve the above drawbacks permanently, a two-vectors-based model-free predictive current control (TVB-MFPCC) is proposed here, which is expected to have a better current-tracking performance. A microcontroller, TMS320F28335, is used to carry out both SVB-MPCC and TVB-MFPCC for testing their performances.

Two-Vector-Based Model-Free Predictive Current Control: Fig. 1 shows the schematic of a voltage source inverter connected with a three-phase load, where G_1–G_6 represent the six power switches in order, R refers to the resistance, and L denotes the inductance. In the SVB-MPCC [1], only eight voltage vectors (denoted as V_0–V_7) are available for current predictions as shown in Fig. 2a, where T_s is the fixed sampling period and the relationships between the voltage vectors and the six switches are listed in Table 1. As observed in Figs. 2a and 3a, the SVB-MPCC can apply only one voltage vector/switching mode during one sampling period T_s, which may significantly reduce the accuracy of current prediction. In order to increase the number of available voltage vectors up to fourteen as shown

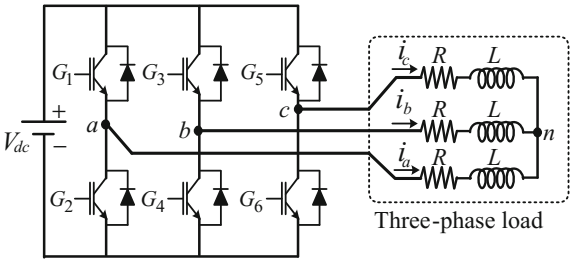


Fig. 1 Configuration of a voltage source inverter

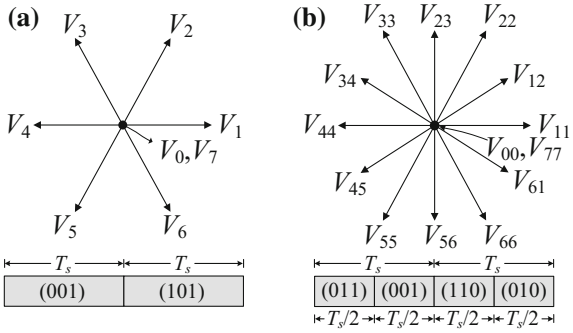


Fig. 2 Voltage vectors and switching status during sampling period T_s
a SVB-MPCC [1]
b TVB-MPCC

Table 1 Eight voltage vectors of the SVB-MPCC [1]

	Voltage vectors with switching functions applied throughout the sampling period T_s							
	V_0 (000)	V_1 (001)	V_2 (101)	V_3 (100)	V_4 (110)	V_5 (010)	V_6 (011)	V_7 (111)
Switches on	G_2, G_4, G_6	G_2, G_4, G_5	G_1, G_4, G_5	G_1, G_4, G_6	G_1, G_3, G_6	G_2, G_3, G_6	G_2, G_3, G_5	G_1, G_3, G_5
Switches off	G_1, G_3, G_5	G_1, G_3, G_6	G_2, G_3, G_6	G_2, G_3, G_5	G_2, G_4, G_5	G_1, G_4, G_5	G_1, G_4, G_6	G_2, G_4, G_6

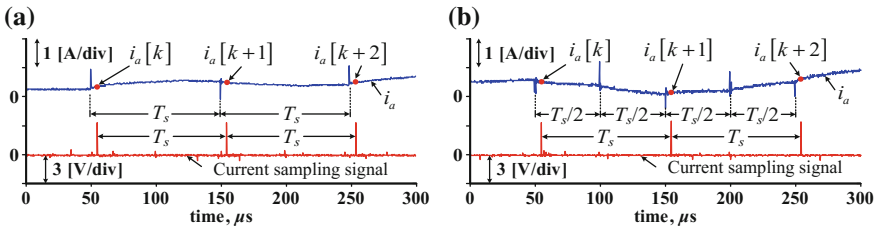


Fig. 3 Measured a-phase current and sampling signal **a** SVB-MPCC [1] **b** TVB-MPCC

in Fig. 2b, a TVB-MFPCC based on synthesized voltage vectors is proposed here. In this study, a synthesized voltage vector is composed of two voltage vectors. The combination of these two voltage vectors create fourteen switching modes (denoted as M_0–M_{13}) and are listed in Table 2. Since modern microcontrollers are sufficiently powerful, it is possible to apply two voltage vectors instead of one [1–3] in each sampling period as shown in Fig. 3b. The symbols used in Fig. 3b will be explained later.

In practice, the three-phase currents shown in Fig. 1 can be precisely measured by current sensors, and the current differences between the two measured currents can be calculated and stored by a microcontroller. To appropriately illustrate the principle of a TVB-MFPCC, the *a*-phase current differences can be calculated and classified according to the fourteen switching modes as follows:

$$\Delta i_{a,\,stored}\big|_{M_j} = i_a\left[n_j+1\right] - i_a\left[n_j\right], \qquad \forall n_j \le k-1, \quad j \in \{0, \ldots, 13\} \tag{1}$$

where the subscript *stored* refers to the previous current differences, n_j and k are positive integers, and $i_a\left[n_j+1\right]$ and $i_a\left[n_j\right]$ specify the currents that were measured at the $(n_j + 1)$th and the (n_j)th sampling instants, respectively.

Next, using the add-and-subtract technique, the $(k + 2)$th sampled *a*-phase current is derived as

Table 2 Fourteen switching modes of the proposed TVB-MFPCC

Switching mode	Applying time		Synthesized voltage vector	Output Voltages		
	$T_s/2$	$T_s/2$		V_{an}	V_{bn}	V_{cn}
M_0	(000)	(000)	$V_{00} = V_0/2 + V_0/2$	0	0	0
M_1	(001)	(001)	$V_{11} = V_1/2 + V_1/2$	$-V_{dc}/3$	$-V_{dc}/3$	$2V_{dc}/3$
M_2	(001)	(101)	$V_{12} = V_1/2 + V_2/2$	0	$-V_{dc}/2$	$V_{dc}/2$
M_3	(101)	(101)	$V_{22} = V_2/2 + V_2/2$	$V_{dc}/3$	$-2V_{dc}/3$	$V_{dc}/3$
M_4	(101)	(100)	$V_{23} = V_2/2 + V_3/2$	$V_{dc}/2$	$-V_{dc}/2$	0
M_5	(100)	(100)	$V_{33} = V_3/2 + V_3/2$	$2V_{dc}/3$	$-V_{dc}/3$	$-V_{dc}/3$
M_6	(100)	(110)	$V_{34} = V_3/2 + V_4/2$	$V_{dc}/2$	0	$-V_{dc}/2$
M_7	(110)	(110)	$V_{44} = V_4/2 + V_4/2$	$\boldsymbol{V}_{dc}/3$	$\boldsymbol{V}_{dc}/3$	$-2\boldsymbol{V}_{dc}/3$
M_8	(110)	(010)	$V_{45} = V_4/2 + \mathrm{F}_5/2$	0	$\boldsymbol{V}_{dc}/2$	$-\boldsymbol{V}_{dc}/2$
M_9	(010)	(010)	$V_{55} = V_5/2 + V_5/2$	$-\boldsymbol{V}_{dc}/3$	$2\boldsymbol{V}_{dc}/3$	$-\boldsymbol{V}_{dc}/3$
M_{10}	(010)	(011)	$V_{56} = V_5/2 + V_6/2$	$-\boldsymbol{V}_{dc}/2$	$\boldsymbol{V}_{dc}/2$	0
M_{11}	(011)	(011)	$V_{66} = V_6/2 + V_6/2$	$-2V_{dc}/3$	$\boldsymbol{V}_{dc}/3$	$V_{dc}/3$
M_{12}	(011)	(001)	$V_{61} = V_6/2 + V_1/2$	$-\boldsymbol{V}_{dc}/2$	0	$V_{dc}/2$
M_{13}	(111)	(111)	$V_{77} = V_7/2 + V_7/2$	0	0	0

$$\begin{aligned} i_a[k+2] &= i_a[k] + \big(i_a[k+1] - i_a[k]\big) + \big(i_a[k+2] - i_a[k+1]\big) \\ &= i_a[k] + \Delta i_a|_{M_k} + \Delta i_a|_{M_{k+1}}, \qquad M_k,\, M_{k+1} \in \{M_0,\, \ldots,\, M_{13}\} \end{aligned} \tag{2}$$

where $\Delta i_a|_{M_k} = i_a[k+1] - i_a[k]$; $\Delta i_a|_{M_{k+1}} = i_a[k+2] - i_a[k+1]$; and $i_a[k]$, $i_a[k+1]$, and $i_a[k+2]$ specify the currents measured at the (k)th, (k + 1)th, and (k + 2)th sampling instants, respectively, as shown in Fig. 3b. During the (k)th sampling period, Eq. (1) can be calculated because $n_j \le k-1$; however, Eq. (2) cannot be calculated because of the unknown currents $i_a[k+1]$ and $i_a[k+2]$ appearing in $\Delta i_a|_{M_k}$ and $\Delta i_a|_{M_{k+1}}$, respectively. A solution to this problem is proposed here. According to Eq. (1), the stored current differences, $\Delta i_{a,\,stored}\big|_{M_k}$ and $\Delta i_{a,\,stored}\big|_{M_{k+1}}$, are used to predict the values of $\Delta i_a|_{M_k}$ and $\Delta i_a|_{M_{k+1}}$, respectively. Further, the fourteen predicted values from (2) according to the fourteen switching modes can be calculated in the (k)th sampling period by

$$i_a^P[k+2]_{M_0,\ldots,M_{13}} = i_a[k] + \Delta i_{a,\,\mathrm{stored}}\big|_{M_k} + \Delta i_{a,\,\mathrm{stored}}\big|_{M_0,\ldots,M_{13}} \tag{3}$$

where the superscript P stands for prediction. Prediction errors are unavoidable, but they can be effectively reduced by setting the sampling period appropriately. The reason is that the values of Eq. (1) are updated and stored in each sampling period, signifying that the shorter the sampling period is made, smaller are the prediction errors obtained. Following the above descriptions, the subscript a used in (3) can be replaced by b and c in order to obtain the b-phase and c-phase current predictions,

denoted as $i_b^P[k+2]$ and $i_c^P[k+2]$, respectively. In order to quantify the errors between the current commands and the predicted currents, a cost function under different switching modes is defined as

$$g[k]_{M_0,\ldots,M_{13}} = \left|i_a^*[k] - i_a^P[k+2]_{M_0,\ldots,M_{13}}\right| + \left|i_b^*[k] - i_b^P[k+2]_{M_0,\ldots,M_{13}}\right| + \left|i_c^*[k] - i_c^P[k+2]_{M_0,\ldots,M_{13}}\right| \tag{4}$$

where $i_a^*[k]$, $i_b^*[k]$, and $i_c^*[k]$ are the (k)th current commands represented in a-phase, b-phase, and c-phase, respectively. Based on (4), a switching mode, denoted by $M_{\min}$, can be determined as follows:

$$g[k]_{M_{\min}\in\{M_0,\ldots,M_{13}\}} = \min\left\{g[k]_{M_0}, g[k]_{M_1}, \ldots, g[k]_{M_{13}}\right\}. \tag{5}$$

The switching mode $M_{\min}$ selected from (5) is applied to control the six power switches (G_1–G_6) directly in the next sampling period. Thus, the design of TVB-MFPCC is completed.

Experimental results: Both the SVB-MPCC [1] and the proposed TVB-MFPCC were tested under the same experimental circuits and RL load as shown in Fig. 4. There are nine parts in Fig. 4a. Part A is a SCM1246MF inverter power module, providing 600 V for the insulated gate bipolar transistor (IGBT) breakdown voltage and 30 A for the maximum continuous output current. Part B includes two LV-25P voltage sensors and a 16-bit four-channel A/D converter, AD7655. Part C includes three LA-25NP current sensors. Part D consists of the two power supplies. Parts E, F, and G are an overcurrent protection circuit, an optical isolation circuit, and an encoder circuit, respectively. Part H is the microcontroller board equipped with a TMS320F28335 chip manufactured by Texas Instruments. Part I is a XDS100V3 USB emulator connecting Part H and a computer. Figure 4b shows the three-phase RL load, consisting of three resistances of 2 Ω each and three inductances of 30 mH each.

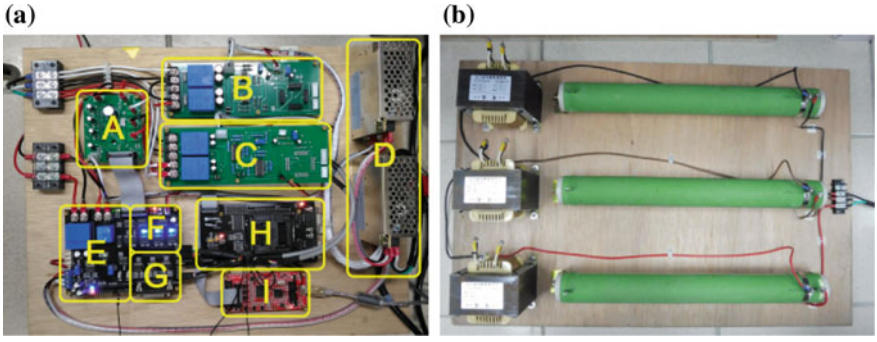


Fig. 4 Implemented circuit of the voltage source inverter and RL load **a** Hardware circuit **b** Three-phase RL load

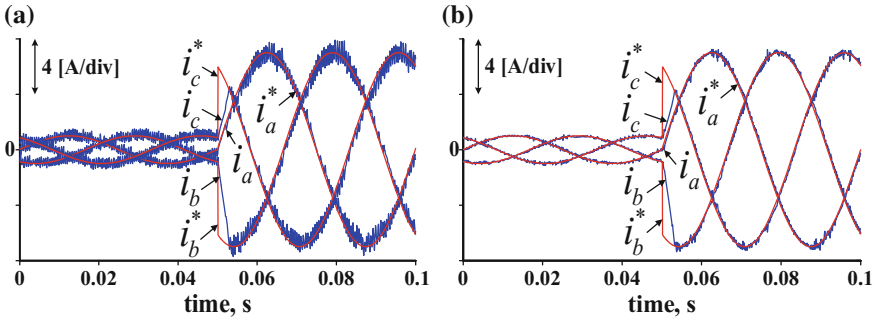


Fig. 5 Experimental results showing current responses with amplitudes of sinusoidal current commands changing from 1 A to 7 A at 0.05 s **a** SVB-MPCC [1] **b** TVB-MFPCC

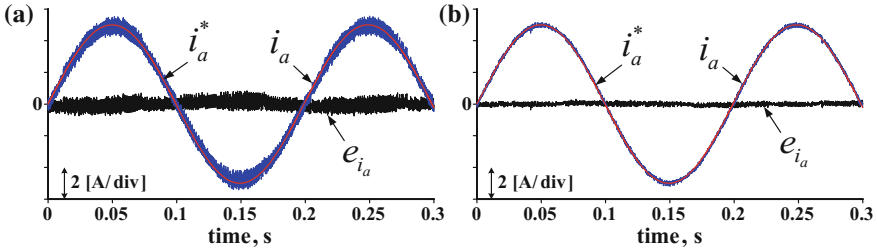


Fig. 6 Experimental results showing steady-state current response under a sinusoidal current command having an amplitude of 5 A **a** SVB-MPCC [1] **b** TVB-MFPCC

Two sets of experimental results are presented in Figs. 5 and 6. As seen in Fig. 5b, the proposed TVB-MFPCC has considerably better performance with smaller current errors as compared to that of the SVB-MPCC, shown in Fig. 5a. Figure 6 shows the steady-state current responses of the two methods, wherein the current error is defined as $e_{i_a} = i_a^* - i_a$. The current ripples generated by the TVB-MFPCC as shown in Fig. 6b are significantly reduced with respect to those in Fig. 6a, achieving a satisfactory steady-state response. In summary, the experimental results display that the current-tracking performance can be effectively improved by the proposed TVB-MFPCC.

2 Conclusions

This paper presents the TVB-MFPCC for a VSI connected with a three-phase load and reports the experimental results to test and verify its feasibility and correctness. As compared to the SVB-MPCC [1], the available voltage vectors of the TVB-MFPCC are increased up to fourteen from eight, implying that the accuracy of

the current prediction of the TVB-MFPCC is expected to be significantly higher. Using the TVB-MFPCC, the next switching mode is selected from fourteen ones through the process of finding the minimum cost function. Besides, in terms of the realization, the proposed method only performs one load current detection in each sampling period for reducing the implementation difficulty. Moreover, the TVB-MFPCC is based on the load current detection, implying that it is less affected by the variations of load parameters. In addition, the TVB-MFPCC retains all benefits of the SVB-MPCC, such as flexible current control design and the absence of pulse width modulation and a linear regulator. The experimental results show that the TVB-MFPCC provides better current-tracking performance than the SVB-MPCC.

References

1. J. Rodriguez, J. Pontt, C.A. Silva et al., Predictive current control of a voltage source inverter. IEEE Trans. Ind. Electron. **54**(1), 495–503 (2007)
2. C.K. Lin, T.H. Liu, J. Yu et al., Model-free predictive current control for interior permanent magnet synchronous motor drives based on current difference detection technique. IEEE Trans. Ind. Electron. **61**(2), 667–681 (2014)
3. C.K. Lin, J. Yu, Y.S. Lai et al., Simplified model-free predictive current control for interior permanent magnet synchronous motors. IET Electron. Lett. **52**(1), 49–50 (2016)

Net Power Output of Open Cathode Fuel Cell Optimization Based on the Over-Temperature Protection

Kai Ou, Ya-Xiong Wang, Wei-Wei Yuan and Young-Bae Kim

Abstract The axial flow fan is the only actuator that simultaneously responses for delivering oxygen and cooling system. In this study, oxygen mass flow rate flowing across cathode channels is regarded as the control variable of the open-cathode fuel cell system. Air flow regulation based on a set of optimal oxygen excess ratios was analyzed in different stack currents. The motor voltage of the axial fan is controlled by the PWM circuit to regulate the air flow speed of the air supply subsystem. The destination of the control strategy centralizes on power optimal and over-temperature protection at the same time while the fuel cell system is running. Based on the fuel cell system model, an integral time absolute error (ITAE) criterion is employed to determine the PI gains of the controller. The verified experiments were implemented on the experiment platform consisted of NI devices.

Keywords Open cathode PEM fuel cell · Control-oriented model · Oxygen excess ratio · Power optimal · Overtemperature protection

1 Introduction

The proton exchange membrane fuel cell (PEMFC) has been considered as the most promising energy supply. It has a low to medium electric power source and is the substitute for a conventional battery system because of its high efficiency, high power density, low process temperature, noiseless operation, and zero-emission characteristics. The PEMFCs also have great potential to become an alternative means for both automotive application and stationary power sectors. PEMFC

K. Ou · W.-W. Yuan · Y.-B. Kim (✉)
Department of Mechanical Engineering, Chonnam National University, Gwangju, Republic of Korea
e-mail: ybkim@chonnam.ac.kr

Y.-X. Wang
School of Mechanical Engineering and Automation, Fuzhou University, 350116 Fuzhou, China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_25

attracts more and more scientists' research interests and researchers' attention because of the abovementioned reasons [1]. With the recent progress of fuel cell technologies, mainly two types of PEM fuel cell are utilized, as follows: closed-cathode and open-cathode PEMFC. An integral fuel cell system comprises a cathode subsystem for oxygen supply and an anode subsystem for hydrogen supply. As introduced by Pukrushpan et al. [2], the air/oxygen supply system of typical closed-cathode PEMFC uses a compressor pressurizing ambient air in high pressure to cathode and uses a humidifier to moisturize and cool the air/oxygen after the pressurization process. Moreover, to satisfy the requirements for fuel cell calibration of a power system, such as optimal temperature, humidity, and oxygen excess ratio (λ_{o_2}), air flow delivered to fuel cell needs to vary with the change of working state. As for closed-cathode PEMFC system, a number of studies have shown improvement with various control strategies of the auxiliary system for optimizing the fuel cell performance. For example, Gruber et al. [3], Wang and Kim [4] presented model predictive control for oxygen excess ratio with the PEMFC balance-of-plant (BOP) model. Effective thermal management of fuel cell system is also a popular topic in fuel cell system [5, 6]. Hu et al. [5] proposed a fuzzy logic control for the fuel cell temperature to ensure system safety and high efficiency operation. Ahn et al. [7] has described a coolant control that would help improve reliability of the fuel cell system. The performance of a fuel cell system is also highly related to water content in the membrane. To avoid drying membrane or flooding the electrode, Riascos has studied humidity control [8]. However, the majority of studies concentrate on closed-cathode fuel cell system.

The main characteristic of an open-cathode PEMFC is that the cathode channels are exposed to the surroundings and only an axial fan serves as the auxiliary system that forces the ambient air into the cathode channels (shown in Fig. 1). Its structure is very simple. Thus, this fuel cell system is comparatively cheaper than the closed-cathode type PEMFC. However, it requires a more delicate control scheme to operate the PMEFC because only a fan is utilized for regulating pressure, temperature, humidity, and air excess ratio. Some researchers have studied advanced control applications to operate open-cathode-type PEMFC properly. For example,

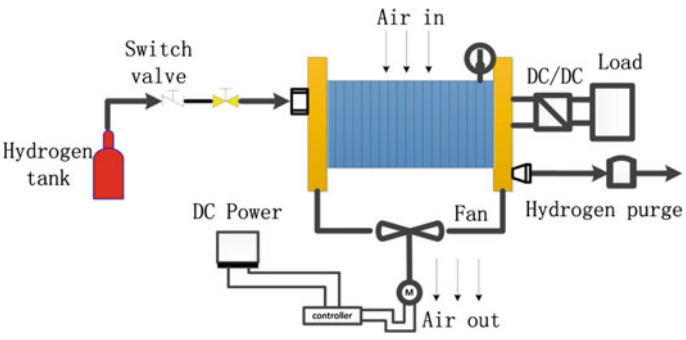


Fig. 1 Open cathode fuel cell system structure

optimal thermal management and analysis have been discussed previously [9, 10]. An experiment has been conducted to study the pressure drop in an open-cathode PEMFC system by Barreras et al. [11, 12]. They showed that air stoichiometry rate calculation of the open-cathode fuel cell system is much more difficult than the oxygen excess ratio of closed-cathode fuel cell system because laminar and turbulence flows coexist simultaneously in the cathode channels. Therefore, the thermal, humidity, and oxygen excess ratio management in open-cathode PEMFC system is difficult because the integrated system supply of the air flow adversely affects fuel cell system performance [13].

The air flow regulation is necessary to deliver air to cathode as soon as possible, especially when the external load has suddenly changed. If the delivery of the air is insufficient, the system would be caught in a starvation phenomenon ($\lambda_{o_2} < 1$), which possibly causes permanent membrane degradation [13, 14]. For closed-cathode fuel cell systems, several studies have presented the methods to ensure that the fuel cell system works in plentiful air/oxygen supply state [2–4]. Although many researchers regard the thermal management as a principal issue to safeguard the system for integration of the open-cathode fuel cell [8–11], the oxygen excess ratio is also critical for fuel cell system operation, especially when the fuel cell systems work in the initial phase. Most suppliers of the open-cathode PEMFC used simple operating rules. For example, before the open-cathode fuel cell system reached an optimal temperature point, the actuator of the fan almost will not rotate or will rotate slowly to allow the system to increase the temperature quickly [6]. In other words, before the fuel cell system reaches the optimal temperature point, the air flow should be regulated to avoid oxygen starvation because the current air supply cannot meet the necessary air flow if higher load is applied during the period of increasing temperature. In another situation, there could be pretty low temperature in the circumstance, heat transferring very quickly between fuel cell and ambient air. And it means that the starvation of oxygen may happen when the cooling system needs less air flow to fix the fuel cell at optimal temperature point but the amount of oxygen as reactant almost doesn't vary. To prevent the system subject to oxygen starvation, an air supply circuit is used to control the axial fan, thereby forcing the ambient air to deliver the air to the fuel cell system. Few control strategies have been validated experimentally, although a few control strategies have been proposed for a closed-cathode PEM fuel cell system [15, 16].

In this work, the axial fan speed is controlled to track the desired oxygen excess ratio which is based on the optimal net power output and protects the fuel cell from coming up overtemperature phenomenon. There are two sets of oxygen excess ratio data. One is that optimizes the net power output. Another one is that avoids over-temperature. Combining the two data, the desired value always selects the higher value. The reason for selection is that the higher value not only considers the optimal net power output but also considers the temperature protection.

2 Modeling PEM Fuel Cell System

2.1 Electrochemical PEMFC Model

The electrochemical model of fuel cell is determined by the thermodynamic reversible potential E_{nernst}, activation over potential V_{act}, ohmic voltage losses V_{ohm}, and mass transport losses V_{mt}, as follows [2, 6, 9, 17]:

$$V_{fc} = E_{nernst} - V_{act} - V_{ohm} - V_{mt} = E_{nernst} - V_{act} - iR_{ohm} - i\left(c_2 \frac{i}{i_{\max}}\right)^{c_3} \quad (1)$$

The Nernst potential is also regarded as open circuit voltage of the single fuel cell expressed by (2).

$$E_{nernst} = 1.229 - (8.5 \times 10^{-4})(T_{st} - 298.15) + 4.308 \times 10^{-5} \times T_{st} \times \left[\ln(P_{H_2}) + \frac{1}{2}\ln(P_{O_2})\right] \quad (2)$$

The Nernst potential depends on fuel cell stack temperature T_{st}, partial pressure of oxygen P_{H_2}, and partial pressure of oxygen P_{O_2}. And i is the current density. The oxygen reduction reaction kinetics is slower than hydrogen oxidation reaction, which produces a non-linear voltage drop called activation polarization (V_{act}). The ohmic over potential is caused by resistance to the flow of electrons through the conductive components and ions through the membrane. The mass transport voltage drop V_{mt} and the consumption of reactant gases at the catalyst layers lead to concentration gradients. The polarization curve of fuel cell under real experimental test condition (at a constant air stoichiometrical operation point) is compared with the simulation result shown in Fig. 2.

In an open-cathode PEM fuel cell, a fan is usually mounted at the end to deliver air into the channels. A schematic with all input and output variables of the cathode flow system is shown in Fig. 3.

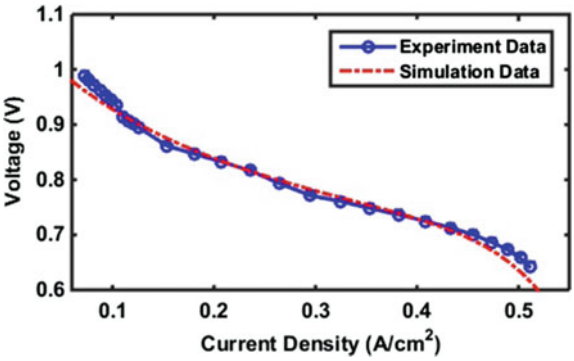


Fig. 2 Polarization curve of fuel cell

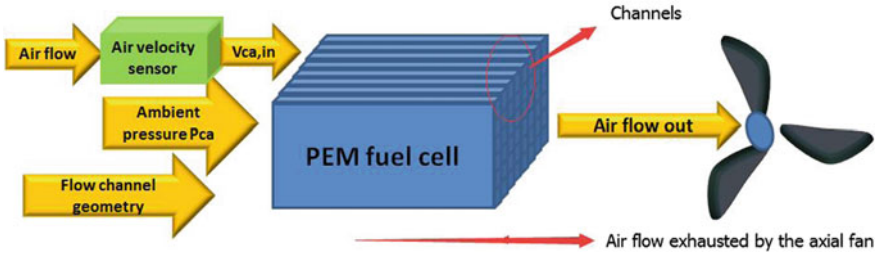


Fig. 3 Air delivery subsystem of the open cathode PEM fuel cell

Barreras et al. mentioned that axial fans can yield relatively large air flows, thereby overcoming pressure drops ranging from 70 Pa to 3 kPa [11]. They also showed that the pressure drop in the channels is around 80–150 Pa [12].

In an open-cathode PEMFC system, the mass flow rate depends on oxygen, nitrogen, and water vapor mass flow rate (kg/s), as follows:

$$\dot{m}_{ca,in} = \dot{m}_{O_2,ca,in} + \dot{m}_{N_2,ca,in} + \dot{m}_{H_2O,ca,in} \tag{3}$$

The total mass of air flow is related to the cross section area of channels, the air velocity in inlet channels, and the gas density. The influence of the pressure and temperature change would be neglected.

$$\dot{m}_{ca,in} = \rho_{air} \cdot v_{ca,in,average} \cdot A_{activechannels} \tag{4}$$

where ρ_{air} is the ambient air density, and $v_{ca,in,average}$ is cathode inlet air velocity measured by air flow velocity sensor, which is installed on the inlet of channels.

$$\dot{m}_{O_2,ca,in} = 0.21\,\dot{m}_{ca,in} \tag{5}$$

Regression analysis with least square method is used to obtain the relationship between sensor voltage and oxygen flow mass rate based on the experiment, as shown in Eq. (11). Figure 4 shows the relationship between flow rate sensor voltage and oxygen flow mass rate.

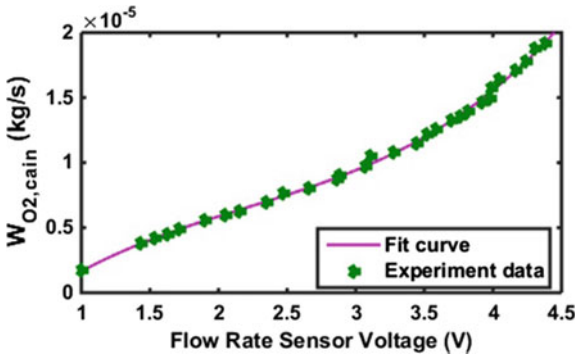


Fig. 4 The relation between flow rate sensor voltage and oxygen flow mass rate

Combining the oxidant and coolant flow of open-cathode PEMFC can predigest the structure of the fuel cell significantly. The oxygen excess ratio is defined as Eq. (6).

$$\lambda_{O_2} = \frac{\dot{m}_{O_2, ca, in}}{\dot{m}_{O_2, reacted}} \tag{6}$$

The rate of oxygen consumption is directly proportional to the system output current, which can be expressed as Eq. (7).

$$\dot{m}_{O_2, reacted} = M_{O_2} \frac{nI_{st}}{4F} \tag{7}$$

2.2 Thermal Model of the Fuel Cell

Temperature dynamics inside the open cathode fuel cell stack is presented employing lumped-parameter energy conversation equation:

$$m_{fc} C_{fc} \frac{dT_{fc}}{dt} = \dot{H}_{reac} - P_{st} - \dot{Q}_{rad, B2\,amb} - \dot{Q}_{cool} \tag{8}$$

The fuel cells' mass m_{fc} and specific heat capacity C_{fc} were defined by experiments, as given in published earlier [6]. H_{reac} is the total energy released from electrochemical reaction which is determined by electrochemical reaction defined as [6–8]

$$\begin{aligned} H_{reac} = {} & m_{H_2, reac}(H^0_{f, H_2} + C_{p, H_2}(T_{an, in} - T_0)) \\ & + m_{O_2, reac}(h^0_f + C_{p, O_2}(T_{ca, in} - T_0)) \\ & - m_{H_2O, gen}(H^0_{f, H_2O} + C_{p, H_2O}(T_{st} - T_0)) \end{aligned} \tag{9}$$

P_{st} is the electrical power output produced by the stack. $\dot{Q}_{rad, B2\,amb}$ is the released energy caused by radiation is determined as follow.

$$\dot{Q}_{rad, B2\,amb} \approx 2.18 \times 10^{-3} T_{st}^2 - 2.44 \times 10^{-3} T_{ca, in} \tag{10}$$

$\dot{Q}_{rad, B2\,amb}$ is highly influenced by the ambient temperature. Q_{cool} is released energy from convection which is mainly related to the rate of air flow passing through cathode channels.

$$\dot{Q}_{cool} = \rho_{air} A_{inlet} v_{air} C_{p,air} (T_{fc} - T_{amb}) \quad (11)$$

ρ_{air} is the ambient air density, $C_{p,air}$, specific heat capacity, v_{air} is the inlet velocity of air, A_{inlet} is crossed-sectional area. The more detail of the thermal models of the open cathode PEM fuel cell system have been presented in [6, 10].

3 Experiment Setup

The experiment is implemented on the platform as shown in Fig. 7. The fuel cell system is an open cathode design, the structure shown in Fig. 1. The axial fan is the only actuator servicing for air delivery subsystem and cooling subsystem simultaneously. The axil fan speed is determined by the supply voltage of fan which is controlled by the PWM (Pulse-Width Modulation) circuit. The core chip is TL494 and the power transistor type is TIP120 as shown in Fig. 5. The HT-01DV temperature sensor and thermal couple are used to test the fuel cell temperature. HT-01DV tiny temperature sensor is used to measure cell temperature and installed at the outlet of the cathode airflow exit. The thermal couple inserts into the fuel cell bipolar plate. And the air flow rate is measured by the D6F-W10A1 air flow velocity senor. The install location of air flow velocity sensor is based on the air velocity distribution. The exact location is shown in Fig. 6. WCS2705 sensor is used to measure the current. Labview data acquisition software was used in conjunction with the National Instruments Analog input/output module NI-PXI 6221 for collecting experiment data/transmitting command. The programmable electric load M9710 is selected simulating the fuel cell load.

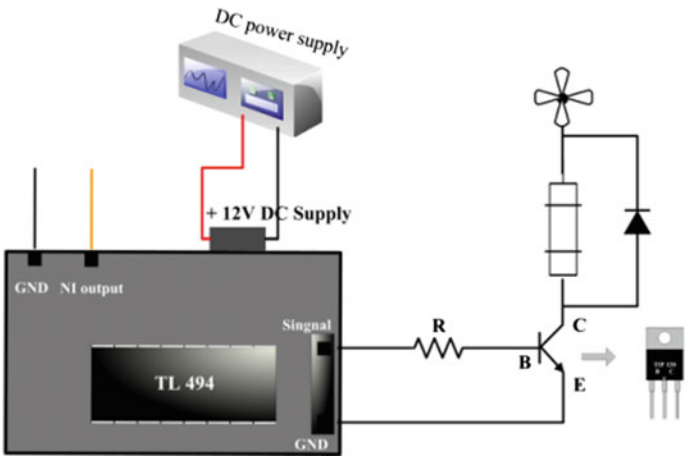


Fig. 5 Axial fan control circuit

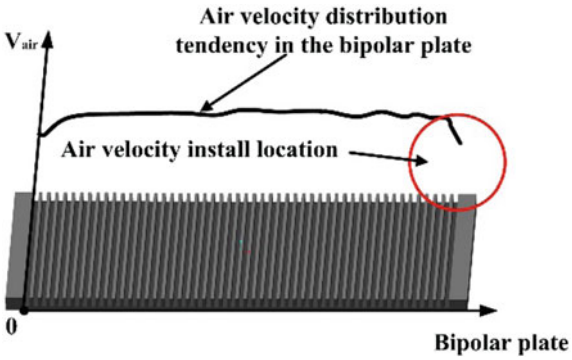


Fig. 6 Air velocity distribution in the bipolar plate

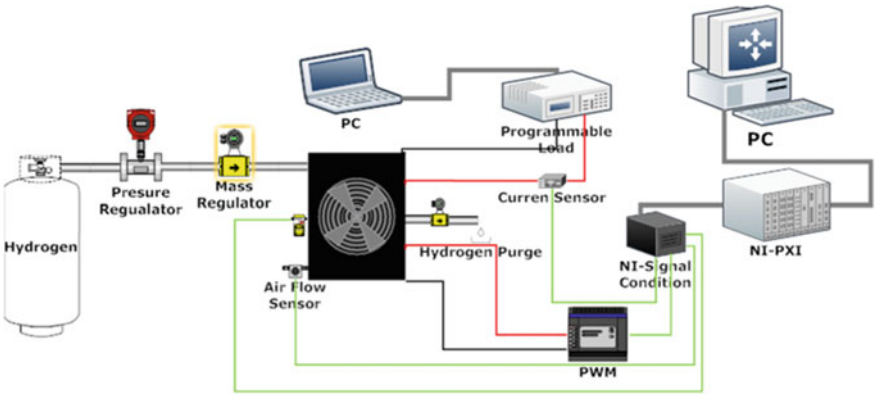


Fig. 7 The schematic diagram of an open-cathode PEM fuel cell system test rig

4 Control Objective and Experiment Analysis

The open cathode PEM fuel cell system is highly nonlinear system [18]. Usually, global optimization requires complicated computation and high accuracy model. In this work, optimal objective is net power output of open cathode PEM fuel cell based on controlling the oxygen excess ratio. At the same time, over-temperature situation is also considered.

To obtain the optimal control gains of PI controller, integral of time absolute error (ITAE) criterion is used by minimizing the integral of time multiply absolute error performance index [19].

$$u(t) = K_P\, e(t) + \int K_I\, e(t)\, dt \tag{12}$$

The conventional PI controller regulates the duty ratio of the PWM circuit to control axial fan speed. The reference governor is a function relevant to stack current [20, 21]. The destination of the function is searching for the desired reference points to optimize the net power output. Fig 8 shows control architecture for optimized net power.

$$P_{net} = P_{st} - P_{fan} = V_{st} * I_{st,net} \tag{13}$$

Based on the experimental data, the relationship between net power output and duty ratio was mapped at different load current. In the right corner of Fig. 9, the optimal duty ratios are located in the Current-Duty coordinate which is the track objective for optimizing the net power output.

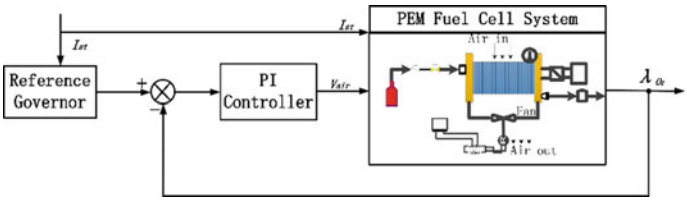


Fig. 8 Control architecture for optimized net power

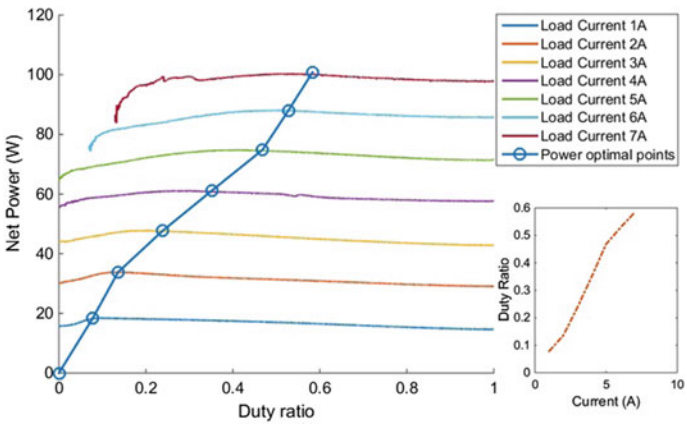


Fig. 9 Net power of the fuel cell of different load current I_{st} and corresponded fit curve in current-duty coordinate

And the experiment was carried out to validate whether the system temperature breaks up safety zone. The experiment result shows in Fig. 10 shows the current load varied condition. And the Fig. 11 indicates the voltage change and the system temperature change when the control system tracking the power optimal points shown in the right corner of the Fig. 9. Judging from Fig. 11, when the open cathode PEM fuel cell works in low current, the temperature of the system increases dramatically. There is existing the possibility that the temperature of the open cathode fuel cell system would break out the maximum temperature if the system always works in low current for long time. The reason causing the phenomenon is the air flow mass supplied by the air delivery subsystem less than the cooling subsystem required, which means these points for optimizing net power output of the fuel cell make fuel cell worked in dangerous situation. Too high temperature will dry the membrane and degrade the fuel cell.

To solve the problem, optimizing oxygen excess ratio based on the overtemperature protection is proposed. Using the same method as mapping optimal net power output as shown in Fig. 12, the Fig. 12 centralizes on finding the balance points between coolant rate and temperature rise rate. In the Fig. 12, the balance points were marked with a circle. And in the right corner, the fit curve was also put in Duty-Current coordinate. Figure 13 shows the process of coordinate system

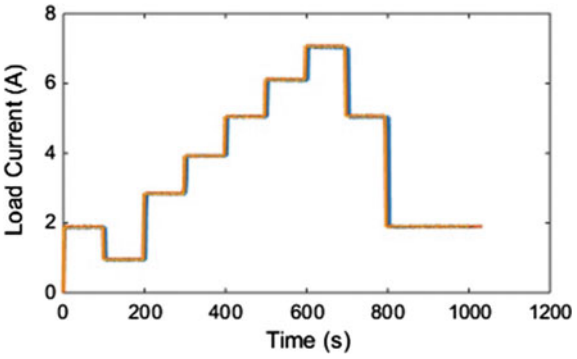


Fig. 10 Load varied curve

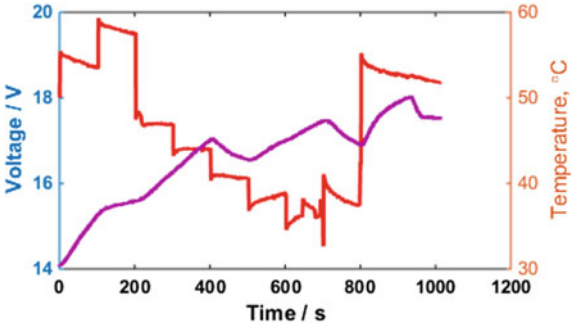


Fig. 11 Net power of the fuel cell and temperature varied

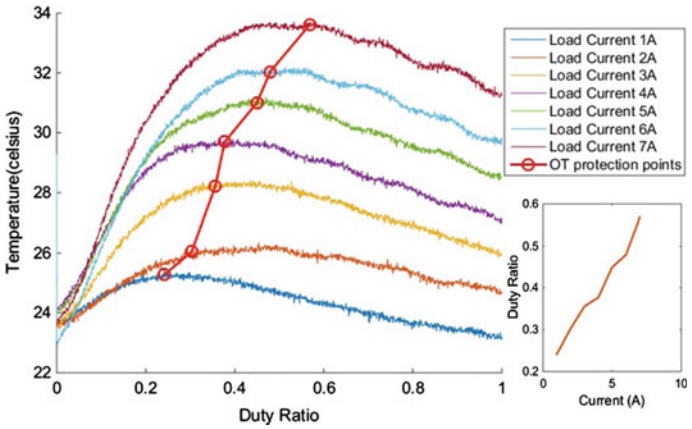


Fig. 12 The balance points between coolant rate and temperature rise rate for avoiding over temperature

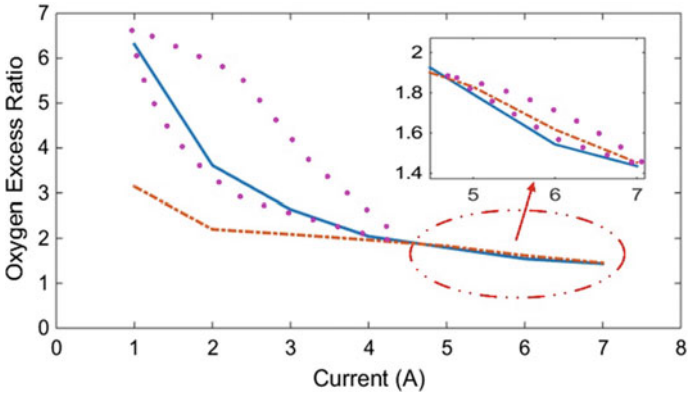


Fig. 13 Converting into oxygen excess ratio

conversion. Form the map of net power, the net power of system will not increase when it reaches the maximum point. Behind the maximum point the air flow mass rate influences the power system weaken than before. However, the influence for system temperature is increasing. The phenomenon can be utilized to restrain the temperature rising when fuel cell works in low current. Combination the two curve in the Fig. 13, the optimal oxygen excess ratio λ_{o_2} can be selected the higher point in the same coordinate system.

Before the cross-point of two lines, the oxygen excess ratio λ_{o_2} of temperature protection is superior to oxygen excess ratio λ_{o_2} of power optimization. Although both of the net power outputs are almost same, the oxygen excess ratio λ_{o_2} of temperature protection can protect the fuel cell from coming up overtemperature.

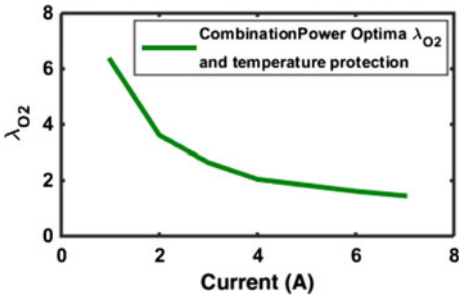


Fig. 14 Optimal oxygen excess ratio

The strategy sacrifices the little power to avoid the fuel cell both of overtemperature and starvation. Behind the cross-point, the oxygen excess ratio λ_{o_2} of power optimization needs more air flow mass than cooling system required. The oxygen excess ratio λ_{o_2} of power optimization is selected in the same reason. What need notes is, the oxygen input rate will increase rapidly when stack works in the high current, even though the difference of oxygen excess ratio λ_{o_2} is small. Because the oxygen consumed rate is fast in the high current. As Fig. 13 shows that the parts circled by pink dish line are selected to form a new serials optimization λ_{o_2} as the reference shown in Fig. 14.

The new optimal oxygen excess ratios are fit using polynomial function as:

$$\lambda_{O_2,\,referenc} = -0.057I_{st}^3 + 0.89I_{st}^2 - 4.669I_{st} + 10.07 \tag{14}$$

The experiment of the new optima oxygen excess ratios as the reference is validate in 100 W open cathode PEM fuel cell system. And the seeking algorithm is programed into the reference governor as shown in Fig. 8. The conventional PI controller parameters are tuned by ITAE criteria. K_P is 0.7 and K_I is 6.8. Figure 15 shows the control result. And using oxygen excess ratios λ_{o_2} of combination power optimization and overtemperature protection has a little lower voltage output than the original reference. But the system temperature can be restrain and the voltage output also has improved compared to the one which only considers restrain temperature (Figs. 16 and 17).

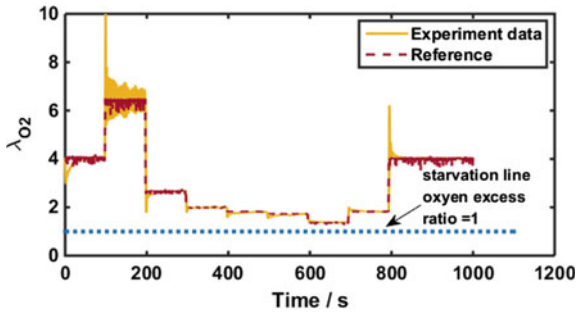


Fig. 15 Oxygen excess ratio regulation

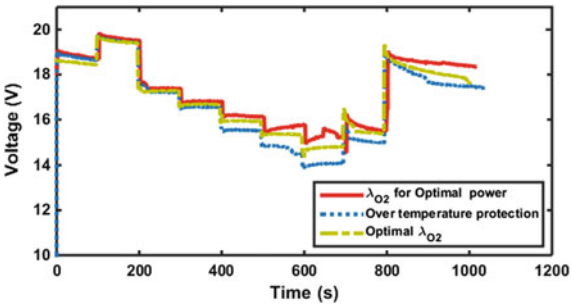


Fig. 16 Net power of the fuel cell with different control strategy

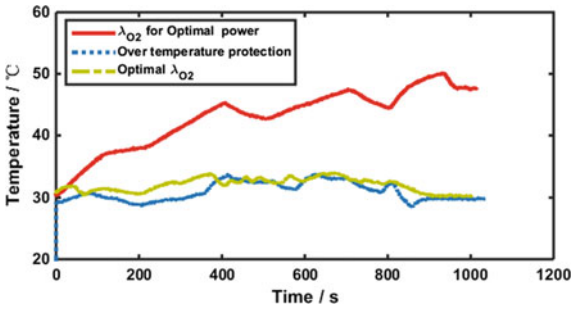


Fig. 17 Temperature varied

5 Conclusions

Open-cathode type fuel cell system has many advantages, such as its relatively cheap price and simple structure. However, appropriate control scheme is necessary to guarantee its performance without sacrificing power loss. One of the appropriate controls is air regulation to avoid oxygen shortage when the fuel cell requires abrupt load increase. The air regulation is performed by the air fan rotating speed, which is controlled by PWM duty ratio. To develop air control algorithms, fuel cell and fan dynamics model are estimated from experiments and empirical equations.

References

1. Y.-X. Wang, K. Ou, Y.-B. Kim, Modeling and experimental validation of hybrid proton exchange membrane fuel cell/battery system for power management control. Int. J. Hydrogen Energy **40**(35), 11713–11721 (2015)
2. J.T. Pukrushpan, H. Peng, A.G. Stefanopoulou, Simulation and analysis of transient fuel cell system performance based on a dynamic reactant flow model, in *Proceedings of ASME IMECE*, Nov. 2002, pp. 637–648
3. J.K. Gruber, M. Doll, C. Bordons, Design and experimental validation of a constrained MPC for the air feed of a fuel cell. Control Eng. Pract. **17**(8), 874–885 (2009)

4. Y.-X. Wang, Y.-B. Kim, Real-time control for air excess ratio of a PEM fuel cell system. IEEE/ASME Trans. Mechatron. **19**(3), 852–861 (2014)
5. P. Hu, G.-Y. Cao, X.-J. Zhu, M. Hu, Coolant circuit modeling and temperature fuzzy control of proton exchange membrane fuel cells. Int. J. Hydrogen Energy **35**(17), 9110–9123 (2010)
6. Y.-X. Wang, F.-F. Qin, K. Ou, Y.-B. Kim, Temperature control for a polymer electrolyte membrane fuel cell by using fuzzy rule. IEEE Trans. Energy Convers., doi:10.1109/TEC.2015.2511155
7. J.W. Ahn, S.Y. Choe, Coolant controls of a PEM fuel cell system. J. Power Sources **179**, 252–264 (2008)
8. L.A. Riascos, Relative humidity control in polymer electrolyte membrane fuel cells without extra humidification. J. Power Sources **184**, 204–211 (2008)
9. S. Strahl, A. Husar, P. Puleston, J. Riera, Performance improvement by temperature control of an open-cathode PEM fuel cell system. Fuel Cells **14**(3), 466–478 (2014)
10. S. Shahsavari, A. Desouza, M. Bahrami, E. Kjeang, Thermal analysis of air-cooled PEM fuel cells. Int. J. Hydrogen Energy **37**(23), 18261–18271 (2012)
11. F. Barreras, A.M. López, A. Lozano, J.E. Barranco, Experimental study of the pressure drop in the cathode side of air-forced open-cathode proton exchange membrane fuel cells. Int. J. Hydrogen Energy **36**(13), 7612–7620 (2011)
12. F. Barreras, A. Lozano1, J. Barroso, V. Roda, M. Maza, Theoretical model for the optimal design of air cooling systems of polymer electrolyte fuel cells. Application to a high-temperature PEMFC. Fuel Cells **13**(2), 227–237 (2013)
13. A. Arce, A.J. del Real, C. Bordons, D. Ramirez, Real-time implementation of a constrained MPC for efficient air flow control in a PEM fuel cell. IEEE Indust. Elec. **57**, 1892–1905 (2010)
14. Y.-X. Wang, D.-J. Xuan, Y.-B. Kim, Design and experimental implementation of time delay control for air supply in a polymer electrolyte membrane fuel cell system. Int. J. Hydrogen Energy **38**(30), 13381–13392 (2013)
15. D.-H. Jeon, K.-N. Kim, S.-M. Baek, J.-H. Nam, The effect of relative humidity of the cathode on the performance and the uniformity of PEM fuel cells. Int. J of Hydrogen Energy **36**(19), 12499–12511 (2011)
16. Q. Meyer, K. Ronaszegi, G. Pei-June, O. Curnick, S. Ashton, T. Reisch et al., Optimisation of air cooled, open-cathode fuel cells: current of lowest resistance and electro-thermal performance mapping. J. Power Sources **291**, 261–269 (2015)
17. F. Barbir, *PEM Fuel Cells*, Theory and Practice, 2nd edn. Sustainable World, ed. by R.C. Dorf (Elsevier, Burlington, MA, 2013)
18. J. Ishaku, N. Lotfi, H. Zomorodi, R.G. Landers, Control-oriented modeling for open-cathode fuel cell systems, American Control Conference, Portland, 2014
19. C.-C. Lee, Fuzzy logic in control systems: Fuzzy logic controller-Part I. IEEE Trans. Syst. Man Cybern. **20**, 404–418 (1990)
20. K. Ou, Y.-X. Wang, Z.-Z. Li, Y.-D. Shen, D.-J. Xuan, Feedforward fuzzy-PID control for air flow regulation of PEM fuel cell system. Int. J. Hydrogen Energy **40**(35), 11686–11695 (2015)
21. S. Caux, W. Hankache, M. Fadel, D. Hissel, On-line fuzzy energy management for hybrid fuel cell systems. Int. J. Hydrogen Energy **35**(5), 2134–2143 (2010)

Part IV
Modern Design Methodologies

General Solution Technique for Electroelastic Problems in Green Materials

M. Ishihara, Y. Ootao and Y. Kameo

Abstract In view of the engineering significance of piezoelectric bodies with D_∞ symmetry, such as bulk wood and film or fibers made of poly-L-lactic acid, an analytical technique is constructed for general solutions to electroelastic problems in these bodies. First, the constitutive equations which describe the symmetry are presented. Then, the displacement and electric field are expressed in terms of two types of displacement potential functions and the electric potential function, and their governing equations are obtained using the fundamental equations for the electroelastic field. As a result, the electroelastic field quantities are found to be expressed in terms of three functions, namely an elastic displacement potential functions and two piezoelastic displacement potential functions, and the former is found to satisfy a fourth-order differential equation and the latter to satisfy a Laplace equation with respect to the appropriately transformed spatial coordinates.

1 Introduction

Because of demands to reduce environmental loads, wooden materials and biodegradable polymers have attracted considerable attention as green materials. To ensure the quality of wooden materials, nondestructive evaluation (NDE) techniques need to be developed [1]. In addition, poly-L-lactic acid (PLLA) has been expected to be used for a human-machine interface, microtweezers, or catheters [2, 3]. Wood and film or fibers made of PLLA have macrosymmetry in

M. Ishihara (✉) · Y. Ootao
Osaka Prefecture University, 1-1 Gakuen-Cho, Naka-Ku, Sakai-Shi,
Osaka 599-8531, Japan
e-mail: ishihara@me.osakafu-u.ac.jp

Y. Ootao
e-mail: ootao@me.osakafu-u.ac.jp

Y. Kameo
Kyoto University, 53 Kawahara-Cho, Shogoin, Sakyo-Ku, Kyoto 606-8507, Japan
e-mail: kameo@frontier.kyoto-u.ac.jp

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_26

common; both of them have D_∞ symmetry, which is characterized by an ∞-fold rotation axis and a two-fold rotation axis perpendicular to it [4]. Bodies with D_∞ symmetry exhibit a unique electroelastic coupling behavior: coupling only occurs between the shear deformation in the plane parallel to the ∞-fold rotation axis and the electric poling perpendicular to the sheared plane.

Motivated by the above-mentioned demands for wood and PLLA, we constructed a general solution technique for the electroelastic fields in bodies with D_∞ symmetry [5, 6], and analyzed various kinds of boundary-value-problems [7–10]. In these studies [5–10], the electroelastic field quantities were described by several potential functions, each of which satisfy a Laplace equation with respect to the appropriately transformed spatial coordinates.

In these studies [5–10], however, a limitation for the material constants was imposed in order to simplify the solution technique: the stiffness for shear in the plane perpendicular to the ∞-fold rotation axis was assumed to be smaller than a critical value determined by other stiffness components. By this assumption, which happened to be justified for the materials treated in [5–10], the potential functions were found to satisfy the Laplace equations as mentioned above and the solution technique was simplified because the well-established body of knowledge on the Laplace equations was made available. In order to achieve the theoretical completeness or to prepare for the emergence of the tailored materials, however, the general solution technique which falls outside the abovementioned assumption must be constructed. In other words, the case in which the stiffness for shear in the plane perpendicular to the ∞-fold rotation axis is greater than the critical value mentioned above must be treated.

In this paper, therefore, an analytical technique is constructed for general solutions to electroelastic problems in bodies with D_∞ symmetry under the assumption that the stiffness for shear in the plane perpendicular to the ∞-fold rotation axis is greater than the critical value determined by other stiffness components. First, the fundamental equations of the electroelastic field and the solution technique under the previous assumption are presented, in which the electroelastic field quantities are expressed in terms of four functions, namely two elastic displacement potential functions and two piezoelastic displacement potential functions. Then, the behavior of characteristic roots in that technique is examined in detail, and the solution technique under the previous assumption is updated to the technique under the present assumption by introducing a new elastic displacement potential function. As a result, two elastic displacement potential functions, which were found to satisfy the Laplace equations, are replaced with one elastic displacement potential function which satisfies a fourth-order differential equation. Furthermore, the singular condition in which the stiffness for shear coincides with the critical value is also treated, and the corresponding solution technique is presented.

2 Solution Technique

2.1 Fundamental Equations of Electroelastic Field

We consider a piezoelectric body with D_∞ symmetry. The Cartesian coordinate system (x, y, z) is defined so that the z axis is parallel to the ∞-fold rotation axis of the body. Let (u_x, u_y, u_z), $(\varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{zz}, \varepsilon_{yz}, \varepsilon_{zx}, \varepsilon_{xy})$, $(\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{yz}, \sigma_{zx}, \sigma_{xy})$, (E_x, E_y, E_z), and (D_x, D_y, D_z) be the components of the displacement, strain, stress, electric field, and electric displacement, respectively. The displacement-strain relations are given as

$$\left.\begin{array}{l} \varepsilon_{xx} = \frac{\partial u_x}{\partial x},\ \varepsilon_{yy} = \frac{\partial u_y}{\partial y},\ \varepsilon_{zz} = \frac{\partial u_z}{\partial z}, \\ 2\varepsilon_{yz} = \frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y},\ 2\varepsilon_{zx} = \frac{\partial u_z}{\partial x} + \frac{\partial u_x}{\partial z},\ 2\varepsilon_{xy} = \frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \end{array}\right\}. \tag{1}$$

The constitutive equations of the body for an isothermal case are given as

$$\begin{Bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{Bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ & c_{11} & c_{13} & 0 & 0 & 0 \\ & & c_{33} & 0 & 0 & 0 \\ & & & c_{44} & 0 & 0 \\ & & & & c_{44} & 0 \\ \text{sym.} & & & & & \frac{c_{11}-c_{12}}{2} \end{bmatrix} \begin{Bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ 2\varepsilon_{yz} \\ 2\varepsilon_{zx} \\ 2\varepsilon_{xy} \end{Bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ e_{14} & 0 & 0 \\ 0 & -e_{14} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} E_x \\ E_y \\ E_z \end{Bmatrix}, \tag{2}$$

$$\begin{Bmatrix} D_x \\ D_y \\ D_z \end{Bmatrix} = \begin{bmatrix} 0 & 0 & 0 & e_{14} & 0 & 0 \\ 0 & 0 & 0 & 0 & -e_{14} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ 2\varepsilon_{yz} \\ 2\varepsilon_{zx} \\ 2\varepsilon_{xy} \end{Bmatrix} + \begin{bmatrix} \eta_{11} & 0 & 0 \\ & \eta_{11} & 0 \\ \text{sym.} & & \eta_{33} \end{bmatrix} \begin{Bmatrix} E_x \\ E_y \\ E_z \end{Bmatrix}, \tag{3}$$

where c_{ij}, η_{kl}, and e_{kj} denote the elastic stiffness constant, dielectric constant, and piezoelectric constant, respectively. For an equilibrium state to be stable, we finally have the conditions required for the material properties:

$$\left.\begin{array}{l} c_{11} > 0,\ c_{33} > 0,\ c_{44} > 0,\ c_{11} - c_{12} > 0,\ c_{11} + c_{12} > 0, \\ c_{11}c_{33} - c_{13}^2 > 0,\ (c_{11} + c_{12})c_{33} - 2c_{13}^2 > 0,\ \eta_{11} > 0,\ \eta_{33} > 0 \end{array}\right\}, \tag{4}$$

which was proved in the previous paper [5]. The equilibrium equations of stresses and the Gauss law are given, respectively, by

$$\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{zx}}{\partial z} = 0,\ \frac{\partial \sigma_{yy}}{\partial y} + \frac{\partial \sigma_{yz}}{\partial z} + \frac{\partial \sigma_{xy}}{\partial x} = 0,\ \frac{\partial \sigma_{zz}}{\partial z} + \frac{\partial \sigma_{zx}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial y} = 0, \tag{5}$$

$$\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} = 0. \tag{6}$$

2.2 Solution Technique Under Harmonic Function Theory

Solution technique

Under the assumption that the stiffness for shear in the plane perpendicular to the ∞-fold rotation axis is smaller than a critical value:

$$c_{44} < \frac{\sqrt{c_{11}c_{33}} - c_{13}}{2}, \tag{7}$$

the components of displacement and electric field to satisfy Eqs. (1)–(6) are described by elastic displacement potential functions φ_i, piezoelastic displacement potential functions ϑ_i, and electric potential function Φ, respectively, as follows [5]:

$$u_x = \sum_{i=1}^{2}\left(\frac{\partial \varphi_i}{\partial x} + \frac{\partial \vartheta_i}{\partial y}\right),\ u_y = \sum_{i=1}^{2}\left(\frac{\partial \varphi_i}{\partial y} - \frac{\partial \vartheta_i}{\partial x}\right),\ u_z = \frac{\partial}{\partial z}\sum_{i=1}^{2} k_i \varphi_i, \tag{8}$$

$$E_x = -\frac{\partial \Phi}{\partial x},\ E_y = -\frac{\partial \Phi}{\partial y},\ E_z = -\frac{\partial \Phi}{\partial z}. \tag{9}$$

φ_i, ϑ_i, and Φ are governed by

$$\left(\Delta_{\mathrm{p}} + \mu_i \frac{\partial^2}{\partial z^2}\right)\varphi_i = 0\ (i = 1, 2), \tag{10}$$

$$\left(\Delta_{\mathrm{p}} + \upsilon_i \frac{\partial^2}{\partial z^2}\right)\vartheta_i = 0\ (i = 1, 2), \tag{11}$$

$$\frac{e_{14}}{c_{44}}\mu_3\frac{\partial\Phi}{\partial z}=\sum_{i=1}^{2}\left(\Delta_{\mathrm{p}}+\mu_3\frac{\partial^2}{\partial z^2}\right)\vartheta_i, \tag{12}$$

where (μ_1,μ_1) and (υ_1,υ_2) are the roots of quadratic equations with respect to μ and υ as

$$c_{11}c_{44}\mu^2-\left(c_{11}c_{33}-c_{13}^2-2c_{13}c_{44}\right)\mu+c_{33}c_{44}=0, \tag{13}$$

$$\upsilon^2-\left[\mu_3\left(1+k_{\mathrm{couple}}^2\right)+\eta\right]\upsilon+\mu_3\eta=0, \tag{14}$$

and

$$k_i=\frac{c_{11}\mu_i-c_{44}}{c_{13}+c_{44}}=\frac{(c_{13}+c_{44})\mu_i}{c_{33}-c_{44}\mu_i}\ (i=1,2), \tag{15}$$

$$\mu_3=\frac{2c_{44}}{c_{11}-c_{12}},\ \eta=\frac{\eta_{33}}{\eta_{11}},\ k_{\mathrm{couple}}^2=\frac{e_{14}^2}{c_{44}\eta_{11}},\ \Delta_{\mathrm{p}}=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}. \tag{16}$$

The components of strain, stress, and electric displacement are obtained by substituting Eqs. (8) and (9) into Eqs. (1)–(3).

Properties of characteristic roots

Then, the properties of υ_1 and υ_2 are discussed. The relations between the roots and coefficients of Eq. (14) are obtained as

$$\upsilon_1+\upsilon_2=\mu_3\left(1+k_{\mathrm{couple}}^2\right)+\eta,\ \upsilon_1\upsilon_2=\mu_3\eta, \tag{17}$$

or Eq. (14) is concretely solved as

$$\left.\begin{aligned}\upsilon_1&=\frac{\left[\mu_3\left(1+k_{\mathrm{couple}}^2\right)+\eta\right]+\sqrt{(\mu_3-\eta)^2+2\mu_3k_{\mathrm{couple}}^2(\mu_3+\eta)+\left(\mu_3k_{\mathrm{couple}}^2\right)^2}}{2},\\ \upsilon_2&=\frac{\left[\mu_3\left(1+k_{\mathrm{couple}}^2\right)+\eta\right]-\sqrt{(\mu_3-\eta)^2+2\mu_3k_{\mathrm{couple}}^2(\mu_3+\eta)+\left(\mu_3k_{\mathrm{couple}}^2\right)^2}}{2},\end{aligned}\right\}. \tag{18}$$

From Eqs. (4), (16), and (17), we have

$$\upsilon_1+\upsilon_2>0,\ \upsilon_1\upsilon_2>0,\ (\mu_3-\eta)^2+2\mu_3k_{\mathrm{couple}}^2(\mu_3+\eta)+\left(\mu_3k_{\mathrm{couple}}^2\right)^2\geq 0. \tag{19}$$

The equal sign in the third equation of Eq. (19) holds only for

$$\mu_3 = \eta,\ e_{14} = 0, \tag{20}$$

which we exclude because we treat a piezoelectric body. Thus, from Eqs. (18) and (19), we obtain the condition:

$$\upsilon_1 \neq \upsilon_2,\ \upsilon_1 > 0,\ \upsilon_2 > 0, \tag{21}$$

which implies that Eq. (11) for $i = 1$ and $i = 2$ are distinct from each other and both of them are the Laplace equations with respect to the appropriately transformed spatial coordinates. It should be noted that Eq. (21) holds regardless of the condition described by Eq. (7).

For μ_1 and μ_2, the relations between the roots and coefficients of Eq. (13) are obtained as

$$\mu_1 + \mu_2 = \frac{c_{11}c_{33} - c_{13}^2 - 2c_{13}c_{44}}{c_{11}c_{44}},\ \mu_1\mu_2 = \frac{c_{33}}{c_{11}}. \tag{22}$$

Equation (13) is concretely solved as

$$\left.\begin{aligned} \mu_1 &= \frac{\left(c_{11}c_{33} - c_{13}^2 - 2c_{13}c_{44}\right) + \sqrt{\left(c_{11}c_{33} - c_{13}^2\right)D_{\text{discriminant}}}}{2c_{11}c_{44}} \\ \mu_2 &= \frac{\left(c_{11}c_{33} - c_{13}^2 - 2c_{13}c_{44}\right) - \sqrt{\left(c_{11}c_{33} - c_{13}^2\right)D_{\text{discriminant}}}}{2c_{11}c_{44}} \end{aligned}\right\}, \tag{23}$$

where

$$D_{\text{discriminant}} \equiv \left(c_{11}c_{33} - c_{13}^2\right) - 4c_{44}(c_{13} + c_{44}). \tag{24}$$

Under the assumption described by Eq. (7), the inequality

$$D_{\text{discriminant}} > 0 \tag{25}$$

holds. Then, both μ_1 and μ_2 are found to be real and distinct from each other, and to have the same sign from Eqs. (4), (22), and (23). From Eqs. (4) and (25), we have

$$D_{\text{discriminant}} + \left(c_{11}c_{33} - c_{13}^2\right) > 0. \tag{26}$$

Further, we can calculate

$$D_{\text{discriminant}} + \left(c_{11}c_{33} - c_{13}^2\right) = 2\left(c_{11}c_{33} - c_{13}^2 - 2c_{13}c_{44}\right) - 4c_{44}^2. \tag{27}$$

From Eqs. (4), (22), (26), and (27), we have

$$\mu_1 + \mu_2 > 0. \tag{28}$$

Therefore, under Eq. (7), we have

$$\mu_1 \neq \mu_2,\ \mu_1 > 0,\ \mu_2 > 0, \tag{29}$$

which implies that Eq. (10) for $i = 1$ and $i = 2$ are distinct from each other and both of them are the Laplace equations with respect to the appropriately transformed spatial coordinates. Therefore the well-established body of knowledge on the theory of Laplace equations can be applied to Eq. (10). It should be noted that the boundary-value-problems in previous papers [5–10] were treated under this framework.

2.3 Solution Technique Under Non-harmonic Function Theory

This subsection treats the counterpart of the condition described by Eq. (7). In other words, the stiffness for shear in the plane perpendicular to the ∞-fold rotation axis is assumed to be greater than a critical value:

$$c_{44} > \frac{\sqrt{c_{11}c_{33}} - c_{13}}{2}. \tag{30}$$

Properties of characteristic roots
From Eqs. (24) and (30), we have

$$D_{\text{discriminant}} < 0. \tag{31}$$

Then, from Eqs. (4) and (23), we have

$$\mu_2 = \bar{\mu}_1; \tag{32}$$

namely, μ_1 and μ_2 are found to be a conjugate pair of complex numbers. From Eqs. (10) and (32), we have

$$\varphi_2 = \bar{\varphi}_1; \tag{33}$$

namely, the corresponding elastic displacement potential functions φ_1 and φ_2 are found to be a conjugate pair of complex-valued functions. In that case, although Eqs. (8)–(11) are still valid, we are forced to treat complex-valued functions. Procedures to overcome this difficulty are presented as follows.

Introduction of new elastic displacement potential function From Eqs. (32) and (33), both $(\varphi_1+\varphi_2)$ and $(\mu_1\varphi_1+\mu_2\varphi_2)$ are found to be real-valued functions. Therefore, real-valued functions defined as

$$\varphi_R \equiv \varphi_1+\varphi_2,\ \varphi_{\mu R} \equiv \mu_1\varphi_1+\mu_2\varphi_2 \tag{34}$$

are introduced. In that case, from Eqs. (8) and (15), the displacement is expressed in terms of φ_R and $\varphi_{\mu R}$ as

$$\left.\begin{aligned} u_x &= \frac{\partial\varphi_R}{\partial x}+\sum_{i=1}^{2}\frac{\partial\vartheta_i}{\partial y},\ u_y = \frac{\partial\varphi_R}{\partial y}-\sum_{i=1}^{2}\frac{\partial\vartheta_i}{\partial x},\\ u_z &= \frac{c_{11}}{c_{13}+c_{44}}\frac{\partial\varphi_{\mu R}}{\partial z}-\frac{c_{44}}{c_{13}+c_{44}}\frac{\partial\varphi_R}{\partial z}\end{aligned}\right\}. \tag{35}$$

From Eqs. (10) and (34), we can calculate

$$\frac{\partial^2\varphi_{\mu R}}{\partial z^2} = \mu_1\frac{\partial^2\varphi_1}{\partial z^2}+\mu_2\frac{\partial^2\varphi_2}{\partial z^2} = -\Delta_{\mathrm{p}}\varphi_1-\Delta_{\mathrm{p}}\varphi_2 = -\Delta_{\mathrm{p}}\varphi_R, \tag{36}$$

and $\varphi_{\mu R}$ and φ_R are found to be related to each other. Therefore, by introducing the real-valued function φ^R, defined as

$$\frac{\partial\varphi^R}{\partial z} \equiv \varphi_R = \varphi_1+\varphi_2, \tag{37}$$

equation (35) is rewritten as

$$\left.\begin{aligned} u_x &= \frac{\partial^2\varphi^R}{\partial z\partial x}+\sum_{i=1}^{2}\frac{\partial\vartheta_i}{\partial y},\ u_y = \frac{\partial^2\varphi^R}{\partial y\partial z}-\sum_{i=1}^{2}\frac{\partial\vartheta_i}{\partial x},\\ u_z &= -\frac{c_{11}}{c_{13}+c_{44}}\Delta_{\mathrm{p}}\varphi^R-\frac{c_{44}}{c_{13}+c_{44}}\frac{\partial^2\varphi^R}{\partial z^2}\end{aligned}\right\}. \tag{38}$$

From Eqs. (10) and (37), we have

$$\left.\begin{aligned} &\left(\Delta_{\mathrm{p}}+\mu_1\frac{\partial^2}{\partial z^2}\right)\left(\Delta_{\mathrm{p}}+\mu_2\frac{\partial^2}{\partial z^2}\right)\left(\frac{\partial\varphi^R}{\partial z}\right)\\ &\quad=\left(\Delta_{\mathrm{p}}+\mu_1\frac{\partial^2}{\partial z^2}\right)\left(\Delta_{\mathrm{p}}+\mu_2\frac{\partial^2}{\partial z^2}\right)(\varphi_1+\varphi_2)=0\end{aligned}\right\}. \tag{39}$$

By integrating Eq. (39) with respect to z and choosing φ^R for the trivial solution $\left(u_x,u_y,u_z,\Phi\right)\equiv(0,0,0,0)$ as

$$\varphi^R \equiv 0, \tag{40}$$

it is found that, from Eq. (39), the governing equation for φ^R is obtained as

$$\left(\Delta_{\mathrm{p}} + \mu_1 \frac{\partial^2}{\partial z^2}\right)\left(\Delta_{\mathrm{p}} + \mu_2 \frac{\partial^2}{\partial z^2}\right)\varphi^R = 0; \tag{41}$$

namely,

$$\left[\Delta_{\mathrm{p}}\Delta_{\mathrm{p}} + (\mu_1 + \mu_2)\Delta_{\mathrm{p}} \frac{\partial^2}{\partial z^2} + \mu_1\mu_2 \frac{\partial^4}{\partial z^4}\right]\varphi^R = 0. \tag{42}$$

Although both μ_1 and μ_2 are complex numbers, $(\mu_1 + \mu_2)$ and $\mu_1\mu_2$ are real numbers, as shown in Eq. (22); therefore, Eq. (42) is found to be the governing equation with real coefficients for the real-valued function φ^R. The general solutions for the electroelastic field are obtained by substituting the general solutions to Eqs. (11) and (42) into Eqs. (1)–(3), (9), (12), and (38).

In summary, by introducing the new elastic displacement potential function defined by Eq. (37), the governing equations with complex coefficients, Eq. (10), are updated to the equation with real coefficients as in Eq. (42), and therefore the solution procedures are simplified to a great extent, which is the prime achievement in this paper.

2.4 Solution Technique for Degenerate Case

This subsection treats the singular condition which separates the conditions described by Eqs. (7) and (30). In other words, the stiffness for shear in the plane perpendicular to the ∞-fold rotation axis is assumed to have a critical value:

$$c_{44} = \frac{\sqrt{c_{11}c_{33}} - c_{13}}{2}. \tag{43}$$

Then, from Eq. (24), we have

$$D_{\text{discriminant}} = 0. \tag{44}$$

Then, because φ_2, μ_2, and k_2 are found to degenerate into φ_1, μ_1, and k_1, respectively, from Eqs. (10), (15), and (23), we introduce an additional elastic displacement potential λ as

$$\left.\begin{aligned} u_x &= \frac{\partial\varphi}{\partial x} + \sum_{i=1}^{2}\frac{\partial\vartheta_i}{\partial y} + z\frac{\partial\lambda}{\partial x},\ u_y = \frac{\partial\varphi}{\partial y} - \sum_{i=1}^{2}\frac{\partial\vartheta_i}{\partial x} + z\frac{\partial\lambda}{\partial y}, \\ u_z &= k\frac{\partial\varphi}{\partial z} + k'\frac{\partial(z\lambda)}{\partial z} - k''\lambda \end{aligned}\right\}. \tag{45}$$

In Eq. (45), φ and k denote the degenerated forms of φ_i and k_i $(i = 1, 2)$, respectively, and, from Eqs. (10), (15), (23), (24), and (44), they satisfy

$$k = \frac{c_{11}\mu - c_{44}}{c_{13} + c_{44}} = \frac{(c_{13} + c_{44})\mu}{c_{33} - c_{44}\mu}, \tag{46}$$

$$\Delta_{\mathrm{p}}\varphi + \mu\frac{\partial^2\varphi}{\partial z^2} = 0, \tag{47}$$

and

$$\mu = \frac{c_{11}c_{33} - c_{13}^2 - 2c_{13}c_{44}}{2c_{11}c_{44}} = \frac{c_{13} + 2c_{44}}{c_{11}}; \tag{48}$$

k' and k'' are unknown coefficients at present. Substituting Eqs. (1), (9), and (45) into Eqs. (2) and (3) and the results into Eqs. (5) and (6), we have

$$\left.\begin{aligned}
&z\cdot\frac{\partial}{\partial x}\left\{c_{11}\left[\Delta_{\mathrm{p}}\lambda + \frac{k'c_{13} + (1+k')c_{44}}{c_{11}}\frac{\partial^2\lambda}{\partial z^2}\right]\right\} \\
&\quad + [(2k' - k'')c_{13} + (2k' - k'' + 2)c_{44}]\frac{\partial^2\lambda}{\partial z\partial x} = 0, \\
&z\cdot\frac{\partial}{\partial y}\left\{c_{11}\left[\Delta_{\mathrm{p}}\lambda + \frac{k'c_{13} + (1+k')c_{44}}{c_{11}}\frac{\partial^2\lambda}{\partial z^2}\right]\right\} \\
&\quad + [(2k' - k'')c_{13} + (2k' - k'' + 2)c_{44}]\frac{\partial^2\lambda}{\partial y\partial z} = 0, \\
&\frac{\partial}{\partial z}\left\{z\cdot[c_{13} + (1+k')c_{44}]\left[\Delta_{\mathrm{p}}\lambda + \frac{k'c_{33}}{c_{13} + (1+k')c_{44}}\frac{\partial^2\lambda}{\partial z^2}\right]\right\} \\
&\quad - k''c_{44}\left[\Delta_{\mathrm{p}}\lambda + \frac{(k'' - 2k')c_{33}}{k''c_{44}}\frac{\partial^2\lambda}{\partial z^2}\right] = 0
\end{aligned}\right\}. \tag{49}$$

It should be noted that, because φ, ϑ_i, and Φ have been determined so as to satisfy Eqs. (1)–(3), (5), and (6), they do not appear in Eq. (49). Here, we choose k' and k'' so that they satisfy

$$k' = k,\ (2k' - k'')c_{13} + (2k' - k'' + 2)c_{44} = 0; \tag{50}$$

namely,

$$k' = k,\ k'' = \frac{2[kc_{13} + (1+k)c_{44}]}{c_{13} + c_{44}}. \tag{51}$$

Then, from Eqs. (46), (50), and (51), the coefficients in Eq. (49) have the relation:

$$\frac{k'c_{13} + (1+k')c_{44}}{c_{11}} = \frac{k'c_{33}}{c_{13} + (1+k')c_{44}} = \mu. \tag{52}$$

Therefore, from Eqs. (50)–(52), Eq. (49) is found to be satisfied when λ satisfies both

$$\Delta_{\mathrm{p}}\lambda + \mu\frac{\partial^2\lambda}{\partial z^2} = 0 \tag{53}$$

and

$$\Delta_{\mathrm{p}}\lambda + \frac{(k'' - 2k')c_{33}}{k''c_{44}}\frac{\partial^2\lambda}{\partial z^2} = 0. \tag{54}$$

By substituting Eq. (51) into Eq. (54), we have

$$\Delta_{\mathrm{p}}\lambda + \frac{c_{33}}{kc_{13} + (1+k)c_{44}}\frac{\partial^2\lambda}{\partial z^2} = 0. \tag{55}$$

k, k', and k'' can be calculated from Eqs. (46), (48), and (51) as

$$k = k' = 1,\ k'' = \frac{2(c_{13} + 2c_{44})}{c_{13} + c_{44}}. \tag{56}$$

By considering Eqs. (52) and (56), Eqs. (53) and (55) are found to be identical. Therefore, the general solutions for the electroelastic field are obtained by substituting the general solutions to Eqs. (47), (11), and (53), and Eq. (56) into Eqs. (1)–(3), (9), (12), and (45).

3 Conclusion

In order to achieve the theoretical completeness or to prepare for the emergence of the new materials, the analytical technique for general solutions to electroelastic problems in bodies with D_∞ symmetry was constructed for the case where the stiffness for shear in the plane perpendicular to the ∞-fold rotation axis is greater than or equal to, the critical value determined by other stiffness components. As a result, the electroelastic field quantities were found to be expressed in terms of two piezoelastic displacement potential functions and a newly introduced elastic displacement potential function. This new function and its governing equation were found to be real-valued, which facilitates the solution procedures of the relevant electroelastic problems.

References

1. T. Nakai, H. Yamamoto, M. Hamatake, T. Nakao, Initial shapes of stress-strain curve of wood specimen subjected to repeated combined compression and vibration stresses and the piezoelectric behavior. J. Wood. Sci. **52**, 539–543 (2006)
2. M. Ando, H. Kawamura, H. Kitada, Y. Sekimoto, T. Inoue, Y. Tajitsu, Pressure-sensitive touch panel based on piezoelectric poly(L-lactic acid) film. Jpn. J. Appl. Phys. **52**, 09KD17 (2013)
3. Y. Tajitsu, Piezoelectricity of chiral polymeric fiber and its application in biomedical engineering. IEEE Trans. Ultrason. Ferroelectr. Freq. Control. **55**, 1000–1008 (2008)
4. S.K. Kim, *Group Theoretical Methods and Applications to Molecules and Crystals* (Cambridge University Press, Cambridge, 1999)
5. M. Ishihara, Y. Ootao, Y. Kameo, Analytical technique for electroelastic field in piezoelectric bodies belonging to point group D_{∞}. J. Wood. Sci. **61**, 270–284 (2015)
6. M. Ishihara, Y. Ootao, Y. Kameo, A general solution technique for electroelastic fields in piezoelectric bodies with D_{∞} symmetry in cylindrical coordinates. J. Wood. Sci. **62**, 29–41 (2016)
7. M. Ishihara, Y. Ootao, Y. Kameo, Electroelastic response of a piezoelectric semi-infinite body with D_{∞} symmetry to surface friction. Int. J. Eng. Res. Appl. **5**(6–2), 26–32 (2015)
8. M. Ishihara, Y. Ootao, Y. Kameo, An electroelastic problem of green materials subjected to surface friction. Int. J. Civil. Struct. Eng. **2**, 49–52 (2015)
9. M. Ishihara, Y. Ootao, Y. Kameo, Electroelastic response of cylindrical fiber with D_{∞} symmetry exposed to local electric field through opposed electrode pair. Polym-Plast. Technol. Eng. **55**, 900–910 (2016)
10. M. Ishihara, Y. Ootao, Y. Kameo, Electroelastic response of a piezoelectric cylinder with D_{∞} symmetry under axisymmetric mechanical loading. Mech. Res. Commun. **74**, 1–7 (2016)

Supply Chain Network Design with Dynamic Scheduling and Cooperative Negotiation

Yoshitaka Tanimizu

Abstract Supply chain management has been studied for controlling material flow and information flow among different manufacturing companies. The recent research trends on supply chain management have been toward flexible and dynamic supply chains. In the dynamic supply chain environments, different organizations, such as law material suppliers, parts suppliers, and assembly manufacturers, can change business partners for finding suitable business partners and entering into profitable contracts. This paper reviews a basic dynamic supply chain model and discusses a minimum model for multi-layered dynamic supply chains.

Keywords Supply chain management · Scheduling · Genetic algorithm

1 Introduction

Many researchers have investigated supply chain management for planning and controlling material flow and information flow among different organizations, such as raw material suppliers, parts suppliers, assembly manufacturers, and customers. It has been attracting attention as a means to increase opportunities for marketing and to reduce unnecessary inventory through the communications among the organizations. Recent development of the Internet technologies provides us with a global and high-speed communication among the independent organizations in the supply chains.

A large number of researches on supply chain management deal with static supply chain configuration. They proposed various methods to reduce unnecessary inventory in the supply chains. These researches assumed that a single central decision maker controlled the whole supply chains. However, the actual supply chains include a set of independent organizations, which have own decision criteria

Y. Tanimizu (✉)
Graduate School of Engineering, Osaka Prefecture University, Osaka, Japan
e-mail: tanimizu@me.osakafu-u.ac.jp

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_27

and objectives. Each organization requires a decision support system for configuring the most suitable supply chains dynamically and continuously.

Researches in recent years focus on dynamic supply chain configurations, which can change the configurations of supply chains within short time intervals. Kaihara and Fujii [1] developed a multi-agent system for dynamic supply chain configurations and a combinatorial auction algorithm for different organizations. Piramuthu [2] and Emerson [3] proposed a framework for dynamically forming and reconfiguring a supply chain as per the dictates of order specifications, and showed performance improvements of the adaptive supply chain configuration framework over static configurations.

A new research target is a group of organizations having a make-to-order or a build-to-order production system, which has no inventories of manufactured and assembled parts or products in advance. Tanimizu et al. [4, 5] proposed a two-layered dynamic supply chain model consisting of two kinds of organizations; those are a client and a supplier. A client requires a product and sends an order of the required product to all suppliers. A supplier generates an offer for the required product and sends it to the client. An offer includes the information about the possible delivery time and bid price of the required product. The possible delivery time and bid price are determined by both the modification process of a supplier's production schedule and the negotiation process between the client and the supplier. A three-layered dynamic supply chain model was provided as a basic model for representing a multi-layered dynamic supply chain. A model component in the middle layer of the supply chain was named as a manufacturer, which sends not only an order for an upper layer but also an offer for a lower layer.

The reminder of this paper is organized as follows. Section 2 reviews the two-layered dynamic supply chain model. Section 3 describes the three-layered dynamic supply chain model. Section 4 shows a summary of the work.

2 Two-Layered Dynamic Supply Chain Model

A two-layered supply chain model includes two components; those are a supplier and a client [4, 5]. We proposed an auction-based method for making decisions and for negotiating between a client and a supplier. The method is used to determine the suitable price and delivery time of a product through the iteration of both the modification process of the supplier's production schedule by using a genetic algorithm (GA) and the negotiation process between the client and the supplier.

Figure 1 shows a negotiation process for determining the price and delivery time of a required product between a client and a supplier. The negotiation process between a client and a supplier is summarized in the following steps.

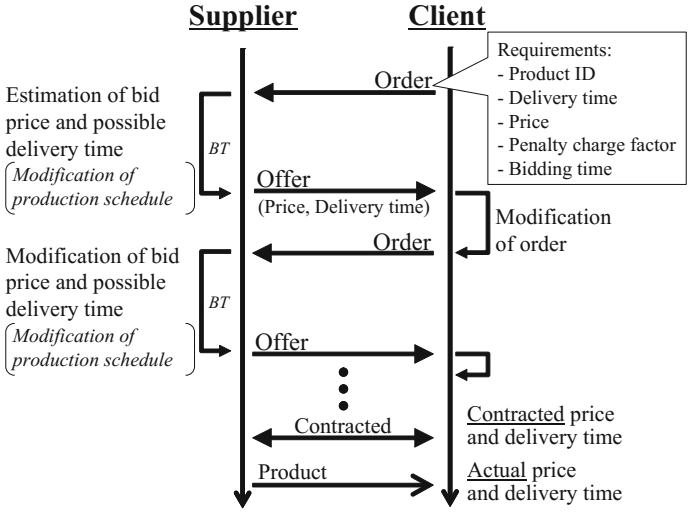


Fig. 1 Negotiation processes between a client and a supplier

1. A client requires a product. Then, the client generates a new order for the product and sends it to all suppliers.
2. Each supplier has a production schedule including the production of pre-ordered products already under contract. After adding the newly ordered product to the existing production schedules, the supplier improves the production schedule by using the GA. Then, the supplier generates an offer for the ordered product. The offer includes the information of the possible delivery time and bid price of the product.
3. All offers sent from suppliers are evaluated by the client. An offer with the lowest bid price is accepted, if the offer satisfies the requirement both in terms of the price and delivery time. However, if no offers satisfy the requirement, the client relaxes it and sends the modified order to suppliers.

This negotiation process is repeated until an offer satisfies the order and one of the suppliers enters into a contract with the client.

3 Three-Layered Dynamic Supply Chain Model

3.1 Basic Model for Multi-Layered Dynamic Supply Chains

Figure 2 shows a configuration of multi-layered dynamic supply chain. We named a model component in the middle layer of the supply chain as a manufacturer in the existing research [6]. A manufacturer has neither the stock of parts nor the stock of

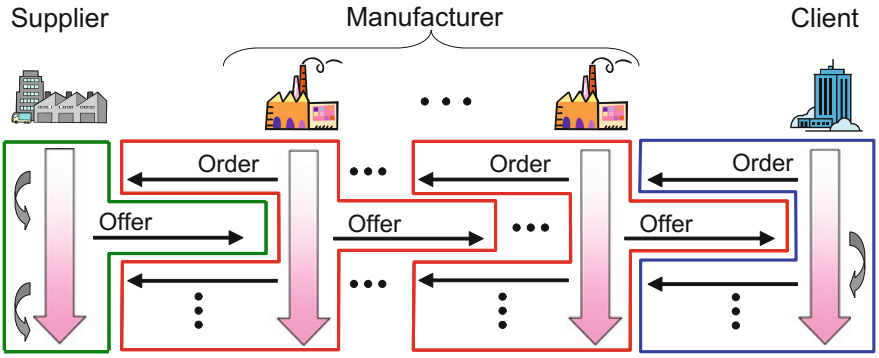


Fig. 2 Multi-layered dynamic supply chain

completed products. The basic function of manufacturer is sending an order for the upper layer and also sending an offer for the lower layer.

We proposed a three-layered dynamic supply chain model as a minimum model of multi-layered dynamic supply chains. The model includes three model components; a supplier, a manufacturer, and a client. A manufacturer sends both an order of a part to a supplier and an offer of a product to a client. A manufacturer determines the suitable price and delivery time of a product through the iteration of both the modification process of all organizations' production schedules and the negotiation process among the three model components.

3.2 Negotiation Protocol

Not only a supplier but also a manufacturer needs enough time to improve their production schedules, since they require generating high profits by entering into many contracts and by minimizing the penalty charge for delays of delivery times. We proposed a concurrent scheduling algorithm in the previous research [6]. Both a supplier and a manufacturer can improve their production schedules concurrently, and keep the consistency among their production schedules in the negotiation process among the three-layered components.

Figure 3 shows a concurrent scheduling process between a manufacturer and a supplier. The vertical axis shows the time of scheduling process.

1. A client generates an order and sends it to all manufacturers.
2. After receiving the order, a manufacturer improves its production schedule for a short time by using the GA. The manufacturer generates an order and sends it to all suppliers. After that, the manufacturer continuously improves the production schedule in consideration of the constraint for the earliest starting time of manufacturing process of the newly ordered product.

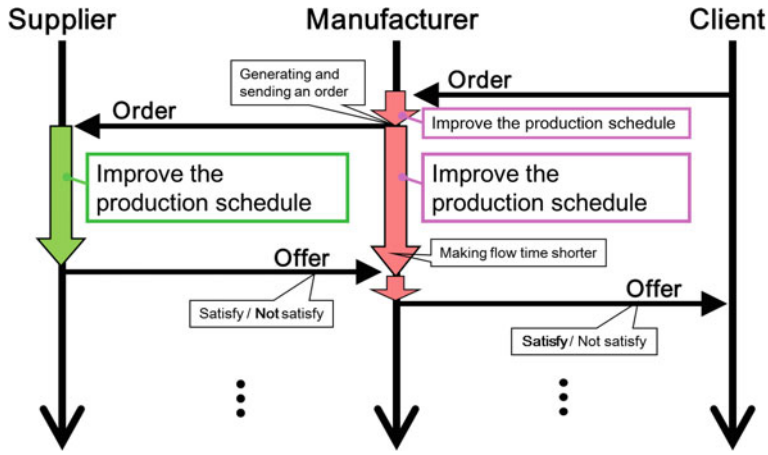


Fig. 3 Negotiation process among three model components

3. A supplier improves its production schedule until the bidding time by using the GA. The supplier generates an offer, and sends it to the manufacturer.
4. According to the possible delivery time described in the offer, the manufacturer modifies the earliest starting time of manufacturing process in its production schedule and improves the production schedule quickly. Then, the manufacturer generates an offer and sends it to the client.

This negotiation process is repeated among the three model components, until the client accepts an offer sent from a manufacturer or cancels the order. If the client accepts an offer sent from a manufacturer, the manufacturer can enter into a contract with the supplier.

3.3 Computational Experiments

We developed a prototype of supply chain simulation system for three-layered supply chains, and evaluated the effectiveness of the proposed model and concurrent scheduling algorithm. Several Windows-based networked computers (Intel Core 2 Duo 3 GHz CPU with 2 GB of RAM) were used to develop the prototype system for ease of applicability to real-world implementation. We implemented all model components; those are clients, two manufacturers, and two suppliers, as agents on different computers, as shown in Fig. 4.

The manufacturers in the simulation system were able to send their orders to both the two suppliers simultaneously. The suppliers had two schedulers not only for receiving the orders from the two manufacturers at the same time, but also for

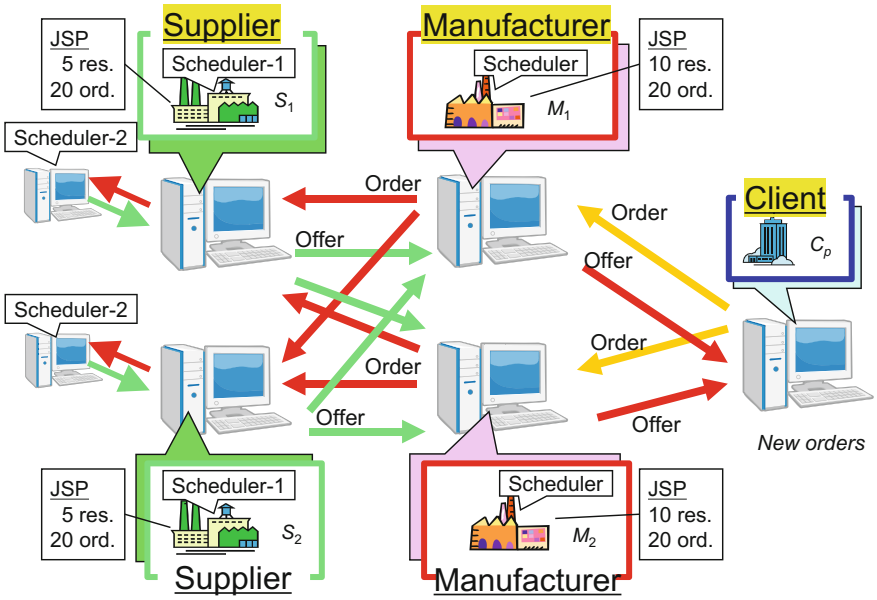


Fig. 4 Three-layered supply chain simulation system

evaluating the most profitable order of all through the rescheduling processes for the additional orders. The effectiveness of the continuous improvement of the production schedules was verified in the three-layered dynamic supply chains from the viewpoint of the profits of both the suppliers and the manufacturers. Additionally, the effectiveness was also evaluated from the viewpoint of the benefits of clients. Both the manufactures and the suppliers were able to enter into a lot of contracts and obtain large profits by improving production schedules of the manufacturers continuously after sending the orders. The clients obtained the ordered products at lower prices, in the case where the manufacturers continued to improve their production schedules.

4 Conclusion

This paper reviews the existing dynamic supply chain models. A two-layered dynamic supply chain model was firstly described as a basic supply chain model. Then, a three-layered dynamic supply chain model was described as a minimum model for the multi-layered dynamic supply chains. The effectiveness of the negotiation protocol was verified through computational experiments by using a supply chain simulation system. The experimental results showed that a concurrent scheduling algorithm was effective in providing not only the profits of suppliers and manufacturers, but also the benefits of clients.

Acknowledgements This research is supported by the Grants-in-Aid for Scientific Research (C) (No. 15K01196) of Japan Society for the Promotion of Science (JSPS).

References

1. T. Kaihara, S. Fujii, in *IT Based Virtual Enterprise Coalition Strategy in Agile Manufacturing Environment*. Proc. of 35th CIRP International Seminar on Manufacturing Systems (2002) pp. 333–338
2. S. Piramuthu, Knowledge-based framework for automated dynamic supply chain configuration. Eur. J. Oper. Res. **165**, 219–230 (2005)
3. D. Emerson, S. Piramuthu, in *Agent-Based Framework for Dynamic Supply Chain Configuration*. Proc. of 37th Hawaii International Conference on System Science, 70168a (2004)
4. Y. Tanimizu, M. Yamanaka, K. Iwamura, N. Sugimura, in *Multi-Agent Based Dynamic Supply Chain Configuration Considering Production Schedules*. Proceedings of 2006 International Symposium on Flexible Automation (2006) pp. 572–578
5. Y. Tanimizu, C. Ozawa, M. Yamanaka, K. Iwamura, N. Sugimura, *Credibility of Supplier in Dynamic Supply Chain*. Proceedings of the 40th CIRP International Seminar on Manufacturing Systems, CD-ROM (2007)
6. Y. Tanimizu, C. Ozawa, Y. Shimizu, B. Orita, K. Iwamura, N. Sugimura, Flexible multi-layered dynamic supply chain models with cooperative negotiation. Int. J. Autom. Tech. **7**(1), 128–135 (2013)

Effects of Sample Tilt on Vickers Indentation Hardness

Ming Liu, Guangyu Zhu, Xiangyu Dong, Jinming Liao and Chenghui Gao

Abstract A theoretical approach is proposed to characterize the sample tilt by the tilt angle and the rotation angle, which can be calculated by the residual imprint morpholo-gy of Vickers indenter that is ideal with its axis along the vertical direction as well as the loading direction. Unique solutions of the three components of the local normal can be obtained by solving the equations with the projected contact lengths of Vickers indenter edges measured from the residual imprint after Vickers indentation. The tilt angle and the rotation angle can be calculated from compo-nents of the normal defined in Cartesian coordinate system. It is found both the tilt angle and the measured hardness have a Gaussian distribution, while the rota-tion angle has a random distribution. The measured hardness is found to lin-early increase with the tilt angle, while independent of the rotation angle.

Keywords Sample tilt · Tilt angle · Rotation angle · Vickers indentation Residual imprint mor-phology · Indentation hardness

1 Introduction

Contact mechanics has been drawing much attention from a wide range of subjects [1] with the widely-used instrumented indentation for replicating contact damage [2], modelling adhesive contact [3], measuring residual stresses [4–7], and characterizing grain boundary [8], interface [9–11], films [12], fracture behaviour [13], fracture toughness [14], and material properties (e.g. hardness [15, 16], yield stress [17], strain hardening [18], viscoelasticity of soft biomaterials [19], creep [20, 21], superelasticity [22], and piezoelectric constants [23]). As indenta-tion technique has become ubiquitous for testing material (e.g. thermosetting pol-ymers [24], composite material [25], nanoscratch test [26]), measurement accu-racy requires

M. Liu (✉) · G. Zhu · X. Dong · J. Liao · C. Gao
School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350116, People's Republic of China
e-mail: mingliu@fzu.edu.cn

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_28

understanding of effects of many factors (e.g. surface topography and roughness [27–32], high plastic strain in the surface region introduced by the polishing process [33], and indenters of imperfect geometries [34]), which can cause large scatter in the measured values [31] and erroneous interpretation of in-dentation results (e.g. artificial length scale [35]).

Most studies are focused on normal contact, in which the loading direc-tion is normal to the surface [36], such as contact between two flat surfaces with surface roughness [37], and contact between a curved surface and a flat [38, 39]. However, the ideal condition that a flat surface is mounted perpendicular to the indenter axis is difficult to be met in indentation experiment, since the top and the bottom surfaces of the sample are not parallel to some extent, and obtaining a flat surface can be very challenging: polishing process with water can alter micro-structure and properties of the coating that is sensitive to moisture; locally flat surface in binary materials is difficult to obtain due to inclusions; an embedded phase can be ploughed out, the ratio of elements at the surface can be changed, and a near-surface deformed layer can be generated during mechanical polishing [40], which can impact microstructures and properties. Inclined contact has re-cently attracted much attention, especially in indentation, in order to quantity ef-fects of sample tilt as well as indenter tilt (e.g. erratic behaviour in unloading curves, unreliable measurements of hardness, and reduced modulus [41]).

Effects of sample tilt on indentation responses were studied theoretically with the derived projected contact area for tilt samples [42]. Both finite element simulation and experiments (e.g. conical nanoindentation [43] and Berkovich indentation [44]) were employed to study inclined indentation. Zhao's group [45] theoretically derived the projected normal contact area as well as the tangential contact area in different tilting cases, and investigated effects of tilting probe on nanoscratch using finite element method. Their results showed influences of the tilt angle on residual scratch profile, projected contact area, and calculated frictional coefficient. Wei et al. [46] carried out Berkovich nanoindentation of a single crystalline diamond, and found anomalous indentation curves for tilt samples. Hu et al. [47] measured material elastic modulus and yield strength by indentation with a cylindrical flat-tip indenter, and found sample's tilt angle measured from horizontal plane could be as large as 1°. Laurent-Brocq et al. [48] experimentally studied influences of sample tilt on Berkovich nanoindentation. Their experimental results showed normal analysis of the load-displacement curve of a tilt sample underestimated the contact area, and thus overestimated the hardness.

Sample tilt can be characterized from the coordinates of three points on the sample surface [49] or by using the residual indent morphology [50]. Surface morphology and surface roughness [51, 52] provide much insight into understanding many fundamental problems (e.g. morphological change caused by thermal annealing [12], post-polymerization of the adhesive during heating [24], bonding [53], and optoelectronic properties affected by surface morphology [52]). Special methods for correcting profile deviation due to stylus tip and eliminating blind regions are needed for profile measurement of steep surface [54]. Ramzan et al. [55] used atomic force microscopy (AFM) to study the morphology of thin

films, and found a larger surface roughness under a smaller grain size associated with a larger grain boundary area. Saravana et al. [56] studied ohmic contact by AFM scan of contact surface, and found the contact resistance varied with surface morphology. Renner et al. [57] found the residual imprint topographies (e.g. pile-up profiles) of Berkovich nanoindentation contained much information on intrinsic material properties at the grain scale.

Perfect alignment between indenter axis (or loading direction) and sample surface normal is difficult to ensure in experiment. In order to take into account effects of sample tilt on measurement, correction factors as functions of the normal of the sample top surface should be known. A theoretical approach is presented, in this paper, to calculate the surface normal of a tilted sample as well as the tilt and the rotation angles defined in Cartesian coordinate system at the tip of Vickers indenter. Unique solutions of the three components of the local surface normal can be obtained by solving the equations containing the projected contact lengths of the residual imprint morphology after Vickers indentation. Effects of the tilt angle and the rotation angle on Vickers indentation hardness are quantified experimentally.

2 Theoretical Approach of Calculating the Tilt Angle and the Rotation Angle

Consider the Cartesian coordinate with the origin being at the tip of the Vickers indenter, the indenter axis being z-axis (i.e. the vertical direction as well as the loading direction), and the projection of one indenter edge on the horizontal plane (x-o-y plane) being y-axis (see Fig. 1). The unit edge vectors of the Vickers indenter are

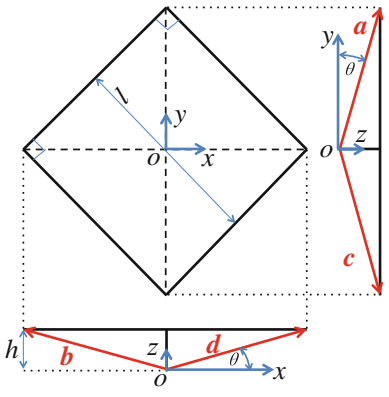


Fig. 1 The projected views of Vickers indenter with geometric parameters ($l = 2h\tan 68°$, $\theta = 15.944°$) in the Cartesian coordinate system (***a***, ***b***, ***c***, and ***d*** are the edge vectors)

$$\begin{bmatrix} \hat{\boldsymbol{a}} \\ \hat{\boldsymbol{b}} \\ \hat{\boldsymbol{c}} \\ \hat{\boldsymbol{d}} \end{bmatrix} = \begin{bmatrix} 0 & \cos\theta & \sin\theta \\ -\cos\theta & 0 & \sin\theta \\ 0 & -\cos\theta & \sin\theta \\ \cos\theta & 0 & \sin\theta \end{bmatrix} \tag{1}$$

Sample tilt is characterized by the unit normal vector (n_x,n_y,n_z) of the sample top surface, and

$$n_x^2 + n_y^2 + n_z^2 = 1,\ n_z > 0 \tag{2}$$

The constraint $n_z > 0$ is because the normal of the top surface is pointing outwards the sample in the established coordinate system (see Figs. 1 and 2). The tilt angle, α, between sample plane and horizontal plane, whose normal is (0,0,1), can be calculated as

$$\cos\alpha = n_z \tag{3}$$

The absolute rotation angle $\xi'(0 \le \xi' < 360^\circ)$ is defined as the angle between the y-axis (i.e. the projection of one indenter edge on x-o-y plane) and the projection of the surface normal on x-o-y plane $\boldsymbol{n}'$ (see Fig. 2)

$$\frac{\pi}{180^\circ}\xi' = \begin{cases} \arctan\left(\frac{n_x}{n_y}\right) & \text{for } n_x \ge 0 \text{ and } n_y > 0 \\ \pi + \arctan\left(\frac{n_x}{n_y}\right) & \text{for } n_y < 0 \\ 2\pi + \arctan\left(\frac{n_x}{n_y}\right) & \text{for } n_x < 0 \text{ and } n_y > 0 \\ 0 & \text{for } n_x = 0 \text{ and } n_y = 0 \\ \pi/2 & \text{for } n_x > 0 \text{ and } n_y = 0 \\ 3\pi/2 & \text{for } n_x < 0 \text{ and } n_y = 0 \end{cases} \tag{4}$$

Considering the geometrical symmetry of the Vickers indenter, different absolute rotation angles $\xi'(0 \le \xi' \le 360^\circ)$ under the same tilt angle can be of equivalence,

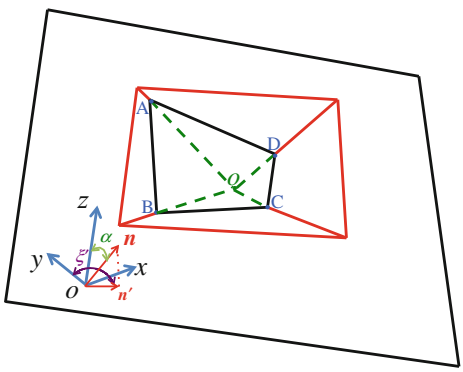


Fig. 2 The geometrical intersection between a tilted plane with the normal $\boldsymbol{n}$ and a Vickers indenter: A, B, C, and D are the intersect points between sample surface and indenter edges, $\boldsymbol{n}'$ is the projection of normal $\boldsymbol{n}$ on x-o-y plane, α is the angle between surface normal $\boldsymbol{n}$ and vertical direction z-axis, and ξ' is the angle between $\boldsymbol{n}'$ and y-axis on x-o-y plane, $0 \le \xi' < 360^\circ$

and a unified rotation angle ξ $(0 \leq \xi \leq 45^\circ)$ is better used for clarity. Each indenter edge vector has its own projection on the horizontal *x*-*o*-*y* plane, and the angles between the four projections and $\boldsymbol{n}'$ are not independent. Due to the symmetry of the Vickers indenter, the relative position in space between the Vickers indenter and the tilted sample can be characterized by two angles: one is the tilt angle α; the other is the angle between $\boldsymbol{n}'$ and the projection of just one edge vector on the horizontal plane. The rotation angle, ξ, is defined as the smallest angle among the four angles between $\boldsymbol{n}'$ and the four projections of edge vectors on the horizontal plane, resulting in the range of ξ from 0 to 45°, and the relationship between ξ and ξ' is expressed as

$$\xi = \begin{cases} \xi' & \text{for } 0 \leq \xi' \leq 45^\circ \\ 90^\circ - \xi' & \text{for } 45^\circ < \xi' \leq 90^\circ \\ \xi' - 90^\circ & \text{for } 90^\circ < \xi' \leq 135^\circ \\ 180^\circ - \xi' & \text{for } 135^\circ < \xi' \leq 180^\circ \\ \xi' - 180^\circ & \text{for } 180^\circ < \xi' \leq 225^\circ \\ 270^\circ - \xi' & \text{for } 225^\circ < \xi' \leq 270^\circ \\ \xi' - 270^\circ & \text{for } 270^\circ < \xi' \leq 315^\circ \\ 360^\circ - \xi' & \text{for } 315^\circ < \xi' < 360^\circ \end{cases} \tag{5}$$

The contact depth, h, is defined as the distance from the indenter tip at the origin (0,0,0) to the intersection (0,0,h) between the indenter axis and the sample surface. Therefore, the plane of the tilted sample can be expressed in point-normal form

$$n_x x + n_y y + n_z (z - h) = 0 \tag{6}$$

Assume the contact boundary on the top surface lies in the same plane with the same normal as that of the tilted sample surface. The intersection points between surface plane and indenter edges can be calculated by purely geometrical consideration (see Fig. 2), and the coordinates of intersect points are

$$\begin{pmatrix} \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \\ \mathrm{D} \end{pmatrix} = \begin{bmatrix} 0 & t_\mathrm{A} \cos\theta & t_\mathrm{A} \sin\theta \\ -t_\mathrm{B} \cos\theta & 0 & t_\mathrm{B} \sin\theta \\ 0 & -t_\mathrm{C} \cos\theta & t_\mathrm{C} \sin\theta \\ t_\mathrm{D} \cos\theta & 0 & t_\mathrm{D} \sin\theta \end{bmatrix} \tag{7}$$

where

$$\begin{pmatrix} t_\mathrm{A} \\ t_\mathrm{B} \\ t_\mathrm{C} \\ t_\mathrm{D} \end{pmatrix} = \begin{pmatrix} \frac{hn_z}{n_y \cos\theta + n_z \sin\theta} \\ \frac{hn_z}{-n_x \cos\theta + n_z \sin\theta} \\ \frac{hn_z}{-n_y \cos\theta + n_z \sin\theta} \\ \frac{hn_z}{n_x \cos\theta + n_z \sin\theta} \end{pmatrix} \tag{8}$$

The projected contact lengths between each Vickers tip edges and the surface plane (see Fig. 3) can be obtained as

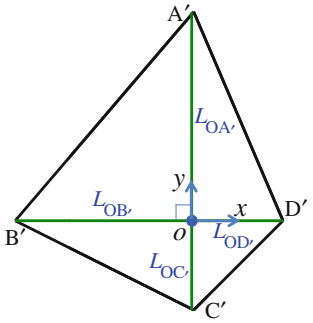


Fig. 3 The projection of contact boundary with projected contact lengths between Vickers indenter tip edges and sample surface on the horizontal plane

$$\begin{pmatrix} L_{OA'} \\ L_{OB'} \\ L_{OC'} \\ L_{OD'} \end{pmatrix} = \cos\theta \begin{pmatrix} t_A \\ t_B \\ t_C \\ t_D \end{pmatrix} \tag{9}$$

From Eqs. (8) and (9), two dimensionless independent equations can be obtained

$$\begin{cases} \frac{2n_z \tan\theta}{n_y + n_z \tan\theta} = \frac{L_{OA'}}{L_{OB'}} + \frac{L_{OA'}}{L_{OD'}} \\ \frac{2n_z \tan\theta}{n_x + n_z \tan\theta} = \frac{L_{OD'}}{L_{OA'}} + \frac{L_{OD'}}{L_{OC'}} \end{cases} \tag{10}$$

With Eq. (2) that is also dimensionless, three independent equations with three unknowns $(\boldsymbol{n}_x, \boldsymbol{n}_y, \boldsymbol{n}_z)$ can be obtained and solved numerically (e.g. by MATLAB software).

The tilt angle α can be calculated from Eq. (3); and rotation angle ξ can be calculated from Eqs. (4) and (5).

3 Results and Discussion

3.1 Experimental Measurement of Sample Tilt

Vickers micro-hardness test is performed on a copper cylinder with 1.5 cm in both height and diameter under normal load of 0.5 kgf and holding time of 10 s. The distance between two indentation tests is 0.4 mm, which is large enough to avoid interaction between neighbouring imprints. The top surface of the copper sample was polished with alumina particles of 7 μm in diameter after being grinded with SiC sandpaper. Figure 4 shows the residual imprint of the Vickers indenter on the sample surface with the projected contact lengths between the four edges of the Vickers indenter and the sample, which can be measured by the optical microscope with high magnification and accuracy. Only the ratios of contact lengths are of relevance if dimensionless equations [Eqs. (2) and (10)] are used. With the length

of each indenter edge projected on *x-o-y* plane, the three independent equations [Eqs. (2) and (10)] with three unknowns can be solved numerically by MATLAB software. The tilt angle and the rotation angle are both zero for the condition that the three projected contact lengths of indenter edges form an equilateral triangle. Figure 5 shows MATLAB code for calculating the surface normal with the tilt angle and the rotation angle (the input lengths are from Fig. 4b). The calculation process of the tilt angle and the rotation angle is summarized is shown in Fig. 6. The solutions are unique under $n_z \geq 0$, since the normal of the surface points upwards (see Fig. 2).

Figure 7a shows the normalized probability distribution of the tilt angle for more than 300 data. The sample surface is not perfectly flat, and the local normal varies with locations. The normalized probability density function (PDF) can be fitted with two Gaussian functions as

$$\mathrm{PDF} = \frac{f_1}{\sigma_1\sqrt{2\pi}}\exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right) + \frac{f_2}{\sigma_2\sqrt{2\pi}}\exp\left(-\frac{(x-\mu_2)^2}{2\sigma_2^2}\right) \qquad (11)$$

where f_1 and f_2 represent the fraction of each Gaussian function, and $f_1 + f_2 = 1$. The two different Gaussian distributions of the tilt angle indicate that the sample surface consists of two different local regions with different surface roughness that can influence contact lengths of the indenter edges, bringing about different tilt angles. The tilt angle lies in the range from 0° to 6°. Figure 7b shows the normalized probability density distribution of the rotation angle. The rotation angle has a random distribution with a uniform probability from 0° to 45°.

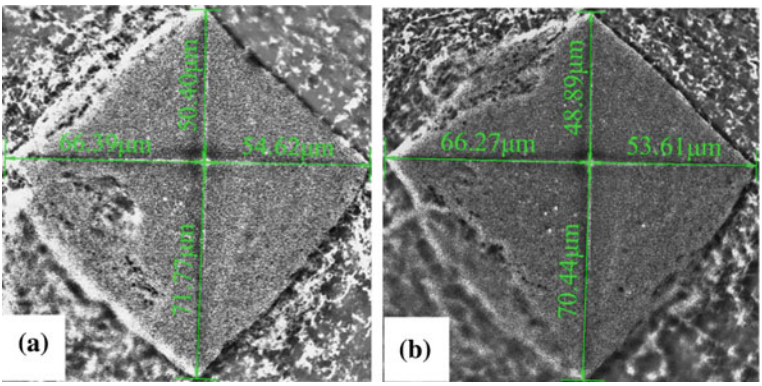


Fig. 4 Projected contact lengths of the four indenter edges measured from the residual imprint of Vickers indenter: (**a**) surface normal is (0.0134, 0.054, 0.998) with a tilt angle of 3.2° and a rotation angle of 14.5°; (**b**) surface normal is (0.014, 0.06, 0.998) with a tilt angle of 3.56° and a rotation angle of 13.1°

Fig. 5 Snapshot of MATLAB code for calculating the tilt angle and the rotation angle of a tilted sample (the input values are from Fig. 4b)

```
a=48.89;                    % OA' measured
b=66.27;                    % OB' measured
c=70.44;                    % OC' measured
d=53.61;                    % OD' measured
A=tan(15.944*pi/180);  % Edge angle
syms x                      % x of normal
syms y                      % y of normal
syms z positive             % z of normal
[x,y,z]=solve(2*z*A==(a/b+a/d)*(y+z*A),...
              2*z*A==(d/a+d/b)*(x+z*A),...
               x^2+y^2+z^2==1);
X=[x,y,z];
Normal=vpa(X,6)
TA=acos(z)*180/pi;
if x>=0 & y>0
   AbsRota=180/pi*atan(x/y);
elseif y<0
   AbsRota=180/pi*(pi+atan(x/y));
elseif x<0 & y>0
   AbsRota=180/pi*(2*pi+atan(x/y));
elseif x==0 & y==0
   AbsRota=0;
elseif x>0 & y==0
   AbsRota=180/pi*(pi/2);
else x<0 & y==0
   AbsRota=180/pi*(3*pi/2);
end
TiltAngle=vpa(TA,6)
AAR=vpa(AbsRota,6);
if AAR>=0 & AAR<=45
   RA=AAR;
elseif AAR>45 & AAR<=90
   RA=90-AAR;
elseif AAR>90 & AAR<=135
   RA=AAR-90;
elseif AAR>135 & AAR<=180
   RA=180-AAR;
elseif AAR>180 & AAR<=225
   RA=AAR-180;
elseif AAR>225 & AAR<=270
   RA=270-AAR;
elseif AAR>270 & AAR<=315
   RA=AAR-270;
elseif AAR>315 & AAR<360
   RA=360-AAR;
end
RotationAngle=vpa(RA,3)
```

[49]

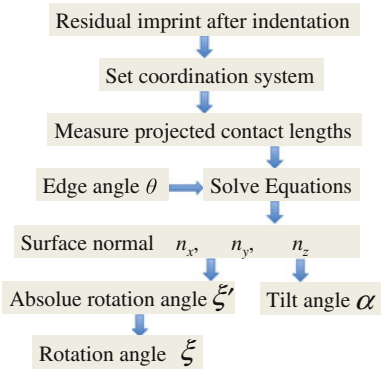


Fig. 6 Flowchart of the calculation process for the normal of a tilted surface as well as the tilt and the rotation angles from the residual imprint of Vickers indentation

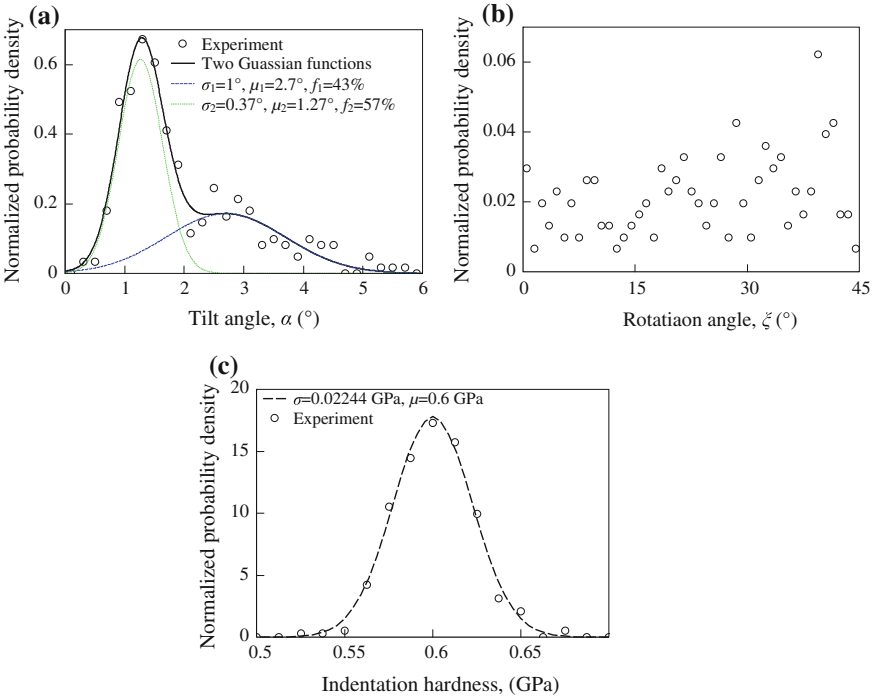


Fig. 7 Normalized probability density distribution of (**a**) the tilt angle; and (**b**) the rotation angle; (**c**) Vickers indentation hardness

3.2 Vickers Indentation Hardness

Figure 7c shows the normalized probability density distribution of Vickers indentation hardness, H_i, which is defined as the ratio of the normal load, P, to the projected contact area

$$H_i = \frac{P}{A} \tag{12}$$

where the projected contact area, A, is computed from the projected contact lengths as

$$A = \frac{(L_{OA'} + L_{OC'})(L_{OB'} + L_{OD'})}{2} \tag{13}$$

The normalized probability density distribution of Vickers indentation hardness can be fitted well with Gaussian function

$$\text{PDF} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{14}$$

The tilt angle should be smaller than 2° in order to limit the influence of sample tilt within 10% of the true hardness [48]. Since the measured tilt angle can be as large as 5°, the variation of Vickers indentation hardness is expected to be larger than 10%. Vickers indentation hardness of copper lies within the range from 0.55 to 0.65 GPa with the mean value of 0.6 GPa. The tilt angle has two different Gaussian distribution, while the indentation hardness has only one Gaussian distribution. It is worth mentioning that the difference between the projected contact lengths of the indenter edges can be caused by many factors: residual stresses in preparing specimens, slope of workbench, material anisotropy and different phases, pores and height difference in the phase boundary [58], imperfect geometry of the indenter due to fabrication of the indenter that is made of anisotropic diamond crystal [59], and misalignment of the indenter rather than the sample, to name a few. It is assumed that the copper used in Vickers test is isotropic; the indenter is ideal, and its axis is vertical and along the loading direction. Although the sample surface experiences the same grinding and polishing processes, different regions can exhibit significantly different surface topography, which is evidenced by the use of two Gaussian functions to fit the probability density distribution of the tilt angle. It is the effects of many factors that make the probability distribution of calculated hardness follow Gaussian distribution.

Figure 8a and b show the variation of Vickers indentation hardness with the tilt angle and the rotation angle, respectively. Figure 8c and d display the variation of Vickers indentation hardness with both the tilt angle and the rotation angle in 3D and contour plots, respectively. It is found that the rotation angle has little effect on

the measured hardness, while indentation hardness increases with the tilt angle in an approximately linear way. The variation of hardness measured in experiment is partly attributed to the tilt angle that is the angle between the normal of the local surface and the vertical direction. Other factors like surface roughness and residual stress make the measured hardness scattered around the linear dependence of hardness on the tilt angle. The mean value of hardness corresponds to the case with the most probability of the tilt angle, and does not correspond to the true hardness, since zero tilt angle hardly exists in reality. Nevertheless, since the tilt angle is small with the mean value about 2°, the mean value of calculated hardness is only a little larger than the true hardness. Microscopic pores, cracks, and shallow voids can result in a reduction of measured hardness, and counteracting effects may make the mean value of measured hardness equal material true hardness. The deviation of measured value from true value can be as large as 10% due to the tilt of local region, and a fairly large number of tests are needed to capture Gaussian distribution of indentation hardness. The micro-Vickers indentation hardness of copper used in this study is found to be 0.6 GPa, which accords well with the reported value of the sample made from copper powders by plasma pressure compaction [60].

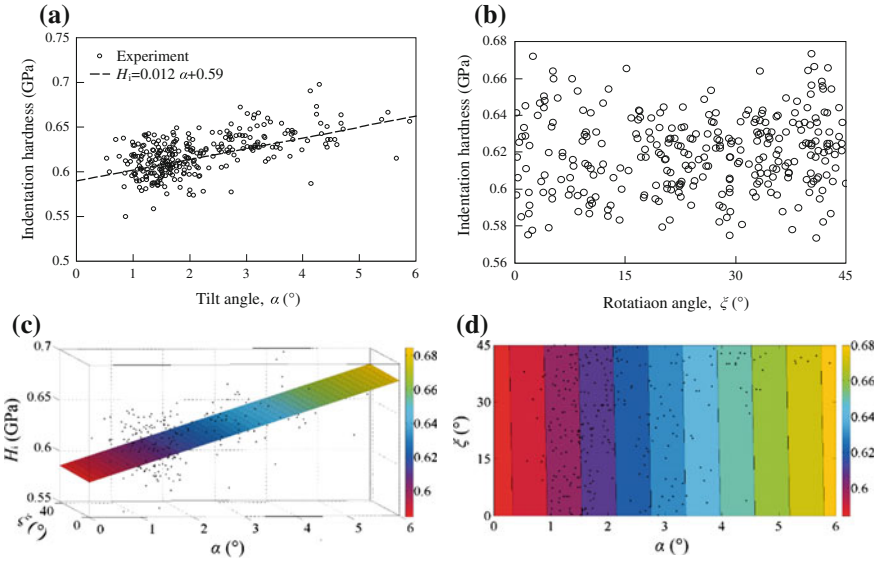


Fig. 8 Variation of Vickers indentation hardness (unit: GPa) due to (**a**) the tilt angle; (**b**) the rotation angle; and both angles: (**c**) 3D plot; (**d**) contour plot

4 Conclusion

Sample tilt is normally inevitable owing to many factors (e.g. sample preparation, manufacturing and assembly processes). A theoretical approach is proposed to measure the normal of sample surface with the tilt angle and the rotation angle by the residual imprint morphology after Vickers indentation. The normal of the local region varies with locations with a random distribution of the rotation angle and Gaussian distribution of the tilt angle. Spatial variability still exists for the polished sample with mirror finish, which is evidenced by multiple Gaussian functions to fit the probability distribution of the tilt angle. The variation of calculated hardness follows Gaussian distribution. With the tilt information between the sample and the indenter, effects of sample tilt on indentation test can be quantified. Experimental results from Vickers indentation of copper show the rotation angle has little effect on the measured indentation hardness, while Vickers indentation hardness increases with the tilt angle in a linear way.

Acknowledgements This project is supported by National Natural Science Foundation of China (Grant No. 51705082). M. Liu is supported by Qishan Scholar program and Minjiang Scholar program in Fuzhou University (No. 0020-650289 and No. 0020-510486). This work was supported by Fujian Provincial Collaborative Innovation Center for High-end Equipment Manufacturing (No. 0020-50006103).

References

1. M. Liu, F. Yang, Modell. Simul. Mater. Sci. Eng. **20**, (2012)
2. B.D. Beake, T.W. Liskiewicz, J.F. Smith, Surf. Coat. Technol. **206**, 1921–1926 (2011)
3. L. Li, W. Song, M. Xu, A. Ovcharenko, G. Zhang, Comput. Mater. Sci. **98**, 105–111 (2015)
4. Z. Lu, Y. Feng, G. Peng, R. Yang, Y. Huan, T. Zhang, Mater. Sci. Eng., A **614**, 264–272 (2014)
5. Y.H. Lee, W.J. Ji, D. Kwon, Exp Mech. **44**, 55–61 (2004)
6. P.-L. Larsson, J. Mater. Process. Technol. **184**, 372–378 (2007)
7. F. Rickhey, J.H. Lee, H. Lee, Mater. Des. **84**, 300–312 (2015)
8. S. Pathak, J. Michler, K. Wasmer, S.R. Kalidindi, J Mater. Sci. **47**, 815–823 (2011)
9. W.-S. Song, S.-G. Kim, Y.-C. Kim, D. Kwon, J. Electron. Mater. **44**, 831–835 (2015)
10. M. Liu, F. Yang, J. Comput. Theor. Nanosci. **11**, 1863–1873 (2014)
11. G. Xiao, X. Yang, G. Yuan, Z. Li, X. Shu, Mater. Des. **88**, 520–527 (2015)
12. D. Lian, P.L. Lin, Microelectron. Reliab. **56**, 66–72 (2016)
13. K. Fu, L. Chang, B. Zheng, Y. Tang, Y. Yin, Vacuum **112**, 29–32 (2015)
14. J. Li, F. Li, M. He, F. Xue, M. Zhang, C. Wang, Mater. Des. **40**, 176–184 (2012)
15. E.S. Gouda, A. Nassar, Eur. Phys. J. Appl. Phys. **68**, 20701 (2014)
16. K.K. Jha, N. Suksawang, A. Agarwal, Comput. Mater. Sci. **85**, 32–37 (2014)
17. A. Clausner, F. Richter, Eur. J. Mech. A. Solids **53**, 294–302 (2015)
18. Q.Q. Huang, Y. Song, W.W. Liu, Y. Chen, F. Qi, D. Zhao, Y.G. Wang, Ceram. Int. **41**, 12349–12354 (2015)
19. G. Mattei, G. Gruca, N. Rijnveld, A. Ahluwalia, J. Mech. Behav. Biomed. Mater. **50**, 150–159 (2015)
20. F. Rickhey, J.H. Lee, H. Lee, J. Mater. Res. **30**, 3542–3552 (2015)

21. R. Mahmudi, M. Shalbafi, M. Karami, A.R. Geranmayeh, Mater. Des. **75**, 184–190 (2015)
22. P. Li, H.E. Karaca, Y.-T. Cheng, J. Alloy. Compd. **651**, 724–730 (2015)
23. M. Liu, F.Q. Yang, Modell. Simul. Mater. Sci. Eng. **21**, (2013)
24. X. Wang, Y. Li, S. Wang, Y. Deng, D. Xing, S. He, Eur. Polymer J. **70**, 360–370 (2015)
25. Z. Yuan, F. Li, F. Xue, M. Zhang, J. Li, J. Mater. Eng. Perform. **24**, 654–663 (2014)
26. F. Zhang, B. Meng, Y. Geng, Y. Zhang, Appl. Surf. Sci. **368**, 449–455 (2016)
27. M. Wolski, P. Podsiadlo, G.W. Stachowiak, Tribology letters, **61**, (2015)
28. M. Qasmi, P. Delobelle, Surf. Coat. Technol. **201**, 1191–1199 (2006)
29. C. Walter, C. Mitterer, Surf. Coat. Technol. **203**, 3286–3290 (2009)
30. J.-Y. Kim, S.-K. Kang, J.-J. Lee, J.-I. Jang, Y.-H. Lee, D. Kwon, Acta Mater. **55**, 3555–3562 (2007)
31. S.W. Wai, G.M. Spinks, H.R. Brown, M. Swain, Polym. Test. **23**, 501–507 (2004)
32. L. Chen, A. Ahadi, J. Zhou, J.-E. Ståhl, Model Numer. Simul. Mater. Sci. **04**, 153–162 (2014)
33. Y. Li, P. Kanouté, M. François, Mater. Sci. Eng., A **642**, 381–390 (2015)
34. W. Chen, M. Li, T. Zhang, Y.-T. Cheng, C.-M. Cheng, Mater. Sci. Eng., A **445–446**, 323–327 (2007)
35. G. Constantinides, E.C.C.M. Silva, G.S. Blackman, K.J.V. Vliet, Nanotechnology **18**, 305503 (2007)
36. M. Liu, F. Yang, Int. J. Solids Struct. **50**, 2542–2547 (2013)
37. P. Todorovic, B. Tadic, D. Vukelic, M. Jeremic, S. Randjelovic, R. Nikolic, Tribol. Int. **81**, 276–282 (2015)
38. M. Liu, H. Proudhon, Mech. Mater. **77**, 125–141 (2014)
39. M. Liu, Int. J. Solids Struct. **51**, 3642–3652 (2014)
40. F. Arjmand, L. Zhang, Surf. Interface Anal. **47**, 1120–1126 (2015)
41. C. Shi, H. Zhao, H. Huang, L. Xu, L. Ren, M. Bai, J. Li, X. Hu, Mater. Trans. **54**, 958–963 (2013)
42. M.S. Kashani, V. Madhavan, Acta Mater. **59**, 883–895 (2011)
43. Z.H. Xu, X. Li, Phil. Mag. **87**, 2299–2312 (2007)
44. S. Saber-Samandari, K.A. Gross, J. Eur. Ceram. Soc. **29**, 2461–2467 (2009)
45. C. Shi, H. Zhao, H. Huang, S. Wan, Z. Ma, C. Geng, L. Ren, Tribol. Int. **60**, 64–69 (2013)
46. Q.L. Wei, X.B. Yue, X.Y. Li, B. Liu, X.F. Zhang, D.J. Lei, Adv. Mater. Res. **1120–1121**, 378–382 (2015)
47. Z. Hu, K. Lynne, F. Delfanian, J. Mater. Res. **30**, 578–591 (2015)
48. M. Laurent-Brocq, E. Bejanin, Y. Champion, Scanning **37**, 350–360 (2015)
49. H. Huang, H. Zhao, C. Shi, X. Hu, Y. Tian, Meas. Sci. Technol. **25**, 017001 (2014)
50. H. Huang, H. Zhao, C. Shi, L. Zhang, Meas. Sci. Technol. **24**, 105602 (2013)
51. Y. Fujii, Powder Diffr. **29**, 265–268 (2014)
52. A. Belfedal, D. Benlakehal, Y. Bouizem, R. Baghdad, M. Clin, A. Zeinert, O. Durand-Drouhin, J.D. Sib, L. Chahed, K. Zellama, Mater. Sci. Semicond. Process. **26**, 231–237 (2014)
53. M.G. Subasi, O. Inan, Lasers Med. Sci. **29**, 19–27 (2014)
54. H. Fang, B. Xu, W. Chen, H. Tang, S. Zhao, J. Nanomaterials **2015**, 1–8 (2015)
55. M. Ramzan, E. Ahmed, N.A. Niaz, A.M. Rana, A.S. Bhatti, N.R. Khalid, M.Y. Nadeem, Superlattices Microstruct. **82**, 399–405 (2015)
56. G. SaiSaravanan, K. MahadevaBhat, S. Dhamodaran, A.P. Pathak, R. Muralidharan, H. P. Vyas, D.V. SridharaRao, R. Balamuralikrishnan, K. Muraleedharan, Mater Sci Semicond. Process. **30**, 62–74 (2015)
57. E. Renner, Y. Gaillard, F. Richard, F. Amiot, P. Delobelle, Int. J. Plast. **77**, 118–140 (2016)
58. F. Zhang, B. Meng, Y. Geng, Y. Zhang, Z. Li, Tribol. Int. **97**, 21–30 (2016)
59. W.J. Zong, D. Wu, Z.Q. Li, Mater. Des. **89**, 1057–1070 (2016)
60. T.S. Srivatsan, B. Ravi, A.S. Naruka, L. Riester, S. Yoo, T.S. Sudarshan, Mater. Sci. Eng., A **31**, 22–27 (2001)

Lattice Boltzmann Method for Turbulent Flows

Kazuhiko Suga and Yusuke Kuwata

Abstract An accurate and robust lattice Boltzmann method for calculating turbulent flows by direct and large eddy simulation is developed. For the discrete velocity model, the D3Q27 model is proven to be desirable for turbulent flow simulation and then a multiple relaxation time (MRT) form is derived. To avoid unphysical kinks in turbulent quantities at grid interfaces by local mesh refinement, a correction method is developed. Coupling this imbalance correction (IBC) mesh refinement method with the D3Q27 MRT LBM, large scale turbulence simulation is effectively performed.

1 Introduction

The lattice Boltzmann method (LBM) is free from difficulties of grid generation and convergence of a pressure solver since it usually applies regular cubic grids and does not require a Poisson type of equation for pressure fields. The LBM is based on a discretized Boltzmann equation by an explicit time marching formula and thus it is very much suitable for parallelization with multi GPUs.

However, since the LBM is young compared with the other CFD methods, it still has many important issues to overcome, particularly for turbulence simulation. Therefore, in the present report, we point out some of those issues and show how to overcome them referring to our recent studies.

K. Suga (✉)
Department of Mechanical Engineering, Osaka Prefecture University, Osaka, Japan
e-mail: suga@me.osakafu-u.ac.jp

Y. Kuwata
Department of Mechanical Engineering, Tokyo University of Science, Chiba, Japan
e-mail: a27900@rs.tus.ac.jp

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_29

2 D3Q27 MRT LBM

For three-dimensional flows, the LBM usually applies the D3Q15, D3Q19 or D3Q27 model indicated in Fig. 1. (D3 corresponds to three-dimensional and Qn corresponds to the number of discrete velocity vectors.) If the simulated results are virtually the same, the smaller number of the discrete velocities is desirable due to the computational costs. Indeed, for laminar pipe flows, as seen in Fig. 2, all those discrete velocity models produce ideally symmetrical velocity profiles. However, when the flow becomes turbulent, the D3Q15 and D3Q19 models produce unphysically patterned profiles. Only the D3Q27 model predicts a reasonable velocity profile. The reason why such anomaly occurs was analysed by the authors [1] as follows.

Applying the error analysis method of Holdych et al. [2] to each discrete velocity model, the error term $\mathscr{E}_i$ arising in the momentum equation:

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{1}{\mathrm{Re}} \frac{\partial^2 u_i}{\partial x_j^2} \mathscr{E}_i, \tag{1}$$

was estimated. After projecting the error terms to the cylindrical coordinate, the circumferential components were written with the relaxation time τ of the LBM:

$$\begin{aligned} \mathscr{E}_\theta^{D3Q15} = \Delta^2 \frac{(-6\tau^2 + 6\tau - 1)}{6} \Bigg[\sin 4\theta \Bigg\{ \frac{1}{4} \left(\frac{\partial^3}{\partial r^3} + \frac{3}{r} \frac{\partial^2}{\partial r^2} - \frac{3}{r^2} \frac{\partial}{\partial r} \right) \left(\overline{u_z}^2 + \overline{u_z' u_z'} \right) \\ - \frac{1}{4} \left(- \frac{\partial^3}{\partial r^3} + \frac{4}{r} \frac{\partial^2}{\partial r^2} - \frac{8}{r^2} \frac{\partial}{\partial r} + \frac{4}{r^3} \right) \\ \times \overline{u_r' u_r'} - \frac{1}{4} \left(- \frac{1}{r} \frac{\partial^2}{\partial r^2} + \frac{5}{r^2} \frac{\partial}{\partial r} - \frac{4}{r^3} \right) \overline{u_\theta' u_\theta'} \Bigg\} \Bigg] + O(\Delta^4), \end{aligned} \tag{2}$$

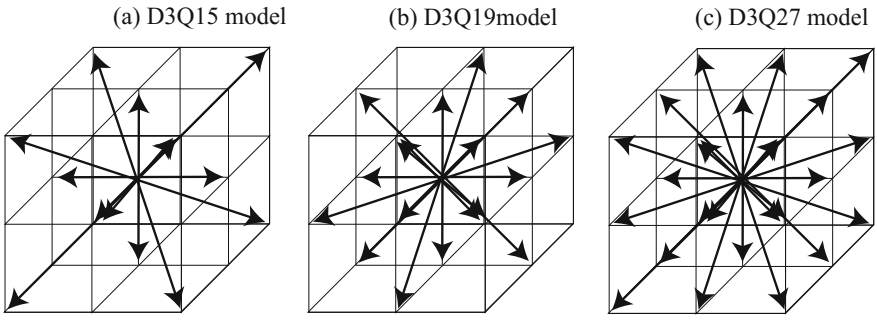


Fig. 1 Discrete velocity models

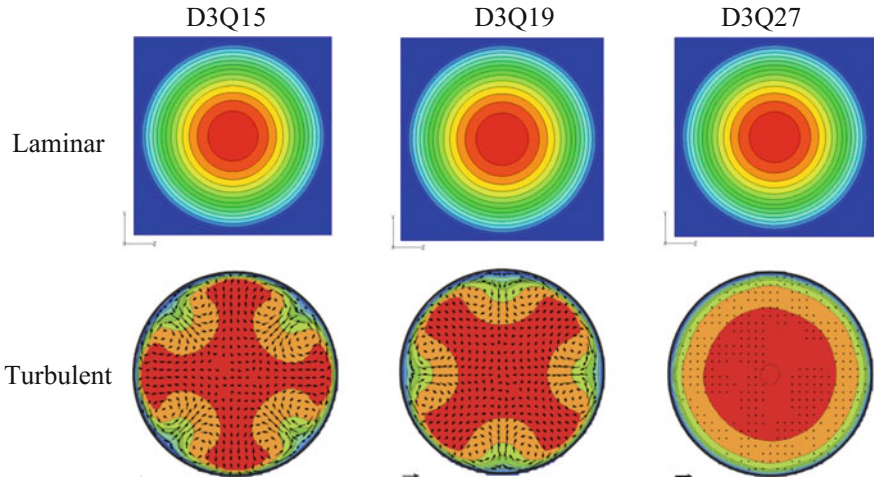


Fig. 2 Streamwise mean velocity contours of pipe flows

$$\mathscr{E}_{\theta}^{D3Q19} = \Delta^2 \frac{(-6\tau^2+6\tau-1)}{6}\left[\sin 4\theta\left\{-\frac{1}{4}\left(\frac{\partial^3}{\partial r^3}+\frac{3}{r}\frac{\partial^2}{\partial r^2}-\frac{3}{r^2}\frac{\partial}{\partial r}\right)\left(\overline{u_z}^2+\overline{u_z'u_z'}\right)\right.\right.$$
$$-\frac{1}{4}\left(-\frac{\partial^3}{\partial r^3}+\frac{4}{r}\frac{\partial^2}{\partial r^2}-\frac{8}{r^2}\frac{\partial}{\partial r}+\frac{4}{r^3}\right)$$
$$\left.\left.\times\overline{u_r'u_r'}-\frac{1}{4}\left(-\frac{1}{r}\frac{\partial^2}{\partial r^2}+\frac{5}{r^2}\frac{\partial}{\partial r}-\frac{4}{r^3}\right)\overline{u_\theta'u_\theta'}\right\}\right]+O(\Delta^4), \tag{3}$$

$$\mathscr{E}_{\theta}^{D3Q27} = \Delta^2 \frac{(-6\tau^2+6\tau-1)}{6}\left[\sin 4\theta\left\{\frac{1}{4}\left(\frac{\partial^3}{\partial r^3}-\frac{4}{r}\frac{\partial^2}{\partial r^2}+\frac{8}{r^2}\frac{\partial}{\partial r}-\frac{4}{r^3}\right)\overline{u_r'u_r'}\right.\right.$$
$$\left.\left.+\frac{1}{4}\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}-\frac{5}{r^2}\frac{\partial}{\partial r}+\frac{4}{r^3}\right)\overline{u_\theta'u_\theta'}\right\}\right]+O(\Delta^4), \tag{4}$$

respectively for the D3Q15, D3Q19 and D3Q27 models. When the pipe flow is laminar, since the Reynolds stresses are negligible, the error term of the D3Q27 model (Eq. 4) vanishes and the total contributions by the parabolic profile of the streamwize mean velocity $\bar{u}_z$ in Eqs. 2 and 3 also vanish. Accordingly, the space accuracy of all those models become $O(\Delta^4)$ or higher. However, in turbulent pipe flows all those contributions by the Reynolds stress and the mean velocity are not negligible and the errors act as asymmetrical force terms. In this sense, even the D3Q27 model is not satisfactory. However, it was found that the magnitude of the error term of the D3Q27 model is two order smaller than those of the other models. Thus, the D3Q27 model is desirable for turbulence simulation.

For high Reynolds number flows, it is known that the multiple-relaxation-time (MRT) form of the LBM is desirable [3].

$$
\begin{aligned}
&|\mathbf{f}(\mathbf{x}+\xi_\alpha\delta t, t+\delta t)\rangle - |\mathbf{f}(\mathbf{x},t)\rangle \\
&= -\mathbf{M}^{-1}\hat{\mathbf{S}}[|\mathbf{m}(\mathbf{x},t)\rangle - |\mathbf{m}^{eq}(\mathbf{x},t)\rangle] + \mathbf{M}^{-1}\left(\mathbf{I}-\frac{\hat{\mathbf{S}}}{2}\right)\mathbf{M}|\mathbf{F}\rangle\delta t,
\end{aligned} \tag{5}
$$

where ξ_α is the discrete velocity and $\mathbf{M}$ is a $Q \times Q$ matrix that transforms the distribution function $\mathbf{f}$ to moments $\mathbf{m}$ as $|\mathbf{m}\rangle = \mathbf{M}\cdot|\mathbf{f}\rangle$. For the D3Q27 model, the present authors' group [4] derived a 27×27 transformation matrix as

$$
\mathbf{M} = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & -1 & 0 & 0 & 0 & 1 & -1 & -1 & 1 & 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 0 \\
0 & 1 & 0 & -1 & 0 & 0 & 1 & 1 & -1 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 0 \\
-1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -2 \\
2 & -1 & 2 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & -2 & 1 & -2 & 1 & -2 & 1 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & 0 & -1 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & -1 & 0 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & -1 & 0 & 1 & 0 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 0 \\
-4 & 0 & 4 & 0 & 0 & 0 & -1 & 1 & 1 & -1 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & 2 & -2 & -2 & 2 & 2 & -2 & -2 & 2 & 0 \\
0 & -4 & 0 & 4 & 0 & 0 & -1 & -1 & 1 & 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & 2 & 2 & -2 & -2 & 2 & 2 & -2 & -2 & 0 \\
0 & 0 & 0 & 0 & -4 & 4 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & -2 & -2 & -2 & -2 & 0 \\
4 & 0 & -4 & 0 & 0 & 0 & -2 & 2 & 2 & -2 & -2 & 0 & 2 & 0 & -2 & 0 & 2 & 0 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 0 \\
0 & 4 & 0 & -4 & 0 & 0 & -2 & -2 & 2 & 2 & 0 & -2 & 0 & 2 & 0 & -2 & 0 & 2 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 0 \\
0 & 0 & 0 & 0 & 4 & -4 & 0 & 0 & 0 & 0 & -2 & -2 & -2 & -2 & 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 4 \\
4 & 4 & 4 & 4 & 4 & 4 & -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -8 \\
-4 & 2 & -4 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 1 & -2 & 1 & -2 & 1 & -2 & 1 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -2 & 0 & -2 & 2 & 2 & 1 & 1 & 1 & 1 & -1 & 0 & -1 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -2 & 2 & -2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 2 & 0 & 2 & 0 & -2 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 2 & 0 & 2 & 0 & -2 & 0 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 0
\end{pmatrix}. \tag{6}
$$

The diagonal collision matrix $\hat{\mathbf{S}}$ is:

$$
\begin{aligned}
\hat{\mathbf{S}} \equiv\ & diag(0,0,0,0,s_4,s_5,s_5,s_7,s_7,s_7,s_{10},s_{10},s_{10},s_{13}, \\
& s_{13},s_{13},s_{16},s_{17},s_{18},s_{20},s_{20},s_{20},s_{23},s_{23},s_{23},s_{26}),
\end{aligned} \tag{7}
$$

and the proposed relaxation parameters [4] are

$$
\begin{aligned}
&s_4 = 1.54,\ s_7 = s_5,\ s_{10} = 1.5,\ s_{13} = 1.83,\ s_{16} = 1.4, \\
&s_{17} = 1.61,\ s_{18} = s_{20} = 1.98,\ s_{23} = s_{26} = 1.74.
\end{aligned} \tag{8}
$$

The relaxation parameter s_5 is related to the fluid viscosity, which is sum of the molecular viscosity v and the sub-grid-scale viscosity v_{SGS} for the LES, $v + v_{SGS} = c_s^2\left(\frac{1}{s_5}-\frac{1}{2}\right)\delta t$, where c_s is the sound speed of the LBM.

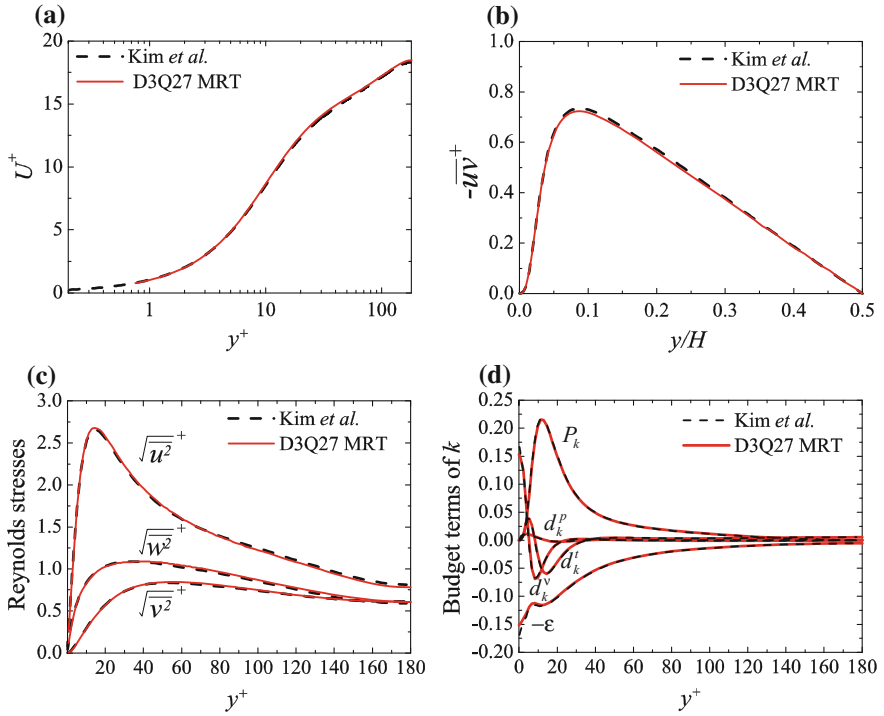


Fig. 3 Turbulent channel flow quantities at $\mathrm{Re}_\tau = 180$: **a** mean velocity, **b** Reynolds shear stress, **c** Reynolds normal stresses, **d** budget terms of the turbulent kinetic energy equation

The accuracy of the D3Q27 MRT LBM for turbulent channel flow DNS is confirmed in Fig. 3. DNS was performed by regular grids of $1539 \times 240 \times 754$ for the region of $4\pi\delta \times 2\delta \times 2\pi\delta$ with the resolution of $\Delta^+ = 1.5$. Indeed, the identical data to those by Kim et al. [5] were obtained.

3 IBC Local Mesh Refinement

To resolve boundary layers, local mesh refinement is essential for the method employing regular grids. In the context of the LBM, the widely applied local mesh refinement applies the communication rule between the distribution functions of course and fine grids: $f_\alpha^{eq,c} \Leftrightarrow f_\alpha^{eq,f}, f_\alpha^{neq,c} \Leftrightarrow n\frac{\tau^c}{\tau^f} f_\alpha^{neq,f}$, with $n = \Delta^c/\Delta^f = \delta t^c/\delta t^f$ where superscripts c and f correspond to ``course'' and ``fine'' grids. The non-equilibrium distribution function is $f_\alpha^{neq} = f_\alpha - f_\alpha^{eq}$.

The well known difficulty of the conventional method by Dupuis & Chopard method [6] is that kinked profiles of turbulent flow quantities are inevitable at the

interface of the grids. The present authors pointed out the reason of the difficulty and proposed the following correction method [7].

The conditions of mass and momentum conservation: $\Sigma_\alpha f_\alpha^{neq} = 0, \Sigma_\alpha \xi_\alpha f_\alpha^{neq} = \mathbf{0}$ are not always satisfied at the interface of the conventional method. To correct such imbalances, the residuals were calculated at the interface shown in Fig. 4 as

$$\varepsilon_\rho = \sum_{\beta c} f_{\beta c}^{neq,c} + \sum_{\hat{\beta} c} f_{\hat{\beta} c}^{neq,f\to c}, \tag{9}$$

$$\varepsilon_{\mathbf{u}} = \sum_{\beta c} \xi_{\beta c} f_{\beta c}^{neq,c} + \sum_{\hat{\beta} c} \xi_{\hat{\beta} c} f_{\hat{\beta} c}^{neq,f\to c}. \tag{10}$$

When $\varepsilon_\rho, \varepsilon_{\mathbf{u}}$ vanish, the conservation of mass and momentum is achieved. Hence, by setting the re-calculated residuals with the corrected density and velocity: $\rho* = \rho + \lambda_\rho \varepsilon_\rho, \mathbf{u}* = \mathbf{u} + \lambda_{\mathbf{u}} \varepsilon_{\mathbf{u}}$, to zero, the correction coefficients were obtained as

$$\lambda_\rho = \lambda_{\mathbf{u}x} = \left\{ 1 - \frac{1}{6}\left(1 - n\tau^c/\tau^f\right)\left(1 + \frac{c^2}{c_s^2}\right) \right\}^{-1}, \tag{11}$$

$$\lambda_{\mathbf{u}y} = \lambda_{\mathbf{u}z} = \left\{ 1 - \frac{1}{18}\left(1 - n\tau^c/\tau^f\right)\left(\frac{c^2}{c_s^2}\right) \right\}^{-1}, \tag{12}$$

where $c = \Delta/\delta t$. Those corrected variables were used for the simulation.

The performance of this imbalance correction (IBC) method is shown in Fig. 5. It does not produce kinks in the Reynolds stress distributions and improves mass conservation drastically.

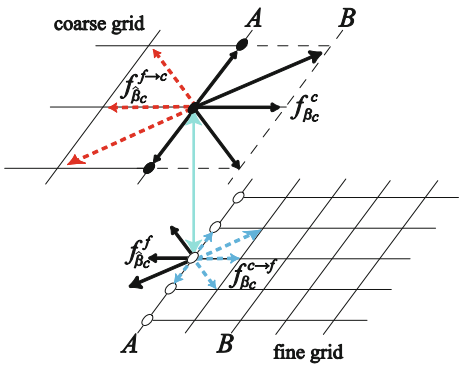


Fig. 4 Stencil of the IBC local mesh refinement

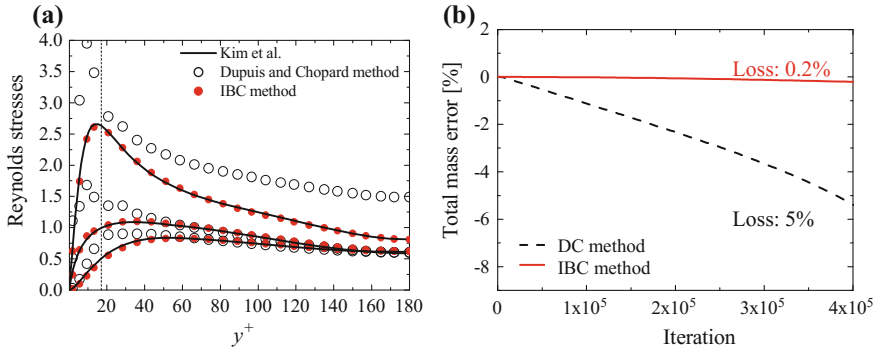


Fig. 5 Performance of the IBC method in LES: **a** Reynolds normal stress profiles, **b** mass conservation history; grid interface is set at $y^+ = 17.7$

4 Conclusions

For turbulent flow simulation by the LBM, the D3Q27 MRT LBM with the IBC mesh refinement method is recommended. It is certainly very powerful for turbulence simulation in a very complex flow region such as that shown in Fig. 6.

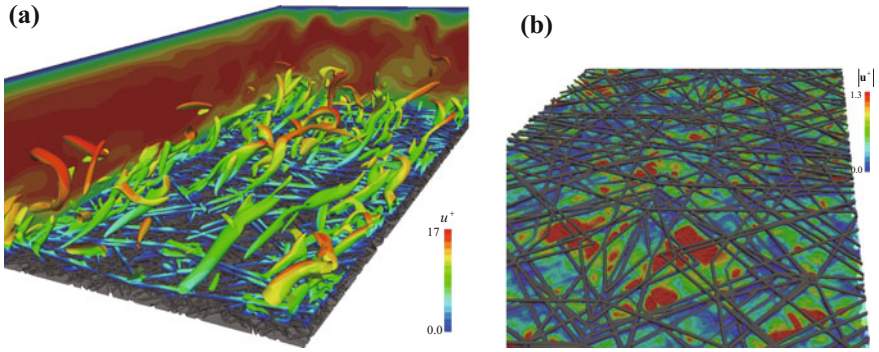


Fig. 6 LES of turbulence inside and over a carbon paper at $\mathrm{Re}_b = 3000$: **a** iso-surfaces of the second invariant of the velocity gradient tensor, **b** velocity magnitudes in carbon paper

Acknowledgements The authors thank their colleague Kei Takashima for his devotion to develop their D3Q27 MRT LBM code. The numerical calculations were carried out on the TSUBAME2.5 supercomputer in Tokyo Institute of Technology.

References

1. Y. Kuwata, K. Suga, Anomaly of the lattice Boltzmann methods in three dimensional cylindrical flows. J. Comput. Phys. **280**, 563–569 (2015)
2. D.J. Holdych, D.R. Noble, J.G. Georgiadis, R.O. Buckius, Truncation error analysis of lattice Boltzmann methods. J. Comput. Phys. **193**, 595–619 (2004)
3. D. d'Humiéres, I. Ginzburg, M. Krafczyk, P. Lallemand, L.-S. Luo, Multiple-relaxation-time lattice Boltzmann models in three dimensions. Phil. Trans. R. Soc. A **360**, 437–451 (2002)
4. K. Suga, Y. Kuwata, K. Takashima, R. Chikasue, A D3Q27 multiple-relaxation-time lattice Boltzmann method for turbulent flows. Comp. Math. Appl. **69**, 518–529 (2015)
5. J. Kim, P. Moin, R. Moser, Turbulence statistics in fully developed channel flow at low Reynolds number. J. Fluid Mech. **177**, 133–166 (1987)
6. A. Dupuis, B. Chopard, Theory and applications of an alternative lattice Boltzmann grid refinement algorithm. Phys. Rev. E **67**, 066707 (2003)
7. Y. Kuwata, K. Suga, Imbalance-correction grid-refinement method for lattice Boltzmann flow simulations. J. Comput. Phys. **311**, 348–362 (2016)

Multi-objective Evolutionary Algorithms for Solving the Optimization Problem of the Surface Mounting

Xuewei Ju, Guangyu Zhu and Shixiang Chen

Abstract The multi-objective particle swarm optimization (MOPSO) is tested by the four ZDT problems. According to the simulation, the MOPSO can locate the global Pareto front (PF) on any instance. Moreover, a three-objective mathematical model is established to verify the ability of MOPSO when solving the practical engineering problems. The simulations show that, compared with the NSGA-II, the MOPSO is able to obtain better optimization results in a short period.

Keywords MOPSO · Pareto front · Test function · Engineering problems

1 Introduction

In practical application, Multi-objective Optimization Problem (MOP) is believed to be difficult and challenging. MOP is different from Single-objective Optimization Problem (SOP) because the objectives in the MOP conflict with each other, and a slight improvement in one objective may affect at least one other objective [1]. Thus, simultaneous optimization of each objective and obtaining the optimal solution of each objective are becoming popular among some researchers.

However, it is difficult for the traditional optimization approaches to solve the MOPs with non-convex, discontinuous and multi-modal solution spaces [2–4]. As a class of stochastic optimization method, Multi-objective Evolutionary Algorithm (MOEA) has been recognized to be well suited to the complex MOP by simulating the process of natural evolution [5]. A classical Pareto-based MOEAs method includes Multi-objective Genetic Algorithm (MOGA) [6], Nondominated Sorting Genetic Algorithm (NSGA) [7], the improved version of NSGA (NSGA-II) [8], etc. Furthermore, some other algorithms including the Particle Swarm Optimization (PSO) algorithm, the Differential Evolution (DE) algorithm, the Bees Algorithm

X. Ju · G. Zhu (✉) · S. Chen
School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou, Fujian 350116, People's Republic of China
e-mail: 724519521@qq.com

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_30

(BeA) and so on, are used to solve the MOPs. Gong et al. [9] proposed a Nondominated Neighbor Immune Algorithm (NNIA) to solve five ZDT problems and a three-objective problem. Chen et al. [10] presented an adaptive PSO to optimize the component placement in printed circuit board (PCB) assembly.

The rest of this paper is organized as follows. Section 2 describes the steps of the multi-objective particle swarm optimization (MOPSO) algorithm. Section 3 gives a detailed description concerning about a three-objective optimization model. Section 4 uses the MOPSO to solve four ZDT problems and the three-objective problem described in Sect. 3. Finally, the conclusions are presented in Sect. 5.

2 The Design of the Algorithm

PSO is a population-based algorithm, which means that PSO begins its search with a randomly generated initial population, and then the evolution of the population is inspired by the principle of the survival of the fittest. Furthermore, the basic PSO based on the mechanism of the Pareto ranking is named as the MOPSO, whose steps are described in detail as follow.

Step 1 Set the parameters of the MOPSO
The corresponding parameters are set in advance, such as the population size M, the inertia weight ω, the cognitive and social parameters c_1 and c_2, the capacity of the external archive $|W_{max}|$, and the maximum number of the iterations T.

Step 2 Initialize the population
At the first iteration, an initial population including M individuals is randomly generated.

Step 3 Update the population

Step 3.1 Calculate the objective vector of each particle
A_i^{gen} in the population represents the ith particle at the genth iteration. Then, the objective vector $F_i(X_i^{gen}) = [f_1(X_i^{gen}), f_2(X_i^{gen}), \ldots, f_n(X_i^{gen})]$ is constructed, where $f_i(i = 1, 2, \ldots, n)$ is the ith objective function.

Step 3.2 Sort the particles based on the mechanism of Pareto ranking
The particles are sorted based on the non-domination and the assigned crowding distance.

Step 3.3 Maintenance of the external archive
Whether the new solutions can enter into the external archive depends on their Pareto ranking. After comparison, the individuals with better fitness values will be added into the external archive, and the poor individuals will be deleted from the external archive as well.

Step 3.4 The operator of updating the particles

The position of the particle A_i^{gen} is represented as $X_i^{gen} = \left(x_{i1}^{gen}, x_{i2}^{gen}, \ldots, x_{in}^{gen}\right)$, and the rate of velocity for the particle A_i^{gen} is represented as v_i^{gen}. P_{id} is the best position of the ith particle among the previous positions, and P_{gd} is the position of the best particle among all the particles in the populations. Then, each particle can obtain a new position and velocity by the formula (1).

$$\begin{cases} v_i^{(gen+1)} = \omega \cdot v_i^{gen} + c_1 \cdot r_1 \cdot (P_{id} - X_i^{gen}) + c_2 \cdot r_2 \cdot \left(P_{gd} - X_i^{gen}\right) \\ X_i^{(gen+1)} = X_i^{gen} + v_i^{(gen+1)} \end{cases} \tag{1}$$

Step 4 Termination

If $gen > T$ or a required level is satisfied, then stop and output the external archive; Otherwise, go to **Step 3 and repeat the whole process mentioned above**.

3 Mathematical Model

As shown in Fig. 1, the configuration of a typically multi-head gantry placement machine is composed of a conveyor, the feeder slots, a moveable arm, two cameras, an automatic nozzle changer and some heads with nozzles.

3.1 Notations

For the convenience of describing the model, the mathematical notations used in the model are summarized as follow.

n	The total number of the components to be mounted.
p	The total number of component types.
L	The total number of pick-and-place cycles.
m	The total number of feeder slot.
$t_{k,k+1}$	The times of nozzle switch.
H	The number of heads.
H_k	Before $(L-1)$th cycle, $H_k = H$. When Lth cycle, $H_k \leq H$.

S	The pickup position where the machine grips the components. Generally, the position is defined as the central position of a feeder slot.
$X = (x_1, \ldots, x_i, \ldots, x_n)^T$	X is the placement sequence for the components,and $x_i (i = 1, 2, 3, \ldots, n)$ is the position of the ith component on the PCB.
x_0	The original position designated before, which is located at the lower-left corner on the PCB.
s^l_{k+1}	The position of the component located at the most left side of the slot in $(k+1)$th cycle.
s^r_{k+1}	The position of the component located at the most right side of the slot in $(k+1)$th cycle.
$d(x_i, x_{i+1})$	The Euclidean distance between the component i and $i+1$.
$d(x_{H_k \times k}, x_{H_k \times k+1})$	The Euclidean distance between the last component of each cycle and the first component of the next cycle.
$d(x_{H_k \times k}, S)$	The Euclidean distance between the last component of each cycle and the pickup position.
$d(S, x_{H_k \times k+1})$	The Euclidean distance between the first component of each cycle and the pickup position.

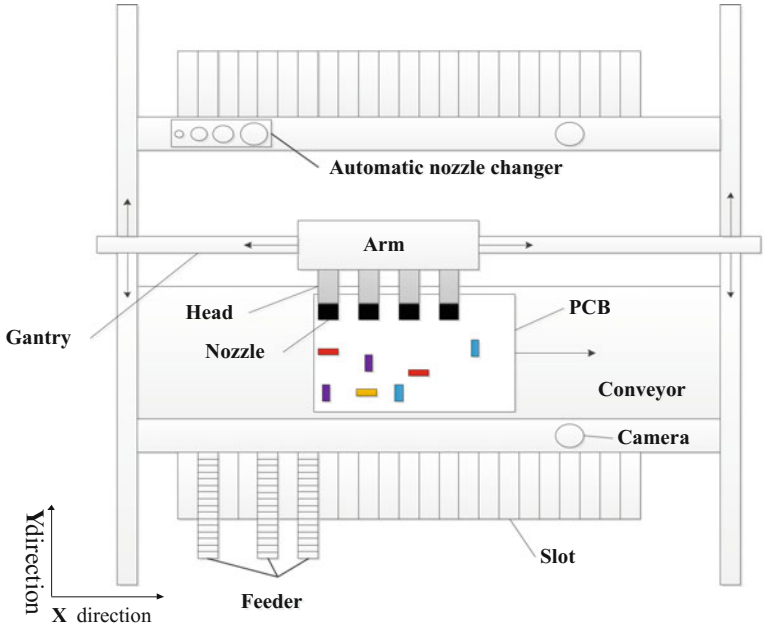


Fig. 1 Schematic diagram of a multi-head placement machine

$d_1\left(x_{H_k\times k}, s^l_{k+1}\right)$	The Euclidean distance between the last component of the kth cycle and the position s^l_{k+1}.
$d_2\left(s^l_{k+1}, s^r_{k+1}\right)$	The Euclidean distance between the position s^l_{k+1} and the position s^r_{k+1}.
$d_3\left(s^r_{k+1}, x_{H_k\times k+1}\right)$	The Euclidean distance between the position s^r_{k+1} and the first component of the $(k+1)$th cycle

3.2 Three-Objective Optimization Model

The component placing problem is divided into three subobjective optimization problems due to its complexity. They are listed as follows:

(1) Pick-and-place sequence (f_1)

The head is required to visit all the components exactly once and return to the starting point. The travelling distance can be described by the formula (2)

$$\begin{aligned} f_1(X) = &\sum_{i=1}^{n-1} d(x_i, x_{i+1}) - \sum_{k=1}^{L-1} d(x_{H_k\times k}, x_{H_k\times k+1}) \\ &+ \sum_{k=1}^{L-1} d(x_{H_k\times k}, S) + \sum_{k=0}^{L-1} d(S, x_{H_k\times k+1}) \end{aligned} \tag{2}$$

(2) Feeder assignment (f_2)

The optimization goal is to minimize the distance that the heads move above the feeder slot when gripping the components, which can be described by the formula (3)

$$f_2(X) = \sum_{k=0}^{L-1} \left\{ d_1\left(x_{H_k\times k}, s^l_{k+1}\right) + d_2\left(s^l_{k+1}, s^r_{k+1}\right) + d_3\left(s^r_{k+1}, x_{H_k\times k+1}\right) \right\} \tag{3}$$

(3) Times of the nozzle exchange (f_3)

The optimization goal is to minimize the times of the nozzle switch, which can be shown by the formula (4)

$$f_3(X) = \sum_{k=1}^{L-1} t_{k,k+1} \tag{4}$$

$$\text{Where} \quad t_{k,k+1} = \begin{cases} 0 & \begin{pmatrix} \textit{if the component type is} \\ \textit{identical in k th and} (k+1) \textit{th cycle} \end{pmatrix} \\ 1 & (\textit{otherwise}) \end{cases}$$

Moreover, a three-objective mathematical model is established based on the above subproblems. The aim is to determine the placement sequence X, and to simultaneously keep the three objective values as small as possible, which is described by formula (5–8).

$$\begin{aligned} F &= min\{f_1(X), f_2(X), f_3(X)\}, \\ X &= (x_1, \ldots, x_i, \ldots, x_n)^T \end{aligned} \tag{5}$$

Subject to:

$$\sum_{k=1}^{L} H_k = n \tag{6}$$

$$1 \le s_{i+1}^{l}, s_{i+1}^{r} \le m \tag{7}$$

$$1 \le \sum_{k=1}^{L-1} t_{k,k+1} \le L - 1 \tag{8}$$

where constraint (6) assures the total number of the components are mounted and constrained at the right place. (7) guarantee the encoding number is within the range of feeder slot. Constraint (8) requires that the total times of the nozzle exchange must be among the total numbers of the pick-and-place cycles.

4 Experiments and Results

In this section, some experiments are conducted to verify the effectiveness of the proposed MOPSO. Firstly, we use the MOPSO to solve four ZDT problems for testing its ability of approximating the Pareto-optimal fronts. Then, compared with NSGA-II which represents the state-of-the-art technology in MOEAs, we use the MOPSO to solve the practical engineering problem demonstrated in Sect. 3. The simulation results are given in Table 1, and some plots for MOPSO solving four ZDT problems are shown in Fig. 2. Moreover, two algorithms and models are programmed through MATLAB 7.11, and the simulations are conducted in a operating environment of Windows 10 professional with an Intel i5-4590 (3.3 GHz) and 4 GB RAM.

Table 1 The simulation results of six instances

Instance ***n***(number)/***p*** (type)	Algorithm	Function values of the best Pareto-optimal solution (f_1, f_2, f_3) (mm, mm, times)	CPU Times (s)
A board 28/6	NSGA-II	(1810.35, 2158.294, 3.4)	10.08
	MOPSO	(1738.203, 1505.741,4)	9.75
B board 60/6	NSGA-II	(3695.36, 4285.79, 6.8)	17.05
	MOPSO	(3598.626, 3601.588, 4.8)	16.95
C board 104/7	NSGA-II	(12, 633.93, 8515.1, 22.5)	23.93
	MOPSO	(12,810.62, 7234.975, 17)	22.06
D board 168/7	NSGA-II	(20, 873.71, 11, 073.48, 37.9)	36.20
	MOPSO	(20, 010.28, 12, 541.35, 16.7)	35.25

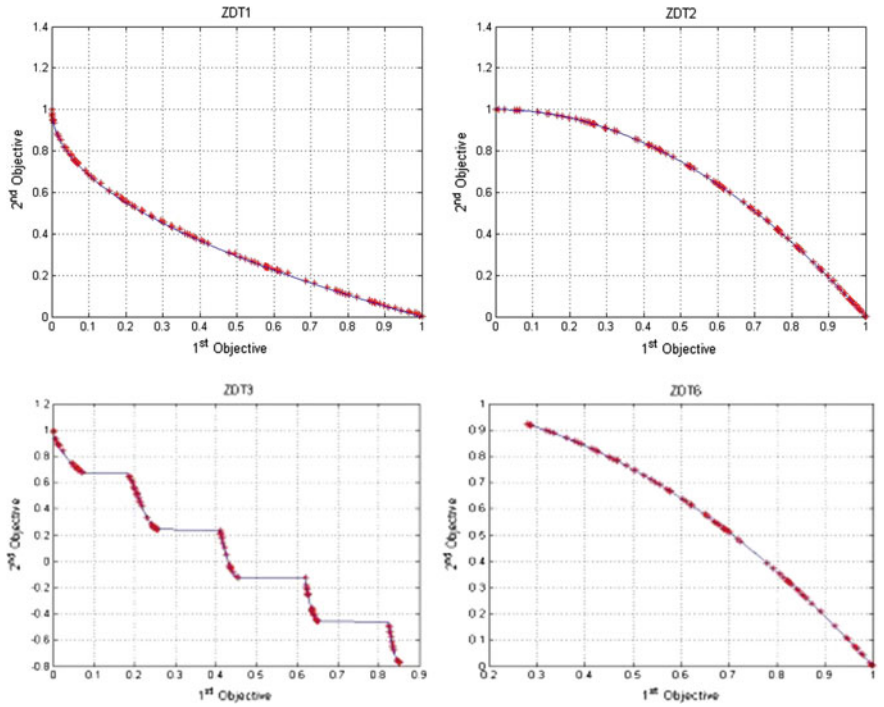


Fig. 2 Plots of the nondominated solutions with the Pareto ranking

4.1 Test Instances

We use four widely used bi-objective ZDT test instances [11], and these test instances are all needed to be minimized. In the objective space, the distribution of the final external archive is shown in Fig. 2, in which the Pareto-optimal fronts of

the four ZDT problems are shown by the black curves. Obviously, MOPSO can locate the global Pareto front (PF) on any instance.

Consequently, we can conclude that MOPSO has an excellent ability of global search, and the Pareto-optimal fronts obtained by MOPSO can be closely deemed as the true Pareto-optimal fronts.

4.2 Practical Engineering Problem

Furthermore, the practical engineering problem established in Sect. 3 is optimized by MOPSO, and the multi-objective optimization results are given in Table 1. In the experiments, MOPSO is tested on four randomly generated PCB instances, in which the number of components (n) ranges from 28 to 168 and the component type (p) reaches up to 7 different types. The optimization results of each instance are the average values of 20 independent runs.

As shown in the third column of the Table 1, from A to D board, function values of the best Pareto-optimal solution obtained by MOPSO are smaller in two objectives than the function values obtained by NSGA-II. In other words, the ability of MOPSO searching the best solutions is excellent. In addition, as shown in the fourth column of Table 1, NSGA-II spends more computational time than MOPSO. Namely, MOPSO is more time-saving in terms of the CPU times.

5 Conclusion

A three-objective optimization model is established in this paper. The proposed MOPSO is used to solve four ZDT problems. Then, a mathematical model established from the practical engineering problem, is used for testing the effectiveness and feasibility of the proposed method. In terms of the benchmark test function, the simulation results indicate that the proposed MOPSO can approximate the Pareto-optimal fronts with sufficient resolution. Furthermore, in terms of the practical engineering problem, the proposed MOPSO is also able to give a better solution for Pareto-optimal problem in an acceptable computation time, when comparing to NSGA-II method. Thus, it can be employed to fast evaluate and solve the multi objective problem with a better overwhelming random search ability and computing efficiency than the ordinary NSGA-II method.

References

1. O. Schütze, V.A.S. Hernández, H. Trautmann, et al., The hypervolume based directed search method for multi-objective optimization problems. J. Heuristics 1–28 (2016)

2. J.D. Schaffer, Multiple Objective Optimization with Vector Evaluated Genetic Algorithms. International Conference on Genetic Algorithms. (Pittsburgh, Pa, USA, July. 1985), 93–100
3. E. Zitzler, L. Thiele, Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. IEEE Trans. Evol. Comput. **3**(4), 257–271 (1999)
4. E. Zitzler, M. Laumanns, L. Thiele, SPEA2: Improving the strength Pareto evolutionary algorithm. Europe. **3242**(103), 95–100 (2001)
5. M. Raghuwanshi, O. Kakde, Survey on multiobjective evolutionary and real coded genetic algorithms, in *The Asia Pacific Symposium on Intelligent and Evolutionary Systems,* (2004), 151–163
6. C.M. Fonseca, P.J. Fleming, Genetic algorithms for multiobjective optimization: formulation discussion and generalization, in International Conference on Genetic Algorithms. (Morgan Kaufmann Publishers Inc. 1999), 416–423
7. N. Srinivas, K. Deb, Muiltiobjective optimization using nondominated sorting in genetic algorithms. Evol. Comput. **2**(3), 221–248 (1993)
8. K. Deb, A. Pratap, S. Agarwal et al., A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. **6**(2), 182–197 (2002)
9. M. Gong, L. Jiao, H. Du et al., Multiobjective immune algorithm with nondominated neighbor-based selection. Evol. Comput. **16**(2), 225–255 (2008)
10. Y.M. Chen, C.T. Lin, A particle swarm optimization approach to optimize component placement in printed circuit board assembly. Int. J. Adv. Manuf. Technol. **35**(5–6), 610–620 (2007)
11. E. Zitzler, K. Deb, L. Thiele, Comparison of multiobjective evolutionary algorithms: empirical results. Evol. Comput. **8**(2), 173–195 (2000)

Molecular Dynamics Studies on Vibration of $SLMoS_2$ Nanoresonantor Under Different Boundary Conditions

Haili HUANG and Minglin LI

Abstract Nanoresonators based on graphenes have attracted worldwide attentions for applications in the field of modern microelectronics and nanoelectronics devices and sensors, due to their high stiffness and outstanding electrical properties. As a graphene-like quasi-two-dimensional (2D) nanomaterials, single-layer molybdenum disulfide ($SLMoS_2$) has the direct bandgap electric property, which is superior than the indirect bandgap of graphene. Compared with the extensive interests on the vibration of graphene nanoresonators with molecular dynamics simulations (MDs), there is lack of knowledge about the vibration of $SLMoS_2$ nanoresonators. Here, we carry out a set of MDs to investigate the fundamental resonant frequency of $SLMoS_2$ nanoresonators under different boundary conditions. Three types of beams, including simply supported beam, clamped-clamped beam, and fixed-fixed beam, are considered in our simulations. The effects of beam length, boundary conditions, and single S point vacancy defects on the fundamental resonant frequency of $SLMoS_2$ are mainly discussed. The interatomic interaction potentials of reactive empirical bond-order (REBO) is adopted in our simulations, after compared with the Stillinger-Weber (SW) potentials. It is found that the calculated frequencies with REBO potentials are higher than that using SW potentials. The resonant frequencies of clamped-clamped beams are found to be greater than other two types of beams. The resonant frequency is found to be inversely proportional to the square of beam length, i.e. decreasing with the increase of beam length, which is in line with the classical vibration theory of continuum beams. The existence of single S point vacancy defects will lead to the drift of resonance frequencies and motivate a new resonance modality.

Keywords $SLMoS_2$ · Vibration · Euler-Bernoulli beam · MD simulations · Frequency

H. HUANG · M. LI (✉)
School of Mechanical Engineering and Automation, Fuzhou University, 2 Xueyuan Road, Minhou, Fuzhou, Jujian 350116, China
e-mail: liminglin@fzu.edu.cn

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_31

1 Introduction

The graphene discovered in 2004 opened a new era in the field of nanotechnology, and it have attracted great attention due to its excellent mechanical and electrical properties [1, 2]. As a new kind of graphene-like two-dimensional (2D) material, single-layer MoS_2 ($SLMoS_2$) has special bandgap characteristics (1.9 eV direct bandgap), high Young's modulus ($\sim$200 GPa), ultralow weight and ultrahigh bending flexibility, which makes it to be a very promising materials for the application of nanodevices [3, 4].

At present, there are extensive study on the vibration characteristics of nano graphene resonators [5–7], while those of nano MoS_2 resonator [3] is just rising up. Jiang [3] investigated the intrinsic energy dissipation in $SLMoS_2$ nanoresonators and compared their dissipation characteristics with those of graphene via classical molecular dynamics simulations (MDs). They found that $SLMoS_2$ nanoresonators exhibit considerably lower energy dissipation than graphene ones under the same conditions, generating both higher Q-factors and figure-of-merit of MoS_2 than those of graphene. Lee [8] demonstrated MoS_2 NEMS resonators with resonances in the high and very high frequency (HF and VHF) bands by a series of experiments, achieving excellent displacement sensitivity and high fundamental-mode frequency-quality factor. They [9] also explored effects of 662 keV γ-ray radiation on multilayer MoS_2 nanoresonators using a 1 mCi 137Cs source and observed the excellent γ-ray radiation sensitivity of the MoS_2 nanoresonators, which endowing its capability to detect radiations in harsh environments. Castellanos [10] fabricated the $SLMoS_2$ mechanical resonators and generated the resonant frequencies in the order of 10 $\sim$ 30 MHz and quality factors of about $\sim$55 at room temperature in vacuum. They demonstrated the attributes of the nonlinear dynamics of the $SLMoS_2$ resonators and found that the mechanical behavior of these single-layer resonators lies on the membrane limit (tension dominated), whereas their thicker counterparts should be modeled as circular plates. Jia [11] fabricated the large arrays of suspended MoS_2 nanomechanical resonators by the means of water-assisted lift-off and all-dry transfer. The resultant large arrays of suspended single- and few-layer MoS_2 drumhead resonators shows the fundamental resonances in the VHF band with outstanding figure-of-merit. Lee [12] demonstrated measurements of pressure dependence in atomically-thin drumhead MoS_2 resonators at high frequencies and found exceptional responsivities of resonance frequency in different pressure regimes. Maitra [13] performed the finite element (FE) simulations on the longitudinal vibration of piezoelectric $SLMoS_2$, with various sizes and under different boundary conditions. They also extracted the longitudinal acoustic velocity of the system from the simulation and verified using the analytical analysis.

Structural defects in $SLMoS_2$ have important influences on the electric, optical, thermal, chemical, and mechanical properties of the material, which can either be formed during growth process or introduced by electron irradiation [15]. Li [14] conducted MDs to investigate the effect of V_{MoS3} point defects on the elastic

properties of $SLMoS_2$ sheets. They revealed that the increase of V_{MoS3} vacancy concentration will lead to the noticeable decrease in the elastic modulus but has a slight effect on Poisson's ratio under the constrained uniaxial tensile test. Ding [15] observed that applying mechanical strain and introducing lattice defects can reduce the thermal conductivity of $SLMoS_2$ via MDs. Zhou [16] investigated four different intrinsic structural defects in chemical vapor phase grown $SLMoS_2$ and studied their energy landscape and electronic properties using first-principles calculations. They found that two different types of 60° grain boundaries can create one-dimensional metallic wires.

As mentioned above, $SLMoS_2$ resonators have attracted more and more attentions from the viewpoints of experimental and theoretical studies. However, so far, few articles about vibration of $SLMoS_2$ resonators with structural defects using continuum beam theory and MDs have been reported. In this work, we model the vibration of $SLMoS_2$ via the Euler-Bernoulli beam theory and perform a set of MDs to investigate the vibration of $SLMoS_2$ resonators under different boundary conditions, i.e. simply-supported, fixed-supported, and clamped-clamped boundary. The effects of interatomic interaction potentials, beam length, boundary conditions, and single S point vacancy defects on the fundamental resonant frequency of $SLMoS_2$ resonators are mainly discussed. The simulation and analysis show that the resonant frequencies of clamped-clamped beams are higher than other two types of beam. The resonant frequency is found to be inversely proportional to the square of beam length, i.e. decreasing with the increase of beam length, which is in line with the classical vibration theory of continuum beams.

2 Models and Methods

2.1 Euler-Bernoulli Beam Vibration of $SLMoS_2$

The model of $SLMoS_2$ based on Kirchhoff plate theory is shown in Fig. 1. Applying a uniformly distributed transverse load on the surface of an infinite long rectangular plate produce a cylindrical bending surface, with the axis of the cylinder parallel to the length of the plate. Here, a unit length strip of the plate (shown as the green atomic group in Fig. 1) is taken as a classical Euler-Bernoulli beam, which can reflect the behavior of the whole plate. In fact, a variety of different boundary conditions are often involved in the engineering application of Euler-Bernoulli beam. Hence, three beam models are considered in our simulations and shown in Fig. 2: (a) simply supported (SS) beam with one fixed end and one slidable end, (b) fixed-fixed (FF) beam with two fixed ends, (c) clamped-clamped (CC) beam with two clamped ends.

According to the elastic dynamics theory, the fundamental frequency for the beams under these three different boundary conditions can be given as

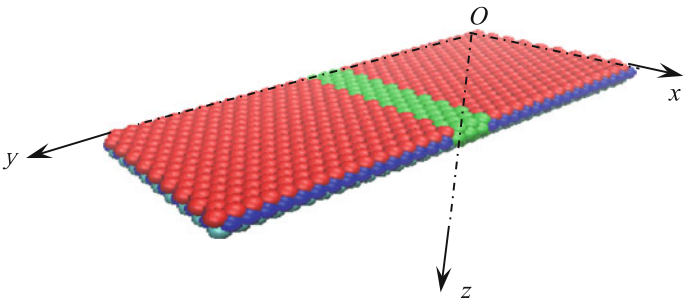


Fig. 1 The Kirchhoff plate model of $SLMoS_2$ with a reference coordinate system. The armchair lattice direction parallel to x axis, while the zigzag direction is consistent with y axis. The unit length strip for *green* atomic groups on the $SLMoS_2$ middle surface can be regarded as a Euler-Bernoulli beam

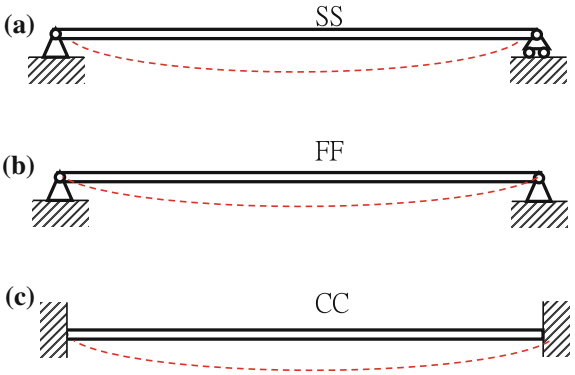


Fig. 2 Euler-Bernoulli beam under different boundary conditions, i.e. (**a**) simply supported, (**b**) fixed, (**c**) clamped boundaries

$$f_i = \frac{i^2\pi}{2l^2}\sqrt{\frac{D}{\rho A}} \quad \text{for SS beam} \tag{2.1}$$

$$f_i = \frac{i^2\pi}{2l^2}\sqrt{\frac{D}{\rho A}} \quad \text{for FF beam} \tag{2.2}$$

$$f_i = \frac{\left(i+\frac{1}{2}\right)^2\pi}{2l^2}\sqrt{\frac{D}{\rho A}} \quad \text{for CC beam} \tag{2.3}$$

where D and A are the bending rigidity and cross-sectional area of the $SLMoS_2$ plate, ρ and i are respectively the mass density and the i th vibrational mode of the beam. In this paper, we consider the first vibrational mode, where i is set to be 1. It is noted that the expression of the fundamental frequency for the SS beam is the same as that for the FF beam, obtained from the theory of continuous mechanical

vibration. Such results are attributed to the negligible distance changes between these two boundary constraints when the vibration occurs, i.e. the small deformation assumption, which leads to identical mathematical boundary conditions in these two beams. In the other words, the deflection and bending moment in both ends of the SS and FF beams are identical to be zero. The fundamental harmonic frequency expression for the CC beam is distinct from the other beams due to its nonzero bending moment boundary conditions for both ends.

2.2 Atomistic Models for Molecular Dynamics Simulation

To mimic the vibration of the three above-mentioned beams in MDs, the atoms in both ends of the beams are constrained respectively, as shown in Fig. 3. For the sake of clarity and simplicity, the zigzag edges of the $SLMoS_2$ plate are adopted to be constrained and the armchair edges is set to be periodical boundary condition to represent the infinite length of the long rectangular $SLMoS_2$ plate. In the SS beam, the bottom S atoms at the left end are fixed in all directions and the bottom S atoms at the right end are set to be free in the xy plane but fixed in the z direction. For the case of FF beam, the top S atoms at both ends are set to be free and the bottom S atoms at both ends are fixed in the all directions. Finally, in the case of CC beam, all edge S atoms at both ends are entirely fixed. In these beams, the cross section includes either only the top and bottom S atoms or only the Mo atoms, which will keep plane as the vibration occurs. For the beams of SS and FF, the top S atoms at both ends have the ability to rotate around the bottom S atoms during the vibration, which is related to the condition of zero bending moments.

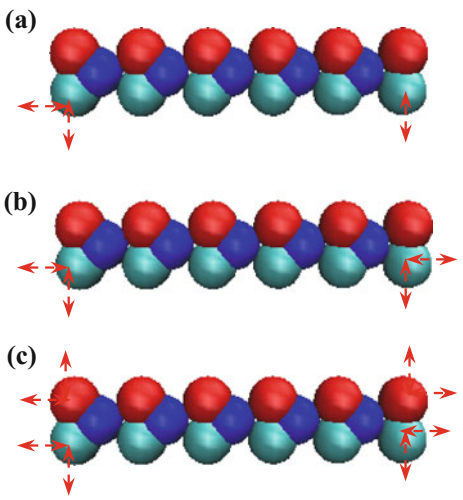


Fig. 3 Three diagrams of the three boundary conditions. *Red arrows* prevent the attached atoms from moving along their axes

2.3 *Molecular Mechanics Simulations*

Before setting the boundary conditions, the atomic beams are relaxed by minimizing the system energy with the stopping tolerance of zero energy and 10^{-10} eV/Å force. The conjugate gradient algorithm is employed for the minimization of the system energy. After setting the boundary conditions, a uniformly distributed lateral load is applied to every atom in the top S-layer and middle Mo-layer of the $SLMoS_2$ plate, to produce an original bending deformation. The bent structure of the beam is close to the fundamental resonant mode. Then the load is released and the beam starts to freely oscillate with the state of microcanonical ensemble, where the deformation energy and the kinetic energy of the beam can be converted into each other. The resonant displacement of the beam is recorded along with the related timestep of 1 fs for 2 ns. Thereafter, the fundamental resonant frequency of the beams can be deduced from the recorded data using the Fast Fourier Transform (FFT) method. The effect of the ambient temperature has not been considered in this work.

3 Result and Discussion

In order to manifest the effectivity of the vibration system with MDs, Fig. 4 show the system energies (including potential energy, kinetic energy and total energy) varying with MDs time using the REBO potential. The total energy of the system, which equals the potential energy plus the kinetic energy, remains a constant of −12,603.194 eV during the oscillation. It can be figured out that the phase angle difference between the potential energy and the kinetic energy is closer to 180 degrees, which suggests the conservation of total energy. The kinetic energy of the system is small enough so that only the linear vibration is activated, which also can be seen in the Fig. 5. In the Fig. 5, the resonant displacement of the SS beam is recorded as a function of the MDs time, and the corresponding resonant frequency

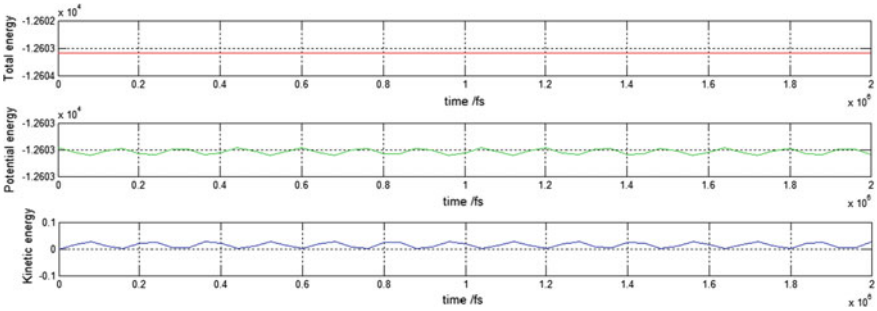


Fig. 4 System energies, i.e. total energy (*red*), potential energy (*green*) and kinetic energy (*blue*) versus MDs time

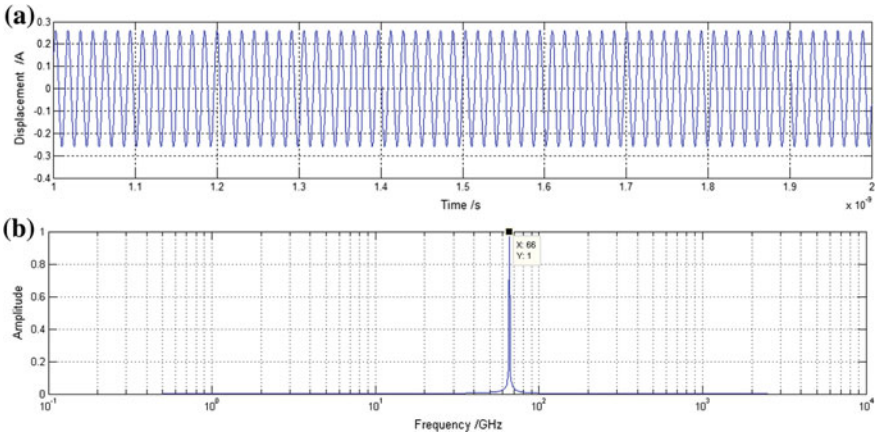


Fig. 5 Resonant displacement (**a**) and frequency (**b**) of the SS beam

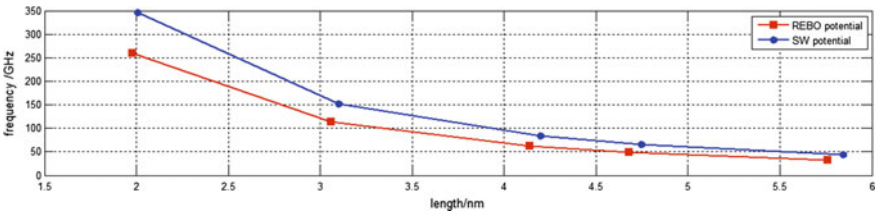


Fig. 6 The resonant frequencies of the $SLMoS_2$ beams varying with the beam length, based on the REBO and SW potentials

is also extracted using the FFT method to be 66.00 GHz. The single peak in the frequency spectrum indicates that only the fundamental resonant frequency is excited regardless of the higher vibrational mode. The equal amplitude oscillations in Fig. 5a illustrates that no energy dissipation occurs during the vibration.

It is well known that the interaction potential between atoms in nanomaterials is highly related to their physical properties. Several empirical potentials for the $SLMoS_2$ have been reported in the literatures, including the SW potential [17], the constant valence force field (CVFF) [18], and the REBO potential [18–21]. Recently the SW and REBO potentials have been adopted in our works to investigate the mechanical properties of the $SLMoS_2$ [22, 23]. Herein, both REBO and SW potentials are further used to simulate the vibrational response of the $SLMoS_2$. Figure 6 shows the resonant frequencies of the $SLMoS_2$ beams varying with the beam length via the REBO and SW potentials. The resonant frequency is found to be inversely proportional to the square of beam length, i.e. decreasing with the increase of beam length, which is in line with the classical vibration theory of continuum beams. It is also clearly observed that the frequencies obtained from the REBO potential are larger than those from the SW potential as the beam length less

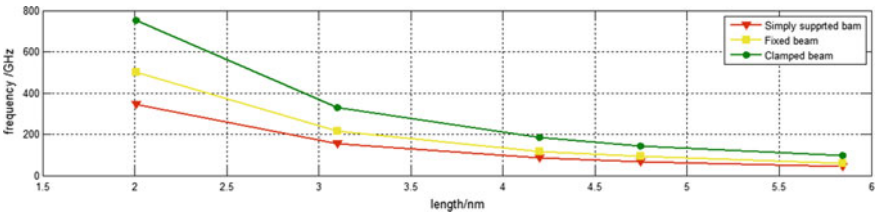


Fig. 7 Resonant frequencies of $SLMoS_2$ beams versus the beam length under different boundary conditions

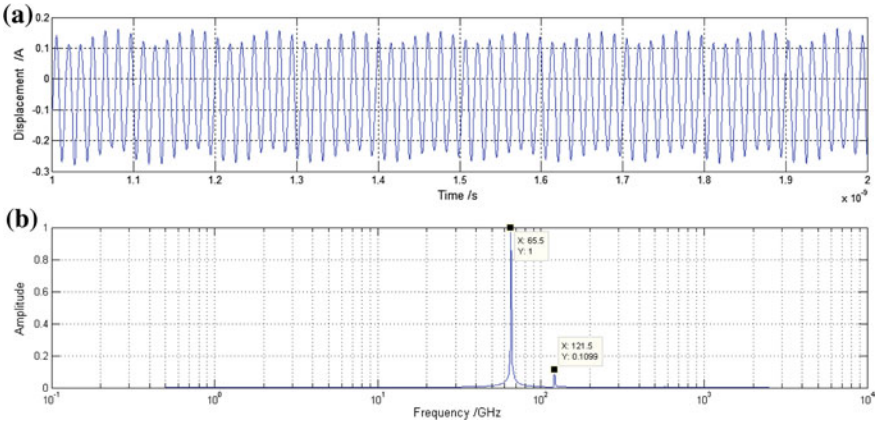


Fig. 8 (**a**) Resonant displacement of the beam with the defect concentration of 0.4% versus the MDs time, (**b**) the frequency spectrogram of the beam with the defect concentration of 0.4%

than 6 nm, which means that the bending stiffness of the $SLMoS_2$ obtained from the REBO potential is larger than that from the SW potential, agreeing with the previous studies. However, the frequencies obtained from these two potentials tend to coincide with each other as the beam length increases. In the following simulations, only the REBO potential is employed to investigate the effect of the boundary conditions and the defect concentration of single S atomic vacancies.

Figure 7 shows the resonant frequencies of $SLMoS_2$ beams varying with the beam length under three different boundary conditions. In accordance with the Eqs. 2.1–2.3, the resonant frequencies of the CC beam are larger than those of the other beams. However, the frequencies of the FF beam are larger than those of the SS beam, which is contrary to the theoretical prediction base on the elasticity continuum theory. It is probably due to the change of the distance between the constrained boundaries in the SS beam. However, the resonant frequencies of these two beams have the trend to converge on the same value as the length increases.

The effect of single S point vacancy defects on the fundamental resonant frequency of $SLMoS_2$ is also investigated in this work. A typical vibrational displacement and frequencies of a $SLMoS_2$ beam with a single S atomic vacancy

Table 1 Resonant frequencies of SLMoS2 beams with various defect concentrations

Defect concentration	The 1st mode frequency/(GHZ)	The 2nd mode frequency/(GHZ)
0	66.00	…
0.0022	65.50	…
0.0039	65.50	121.50
0.0062	65.50	121.50
0.0079	65.50	121.00
0.0101	65.50	…
0.0118	65.00	120.50
0.0140	65.00	122.00
0.0163	65.00	121.50
0.0180	65.00	…

concentration of 0.4% are shown as a function of the MDs time, given in Fig. 8. It is worth noting that there are two frequencies present in the frequency spectrogram. First frequency is close to the fundamental one of the defect-free $SLMoS_2$ beam, howbeit a slight draft is observed probably due to the defect. In addition, the newcome frequency of the higher-frequency mode can be confirmed due to the effect of the introduced defect. However, not all the beams with defects will induce the higher frequency in our simulations. The resonant frequencies of $SLMoS_2$ beams with various defect concentrations are given in Table 1.

4 Conclusions

We investigate the vibrational response of the $SLMoS_2$ beam under different boundary conditions, and the effect of the beam length, the concentration of single S atomic vacancy defects, and the interatomic interaction potentials, using MDs. It is found that the calculated frequencies with REBO potentials are higher than that using SW potentials. The resonant frequencies of CC beams are found to be higher than other two types of beams. The resonant frequency is found to be inversely proportional to the square of beam length, i.e. decreasing with the increase of beam length, which is in line with the classical vibration theory of continuum beams. The existence of single S atomic vacancy defects will lead to the drifting of resonant frequencies and motivate a newcome higher frequency.

Acknowledgements M. Li gratefully acknowledges the financial support from the National Natural Science Foundation of China (Grant No. 50903017), and M. Li thanks the financial support from the Fujian Collaborative Innovation Center of High-End Manufacturing Equipment.

References

1. R. Chowdhury, S. Adhikari, F. Scarpa et al., J. Phys. D **44**, 205401–205411 (2011)
2. B. Arash, Q. Wang, J. Nanotechnol. Eng. Med 2, 011012–011012-7 (2011)
3. J.-W. Jiang, H.S. Park, T. Rabczuk, Nanoscale **6**, 3618–3625 (2014)
4. R. Yang, Z.-H. Wang, and X.-L. Feng, IEEE MEMS, 59–62 (2016)
5. J.W. Kang, H.-W. Kim, K.-S. Kim et al., Curr. Appl. Phys. **13**, 789–794 (2013)
6. O.K. Kwon, K.-S. Kim, J. Park et al., Comput. Mater. Sci. **67**, 329–333 (2013)
7. O.K. Kwon, J.H. Lee, J. Park et al., Curr. Appl. Phys. **13**, 360–365 (2013)
8. J. Lee, Z.-H Wang, K.-L He et al., ACS Nano 7, 6086–6091 (2013)
9. J. Lee, M.J. Krupcale, P.X.L. Feng, Appl. Phys. Lett. **108**, 023106 (2016)
10. A. Castellanos-Gomez, R. V. Leeuwen, M. Buscema et al., Adv. Mater. 25, (6719–6723)
11. H. Jia, R. Yang, A.E. Nguyen, S.N. Alvillar et al., Nanoscale **8**, 10677–10685 (2016)
12. J. Lee, X. L Feng, IEEE FCS, 1–4 (2014)
13. M. Maitra, H.B. Nemade, Procedia Eng. **144**, 682–688 (2016)
14. M. Li, Y. Wan, L. Tu et al., Nanoscale Res. Lett. **11**, 155 (2016)
15. Z. Ding, Q.-X. Pei, J.-W. Jiang et al., J. Phys. Chem. C **119**, 16358–16365 (2015)
16. W. Zhou, X.-L Zou, S. Najmaei et al., Nano Lett 13. 2615–2622 (2013)
17. J.-W. Jiang, H.S. Park, T. Rabczuk, J. Appl. Phys. **114**, 064307 (2013)
18. S. Xiong, G. Cao, Nanotechnology **26**, 185705 (2015)
19. S. Xiong, G. Cao, Nanotechnology **27**, 105701 (2016)
20. T. Liang, S.R. Phillpot, S.B. Sinnott, Phys. Rev. B **79**, 245110 (2009)
21. J.A. Stewart, D. Spearot, Modell. Simul. Mater. Sci. Eng. **21**, 045003 (2013)
22. J. Zhao, J.-W. Jiang, T. Rabczuk, Appl. Phys. Lett. **103**, 231913 (2013)
23. J.-W. Jiang, H.S. Park, Appl. Phys. Lett. **105**, 033108 (2014)

First-Principles Investigation of Magnesium Ion Adsorptions and Diffusions on 1H-Monolayer MoS2 for Energy Storages

Jing Luo and Minglin Li

Abstract Effective energy storage and conversion technologies are indispensable for developing advanced rechargeable batteries. Rechargeable Mg batteries have attracted increasing research interest as potential alternative to lithium-ion batteries, due to high energy capability, earth abundant, good operational safety, environmental friendliness, and low cost. Since monolayer MoS_2 have been synthesized through various methods, MoS_2 nanostructures have been applied to enhance the energy storage in the alkali-ion batteries. However, few works focus on the theoretical investigation of the Mg-ion batteries based on 1H-MoS_2 monolayer. In this paper, we conducted the first-principles calculations to investigate the Mg ions adsorption on and diffusion through the 1H-MoS_2 monolayer. Three adsorption sites, including top of S atom, top of Mo atom, and hole site (centre of hexagonal lattice), were considered for measuring the binding energy of Mg-ions. Result show that the binding energy at the top of Mo atom is the larger than those at the other adsorption sites, which indicates the possibly stable absorption configuration. Energy barriers of Mg-ions passing through and diffusing over the MoS_2 monolayer are also calculated. It was found that the Mg-ion has a largest energy barrier to pass through the 1H-MoS_2 layer and a smallest energy barrier to diffuse above the 1H-MoS_2 layer. The band structures and density of states of MoS_2 before and after Mg-ion absorption were also calculated and discussed.

Keywords First principle · MoS_2 · Adsorption · Diffusion · Energy storage

J. Luo · M. Li (✉)
School of Mechanical Engineering and Automation, Fuzhou University,
2 Xueyuan Road, Minhou, Fuzhou 350116, Fujian, China
e-mail: liminglin@fzu.edu.cn

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_32

1 Introduction

After rapid development over past decades, Lithium ion battery (LIB) technology has grown to be one of the most mature energy storage technology after several years of rapid growth. LIB has been widely used in all walks of life, especially in fields of electric vehicle and energy storage [1], which has the advantages of large electrical storage capacity, high charging and discharging cycles, as well as operational safety and environmental friendliness [2]. In recently years, the development of LIB is moving toward the direction of lowering cost, enhancing storage capacity and high power, prolonging its life cycle, etc. It is widely accepted that the electrode is of crucial for the design of the battery and the innovation of battery materials [3]. At present, the commonly used anode material for LIB is graphite [4], which has wide resource, good electrical conductivity, and suitable layer structure. Layered structures are suitable to act as a host material for anode, which allows the ion to be embedded and deintercalated expediently. Graphene-based composites have been demonstrated to show larger capacity (600 $\sim$ 900 mA h/g) than that of graphite (372 mA h/g), however, the monolayer MoS_2 based composites have exhibited the highest specific capacity ($\sim$1200 mA h/g) [23].

MoS_2 has been considered to be an ideal electrode material for the advanced LIB, due to its distinctive layered structure [5–7]. The layered structure of MoS_2 are favorable for reversible intercalation/deintercalation and diffusion of Li ion without significantly increasing the material volume. However, when applied as a cathode material for LIB, MoS_2 has relatively low average voltage and energy density. Decreasing the layer number down to monolayer leads to improved energy storage of MoS_2. To address the underlying mechanism and the potential application of MoS_2 monolayer to be cathode materials, Li et al. [8] systematically studied the adsorption and diffusion of the Li atom in 1H-MoS_2 monolayer, and revealed that the unique edge state in MoS_2 nanoribbons have the capability of enhancing binding interaction with Li atoms. In spite of great commercial success, lithium-base batteries have several drawbacks, e.g., potential safety issues, high cost, and resource scarcity.

In order to further improve the safety and reduce the cost of LIB [9], sodium-ion (NIB) and magnesium-ion (MIB) batteries, which are still in their infancy, have recently attracted considerable interest as candidates for energy storage applications due to their low price, abundant resource, and better safety. In addition, magnesium ions as the divalent metallic cations can provide twice charge transportation [10]. However, compared to the success of LIB in the commercial field, the other two kinds of battery technology need to overcome some problems and difficulties, such as greater ionic radius [11]. Several experimental and theoretical studies have focused on MoS_2 in various morphologies as a cathode material for MIB to help guide future applications. In an ideal alkali ion batteries, the cathode material should have high ionic intercalation voltage, whereas the anode one should have low ionic deintercalation voltage [12]. In order to improve the performance of LIB, and MIB based on 1H-MoS_2 monolayer, understanding the alkali ionic processes in 1H-MoS_2 monolayer at the atomic scale is very important. In this work, we utilized

first-principles calculations to compare the atomistic adsorption and diffusion of lithium, and magnesium ions. Their band structures and partial densities of states (PDOS) were also investigated.

2 Model and Computational Method

In order to compare the adsorption and diffusion of Li and Mg atoms in monolayer MoS_2, we constructed a periodic 3 × 3 × 1 supercell of MoS_2, which consists of 27 atoms in the supercell, as shown in Fig. 1. Meanwhile, this supercell can avoid the interaction of the adsorbate ions with its images in neighboring supercells. Along with the thickness direction of monolayer MoS_2, a vacuum layer of 18.41 Å is added to the supercell. The calculations are performed using the CASTEP code [13], which is based on density functional theory and plane-wave pseudopotential method [14]. We used PBE exchange-correlation function which adopted in the generalized gradient approximation(GGA) to solve electronic system [15]. After verification and reference, a 3 × 3 × 1 grid for k-point for geometry optimization

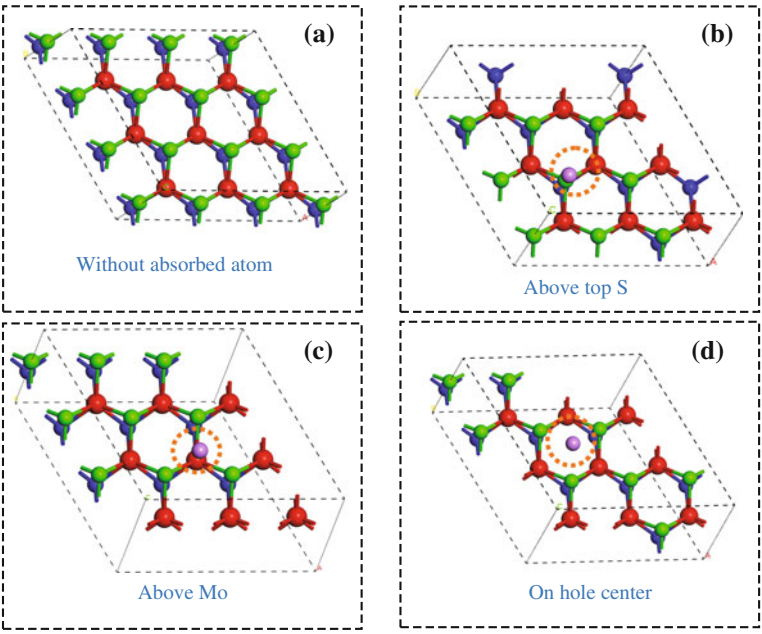


Fig. 1 **a** Defect-free monolayer MoS_2. **b–d** Adsorbed atoms on S top, Mo top, hole sites. Top S atoms colored by green, bottom S atoms are blue, Mo atoms are red, and the adsorptive atom is purple. The circle represents the adsorption sites

and an energy cut-off of 450 eV were consistently used in our calculations. The energy tolerance was set to be 2.0 x 10^{-5} eV/atom.

3 Results and Discussion

The perfect monolayer 1H-MoS_2 supercell was initially optimized to obtain more accurate lattice constant. The calculated lattice constant is $a = b = 3.18$ Å (experimentally $a = b = 3.16$ Å [16]). To study the adsorption of Li, Mg atoms on the monolayer, the hexagonal symmetry and selected three major adsorption sites were considered. Three representative sites, top of a Mo atom, top of a S atom in upper plane, and top of six-membered ring centers, for the adsorption of ions on the surface of the monolayer 1H-MoS_2 were investigated (Fig. 1). For every adsorption site, the binding energy of alkali atoms defined by:

$$E_{ads} = E_{ori} + xE_{ion} - E_{tot} \quad (1)$$

Where E_{tot} and E_{ori} are total energies of the 1H-MoS_2 supercell with and without adsorbate atoms, respectively. E_{ion} is the energy of adsorbate atoms, and x is the number of the adsorbate atoms. In terms of this definition, a higher binding energy indicates a more favorable adsorption between monolayer MoS_2 and alkali (Li and Mg) atoms. As for the adsorbate sites, the calculated binding energy, shown in Table 1, Clearly shows that adsorption atoms are preferable to bind on the top of Mo atoms, because the formation energy for the site of top of Mo atoms is higher than those for the other sites.

Compared with the Li atom, the binding energy of Mg atom is the lower, which indicates the stabilities of the composite are in the order LIB > MIB. Next, we focused on the lattice parameter, and the result indicated that the adsorption of Li atoms barely increases the planar lattice constant to 3.196 Å (0.5%), and the adsorption of Mg atom did hardly change the lattice much. It can be ascribed to the monolayer feature of 1H-MoS_2, which allows the adsorbate atoms to be relaxed along the vertical direction. The height in Table 1 is defined as the distance of z position between the adsorbate atoms and Mo atoms. It can be unraveled that the binding energies between the adsorbate atoms and the adsorbate site atom are related to their binding heights. It is noted that no distortion of monolayer MoS_2

Table 1 The binding energy at representative adsorption sites and the binding height

Atoms	Top of S		Top of Mo		Hole	
	Energy (eV)	Height (Å)	Energy (eV)	Height (Å)	Energy (eV)	Height (Å)
Li	1.3075	2.28	2.0261	2.98	1.8870	3.11
Mg	0.3318	3.10	0.5374	3.72	0.5256	3.73

induced by the adsorption of alkali atoms are distinctly observed. It means that the pristine monolayer MoS_2 is relatively stable, which cannot be easily affected by the external perturbation.

In terms of the lithiation/delithiation procession, the performance of these alkali ionic batteries are closely related with the electronic conduction behavior of the electrode materials. Therefore, the band structures and density of states (DOS) were calculated for these alkali atoms adsorbed monolayer 1H-MoS_2. The calculated band structures show that perfect MoS_2 monolayer is an intrinsic direct semiconductor with a gap of 1.66 eV as shown in Fig. 2a, which is lower than the recent experimental result [17] but consistent with the published calculation results [10, 18–22]. Figure 2b, c show that the conduction band drafts to cross the Fermi level when one alkali atom binds on the 1H-MoS_2 monolayer. Compared with the direct bandgap semiconductor of pristine monolayer MoS_2, the whole system is both metallic for the Li and Mg adsorption. The difference of band structure among these systems also reflects different valence electrons in the outer layer of adsorbed atoms. These band structures are similar to that of the N-type semiconductors, which indicates that more electrons will transmit to the conduction band after being excited.

The difference in the electronic properties between the Li- and Mg-adsorbed 1H-MoS_2 monolayers can be confirmed through the PDOS. Figure 3 shows PDOSs of electrons in Li- and Mg-adsorbed MoS_2 monolayers. The energy bands near the Fermi

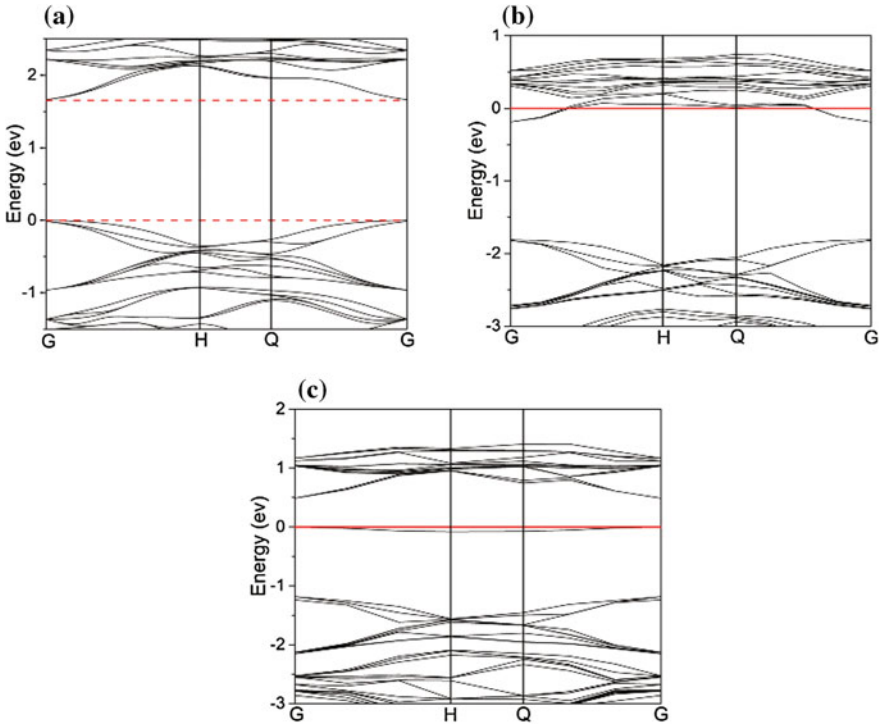


Fig. 2 Band structures of **a** perfect and **b** Li-covered MoS2 monolayer. **c** Mg-covered 1H-MoS2 monolayer

level shown in Fig. 3a, b are mainly provided by the p-orbital of S atoms and d-orbital of Mo atoms. From Fig. 3c, we can see the s-orbital of Mg is higher, and s-orbital of Li is closely to fermi level. Figure 3 shows the PDOS of electrons in Li- and Mg-adsorbed on top Mo-MoS_2 systems. Xu et al. [23] have already shown that the energy band near the Fermi level is mainly provided by the p-orbital of S atoms and d-orbital of Mo atoms, and a large pseudogap below the fermi level. With the increase of the number of adsorbed atoms, the metal properties of system will be enhanced.

To investigate the capacities of the intercalation/deintercalation of alkali atoms, we further calculated the diffusion of adsorbate Li and Mg atoms through and over the MoS_2 monolayer. Here, the atoms passing through the hole sites of the monolayer is defined as diffusion path 1, as shown in Fig. 4a. The atoms moving over the surface from one top site of the Mo atom to the other top site of the nearest Mo atom is also defined as diffusion path 2, as shown in Fig. 5a.

When the alkali atoms passing through the MoS_2 monolayer, we took the z position of the Mo atom as reference. The maximum diffusion energy barriers for Li and Mg atoms through the monolayer, seen in Fig. 4b, are 7.77 eV, and 18.49 eV, respectively. Therefore, the difficulties for these alkali atoms to pass through the monolayer are in the order Li < Mg, which is in accordance with the

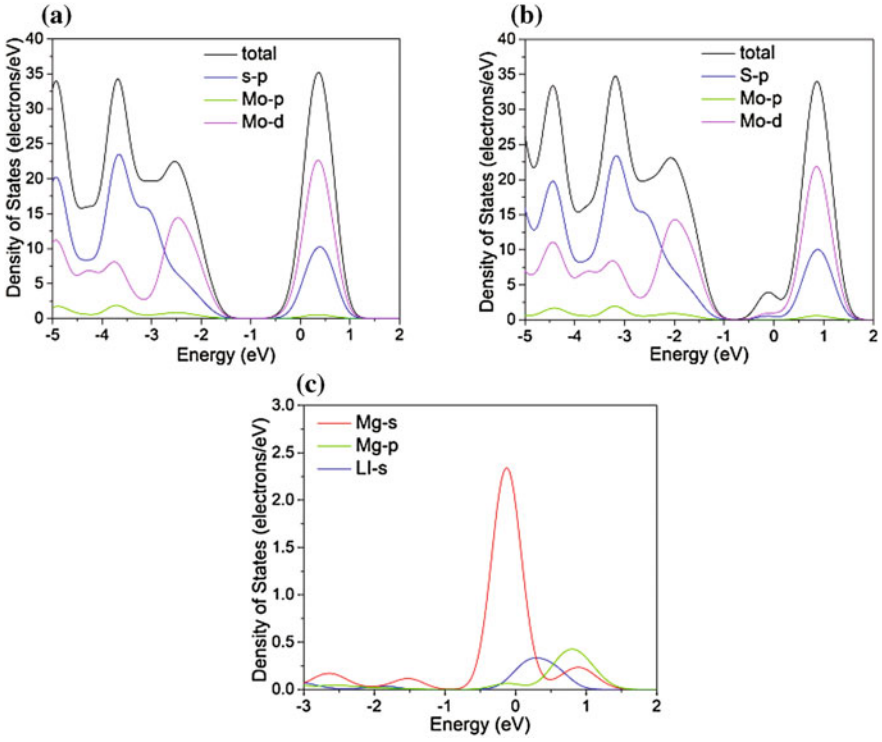


Fig. 3 **a** Li-covered MoS2 PDOS. **b** Mg-covered MoS2 PDOS. **c** The PDOS Mg atom next to the fermi level

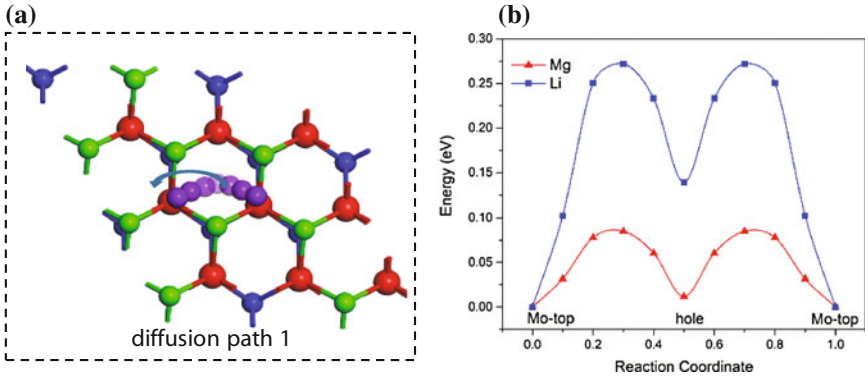


Fig. 4 **a** The path of atomic diffusion. **b** Diffusion barriers of Li, and Mg in MoS2 monolayer along diffusion path1. Color codes are blue = Li, and red = Mg

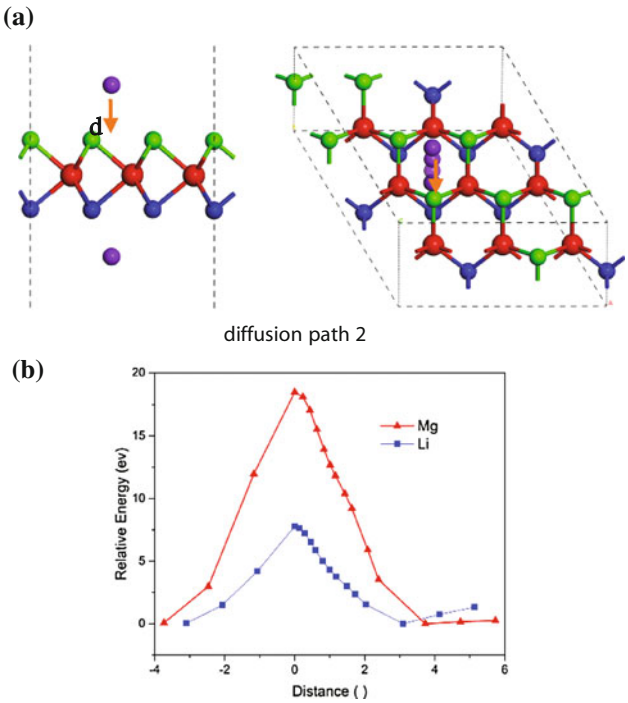


Fig. 5 **a** The path of atomic diffusion on the surface of monolayer MoS_2. **b** Energy curve along the atoms diffusion path 2

order of the element in the periodic table. The maximum diffusion energy barriers are present at the zero position, in which the distances between the alkali atom and the nearest Mo atoms are the shortest, smaller than the distance between the alkali atom and the S atoms. Hence, we can suggest that the maximum energy barrier mainly stems from the interaction between the alkali atom and the Mo atoms.

Figure 5b shows the energy barriers of alkali atoms diffusing over the surface of MoS_2 monolayer along the path 2. When the alkali atom moves from the top of Mo atom to the hole site, the relative energy barrier increases first and then decreases due to changing the distance between the alkali atom and the Mo and S atoms. It can be extracted that the energy barriers are 0.27 eV for Li atom diffusion, and 0.08 eV for Mg atom, lower than commercial electrode materials [23]. In particular, the energy barriers for Mg atoms are lower than that over black phosphorous [3], indicating that the Mg atoms are easier to diffuse over the MoS_2. Monolayer than over the black phosphorous and commercial electrode materials. The 1H-MoS_2 monolayer is prefer to be anode materials for MIB.

4 Conclusion

In summary, the MoS_2 monolayer for Li and Mg atoms adsorption and diffusion were investigated by first-principle calculations. The binding energies at three adsorbate sites are calculated, and it is found that the most stable adsorption site is at the top of Mo atom. The length of Mo-S bond changes little after the atomic adsorption. After the adsorption of atom, the structure of the material is steady, with insignificant distortion. The band structure and PDOS of the system are analyzed to investigate the electrical properties. With the increase of the number of adsorbed atoms, the metallicity character of system will be enhanced. The energy barrier of the alkali atoms passing through the monolayer MoS_2 and diffusing on the surface of the monolayer MoS_2 are investigated. Results show that the alkali atoms can easily diffuse on the surface of the monolayer MoS_2 based on lower energy barriers, compared with passing through the monolayer MoS_2. Among these alkali atoms, it is easier (more difficult) for Mg (Li) atoms to penetrate through the monolayer and diffuse on the surface of the monolayer than the other alkali atoms, which indicates that the MoS_2 monolayer is prefer to be anode materials for MIB and be cathode materials for LIB.

Acknowledgements M. Li gratefully acknowledges the financial support from the National Natural Science Foundation of China (Grant No. 50903017), and M. Li thanks the financial support from the Fujian Collaborative Innovation Center of High-End Manufacturing Equipment.

References

1. V. Etacheri, R. Marom, R. Elazari, G. Salitra, D. Aurbach, Energy Environ. Sci. **4**(9), 3243 (2011)
2. L. Lu, X. Han, J. Li, J. Hua, M. Ouyang, J. Power Sources **226**, 272–288 (2013)
3. R. Marom, S.F. Amalraj, N. Leifer, D. Jacob, D. Aurbach, J. Mater. Chem. **21**(27), 9938 (2011)
4. H. Hwang, H. Kim, J. Cho, Nano Lett. **11**(11), 4826–4830 (2011)
5. J.-W. Jiang, Frontiers of Physics **10**(3), 287–302 (2015)
6. Q. Yue, J. Kang, Z. Shao, X. Zhang, S. Chang, G. Wang, S. Qin, J. Li, Phys. Lett. A **376**(12–13), 1166–1170 (2012)
7. A. Castellanos-Gomez, R. van Leeuwen, M. Buscema, H.S. van der Zant, G.A. Steele, W. J. Venstra, Adv. Mater. **25**(46), 6719–6723 (2013)
8. Y. Li, D. Wu, Z. Zhou, C.R. Cabrera, Z. Chen, J. Phys. Chem. Lett. **3**(16), 2221–2227 (2012)
9. R.C. Massé, E. Uchaker, G. Cao, Sci. China Mater. **58**(9), 715–766 (2015)
10. S. Yang, D. Li, T. Zhang, Z. Tao, J. Chen, The Journal of Physical Chemistry C **116**(1), 1307–1312 (2012)
11. K.P.S.S. Hembram, H. Jung, B.C. Yeo, S.J. Pai, H.J. Lee, K.-R. Lee, S.S. Han, Phys. Chem. Chem. Phys. **18**(31), 21391–21397 (2016)
12. Y.S. Meng, M.E. Arroyo-de, Dompablo. Energy Environ. Sci. **2**(6), 589–609 (2009)
13. S.J. Clark, M.D. Segall, C.J. Pickard, P.J. Hasnip, M.I.J. Probert, K. Refson M.C. Payne, Zeitschrift für Kristallographie—Crystalline Materials **220** (5/6) (2005)
14. G. Kresse, I. fu, T.U. ¨r Theoretische Physik, W.H.-. ¨t Wien, A-1040 Wien, Austria, D. Joubert, U. o. t. W. Physics Department, P.O. Wits 2050, Johannesburg, South Africa and R. J. 1998
15. S. Grimme, J. Comput. Chem. **27**(15), 1787–1799 (2006)
16. T.K. Gupta, Physical Review B **43**(7), 5276–5279 (1991)
17. K.F. Mak, C. Lee, J. Hone, J. Shan, T.F. Heinz, Phys. Rev. Lett. **105**(13), 136805 (2010)
18. J. Kang, H. Sahin, F.M. Peeters, Phys. Chem. Chem. Phys. **17**(41), 27742–27749 (2015)
19. E.W. Keong Koh, C.H. Chiu, Y.K. Lim, Y.-W. Zhang, H. Pan, Int. J. Hydrogen Energy **37**(19), 14323–14328 (2012)
20. H.P. Komsa, J. Kotakoski, S. Kurasch, O. Lehtinen, U. Kaiser, A.V. Krasheninnikov, Phys. Rev. Lett. **109**(3), 035503 (2012)
21. S. Xiong, G. Cao, Nanotechnology **26**(18), 185705 (2015)
22. Q. Yue, S. Chang, S. Qin, J. Li, Phys. Lett. A **377**(19–20), 1362–1367 (2013)
23. B. Xu, L. Wang, H.J. Chen, J. Zhao, G. Liu, M.S. Wu, Comput. Mater. Sci. **93**, 86–90 (2014)

Simulation Study of Two-Phase Flow in the Siphon Pipeline

Yao Lei, Changwei Wang and Yuxia Ji

Abstract With the development of world economy and the growing population, the shortage of water resources has become a major issue which affects survival and development of human beings. Hence, water-saving toilet is promoted and considered as the most effective way to solve this problem. In this paper, a new two-fluid approach was introduced to simulate the gas-liquid flow in the siphon pipeline. In the meantime, the VOF (Volume of Fluid) method and the RNG k-ε model were adopted to analyze the relative factors such as volume fraction, pressure and velocity magnitude, and thus the diagram of curves was produced accordingly. The study also shows that the negative pressure and velocity magnitude effect the siphon action directly which is formed by high-speed flow from the jet. Furthermore, the pressure gradient caused by jet and outlet confirmed the stability of the siphon. Additionally, it has not only approved the feasibility of applying the two-fluid method to siphon pipeline, but also provided the theory foundation for improving the performance of water-saving toilet.

Keywords Siphon pipeline · Two-phase flow · VOF method · RNG k-ε model

1 Introduction

Due to the raised living standard and population, more and more people suffer from water shortage, which is increasingly threatening people's survival and development. Especially in our country, it is a serious shortage of water resources with per

Y. Lei (✉) · C. Wang · Y. Ji
School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou, China
e-mail: yaolei@fzu.edu.cn

C. Wang
e-mail: wangchangwei789@163.com

Y. Ji
e-mail: fzu_yuxiaji@163.com

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_33

capita fresh water volume being only one fourth of the world average [1, 2]. Therefore, the problem of water-saving is becoming increasingly pronounced. Among the ways of solving this problem, water-saving toilet is considered as the most effective way. Since sanitary products that do not conform to the standard and requirements for technical specifications may be not produced or sold, the high frequency of use in all sanitary products that it is urgent to develop the water-saving toilet.

In recent years, researchers mainly depended on experience to improve the performance of the water-saving toilet, which requires a lot of time and resources, heavy workload with low efficiency. However, simulations with commercial CFD (Computational Fluid Dynamics) software are emerged with higher efficiency and accuracy which analyzes the flow field and distribution of inner velocity of the siphon pipeline more accurate [3–5]. To design the water-saving toilet, the main problem is the influence of the shape of pipes change on siphon and lines optimization. The optimal design of the siphon pipeline involves the fluid dynamics and the coupling relationship between the structures. Only by fully considering the influence among them, the good product can be designed. In addition, the analytic results could be regarded as the basis for engineering design of water-saving toilet.

2 Theoretical Calculation

Based on the law of mass conservation and hydrodynamics, basic mathematical equation of flow model was established, otherwise known as the continuity equation.

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}(\rho u_i) = 0 \tag{2.1}$$

Where u_i is the corresponding velocity component of the x_i axis in the Descartes coordinate system, t is the flow time and ρ is the fluid density.

The unsteady flow-field of the momentum equation is calculated by using time-averaged method. According to the assumption of Boussinesq's turbulent viscosity in the eddy viscosity model, the corresponding turbulent Reynolds time-averaged equation can be obtained as followed [6]:

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = f_i - \frac{1}{\rho}\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j}\left[v\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) - \overline{u_i' u_j'} \right] \tag{2.2}$$

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = f_i - \frac{1}{\rho}\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j}\left[(v + v_t)\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \right] \tag{2.3}$$

The velocity, pressure and other physical quantities are the time-averaged physical quantities.

The conservation law of momentum should be satisfied in the flow process, and the conserved form of Navier-Stokes equation can be written as [7]:

$$\frac{\partial p}{\partial t}(\rho u_i) + \frac{\partial}{\partial x_j}(\rho u_i u_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_i}\left[(\mu + \mu_t)\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\right] + F_i \tag{2.4}$$

Where P is the pressure, μ is the fluid viscosity, μ_t is the turbulent viscosity and F_i is the mass force.

Furthermore, if the fluid viscosity is not considered, the Euler equation is an equation of motion of ideal fluid which can be obtained:

$$\frac{du}{dt} = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} = f_x - \frac{\partial p}{\rho \partial x} \tag{2.5}$$

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} = f_y - \frac{\partial p}{\rho \partial y} \tag{2.6}$$

$$\frac{dw}{dt} = \frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z} = f_z - \frac{\partial p}{\rho \partial z} \tag{2.7}$$

Where u, v, and w are the velocity components of the velocity vector on the x, y and z axes, respectively.

The RNG k-ε model is widely applied to calculate several different turbulence and the equations of turbulent can be written as [8]:

$$\frac{\partial}{\partial t}(\rho k) + \frac{\partial}{\partial x_i}(\rho k u_i) = \frac{\partial}{x_j}\left[a_k \mu_{eff}\frac{\partial k}{\partial x_j}\right] + G_k + \rho\varepsilon \tag{2.8}$$

$$\frac{\partial}{\partial t}(\rho \varepsilon) + \frac{\partial}{\partial x_i}(\rho \varepsilon u_i) = \frac{\partial}{x_j}\left[a_\varepsilon \mu_{eff}\frac{\partial \varepsilon}{\partial x_j}\right] + \frac{C^*_{1\varepsilon}}{k}G_k - C_{2\varepsilon}\rho\frac{\varepsilon^2}{k} \tag{2.9}$$

$$C^*_{1\varepsilon} = C_{1\varepsilon} - \frac{\eta(1-\eta/\eta_0)}{1+\beta\eta^3} \tag{2.10}$$

Where μ_{eff} is the coefficient of effective viscous, E_{ij} is the time-averaged strain rate and the coefficients a_k and a_ε are the Prandtl numbers corresponding to the turbulent kinetic energy k and the dissipation rate ε, respectively.

3 Numerical Simulation

The sketch of Numerical Simulation is showed in Fig. 1.

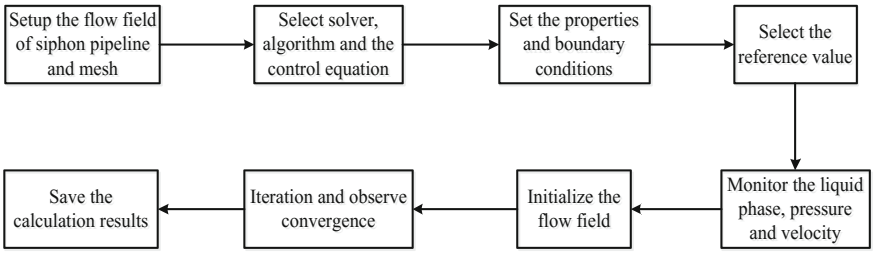


Fig. 1 The sketch of Numerical Simulation

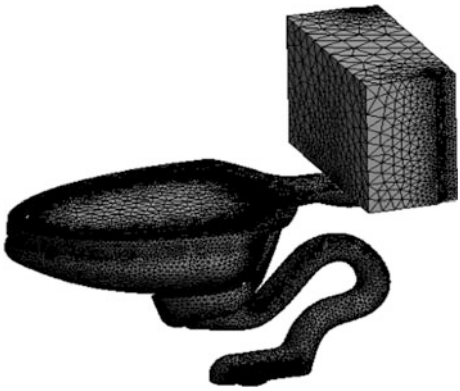


Fig. 2 The grid distribution of siphon toilet

3.1 Grid Partition

Since the structure is complicated, how to generate a reasonable grid to save the time on the grid partition for CFD is very important. The quality of grid directly affects the precision of simulation and the convergence of computing. In addition, this effect is even decisive in many cases. In this paper, the meshes of the flow channel model are divided by Mesh module in ANSYS Workbench. Considering the time and numerical dissipation of initialization, we use triangular and tetrahedral meshes to divide the grid so that the number of grid is more reasonable with higher efficiency. Owing to the shape of the basin, the siphon pipe and the connections of the components are irregular, the quality of several parts of the mesh is not up to the requirement of calculation, and it needs to refine and adjust the local position. The grid distribution of siphon toilet is shown in Fig. 2.

3.2 Calculation Conditions

Considering the flushing process is a two-phase transient flow with free surface, the absolute velocity equation, the three-dimensional transient model and the pressure-based method are used for the iteration. For the actual process of flushing, part of the water descends by its own gravity to siphon pipeline, thus the influence on the function coming from the gravitational field must be considered to establish the equation. Through analyzing the characteristics of multiphase flow, the VOF model and the implicit scheme are selected to simulate the two-phase flow in the siphon pipeline [9, 10]. Due to the larger Reynolds number which is more than 20,000 in siphon pipeline, the flow state is defined as the turbulence and the RNG *k*–epsilon turbulence model is employed in the calculation. In order to improve the accuracy of calculation, discrete format and under-relaxation factors are considered in the algorithm [11]. Moreover, the mass and momentum conservation equations are adopted with the Simple algorithm. Also, turbulent kinetic energy and turbulent dissipation rate equation are calculated using the first order upwind scheme. In the initial state, there is an aqueous phase in the tank and the siphon pipeline, so the initial liquid phase is settled in the process. Additionally, in order to better observe the pressure and the velocity in the siphon pipeline, the surface monitors on the symmetrical surface are also settled in the simulation.

3.3 Materials and Boundary Conditions

According to siphon toilet flushing process, the simulation of materials and boundary conditions are set as followed:

(1) The air is set as the primary phase and the liquid water is set as the secondary phase.
(2) The density of the air is 1.225 kg/m^3 and the viscosity is 1.7894×10^{-5} kg/(m·s).
(3) The liquid phase is water, its density is 998.2 kg/m^3 and viscosity is 1.003×10^{-3} kg/(m·s).
(4) The operating pressure is 101,325 Pa and the reference pressure location should be established in the area filled with air.
(5) Gravitational acceleration is 9.8 m/s and the specified operating density selects the density of the gas phase.
(6) Wall conditions should be set to non-sliding; the upper surface of the tank is defined as a pressure inlet; the seat outlet and siphon outlet are defined as pressure outlets.

4 Simulation Result

Figure 3 illustrates the flushing process of liquid flow in the siphon toilet. As shown in the Figure, the simulation results are divided into four stages: the beginning of the overflow, the formation of siphon, the stability of siphon and the instability of siphon. The color on the left of the figure is corresponded to the percentage of the liquid volume fraction: the red areas show the pure liquid phase, and the blue areas show the pure gas phase. When the flushing process begins, the flushing valve is opened, and most of the water rapidly shot from the injection orifice along the pipeline. In the meantime, the water column is rapidly flowing and mixed with air from the jet to the inner wall of the pelvic cavity. Through the inner wall of pipeline, the water began to overflow, and then continuously brings out the air which is forming unstable negative pressure and resulting in siphon phenomenon. At last, the water filled the entire siphon pipeline quickly to form a stable negative pressure to achieve the steady state of siphon. With the help of negative pressure, the water constantly flow out of pelvic until the water is not enough to maintain the

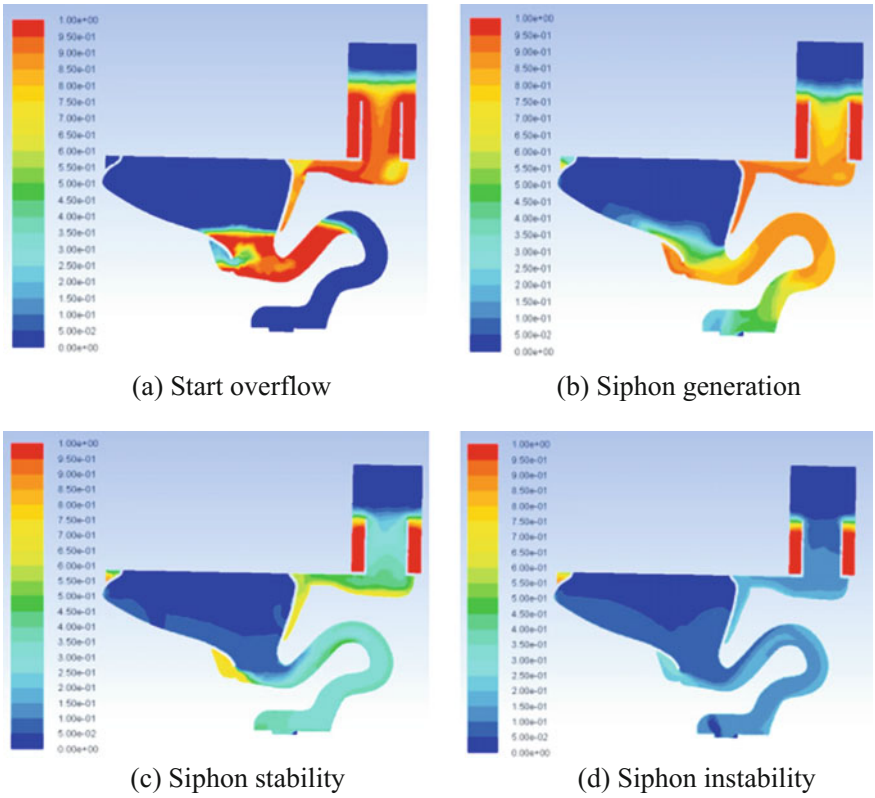


Fig. 3 Distribution of liquid phase on the symmetrical section of the toilet

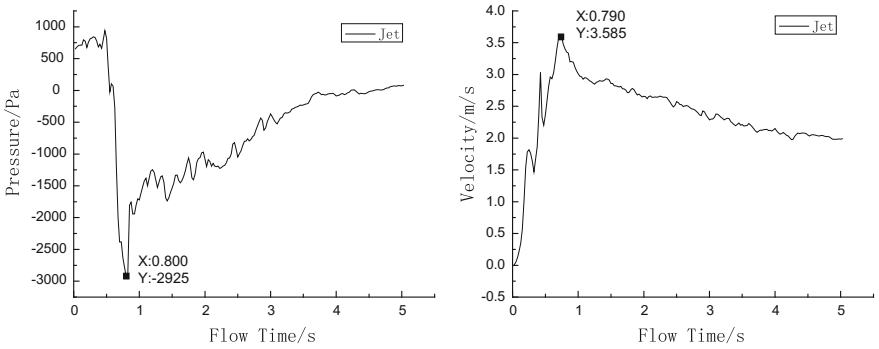
steady state of siphon. Eventually, the unsteady state of siphon occurs due to cavities collapsing with water column until the flushing process is completed.

Based on above analysis, it is clear that the siphon action is mainly decided by the negative pressure, the flow velocity of the jet and the top of siphon pipeline. Therefore, monitoring points are set up at the center of the jet and the top of the pipeline respectively in order to obtain the pressure and fluid flow.

Variation of the pressure and velocity on the Jet is presented in the Fig. 4. It can be seen that the pressure is positive at the beginning and descends rapidly after 0.5 s. At 0.8 s, it reaches a maximum of −2925 Pa. In the meantime, the flow velocity is rising rapidly to reach the maximum flow velocity of 3.585 m/s, and then slowly decline. This is because the water mixed with amounts of air under the action of gravity quickly flow from the jet. In this process, most of the potential energy can be converted into pressure energy. The considerable negative pressure is generated, which leads to the formation of siphon.

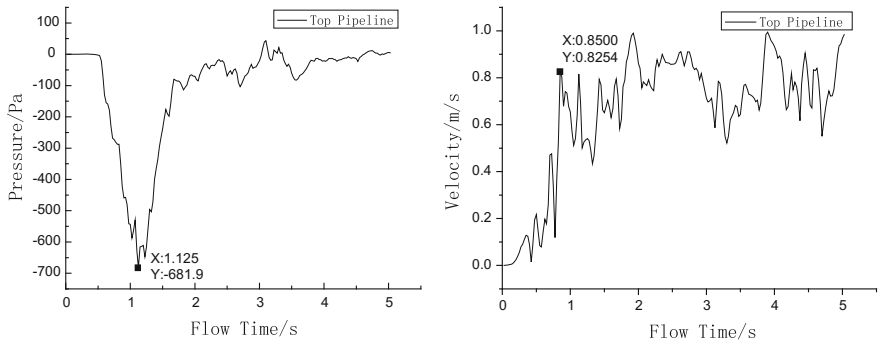
Figure 5 shows variation of the pressure and velocity on the top of pipeline. It is interesting to note that the rapid decline in pressure after about 0.5 s and reaches the lowest negative and then suddenly rapid increases, however, it keeps stable about 50 Pa fluctuations in a small range. Also, the flow velocity is volatility rises first at 0.8 m/s with in small scope. This is because the top of the siphon pipe with the largest pressure force the overflow of water continue to bring out the air, which making the area to form a negative pressure.

In conclusion, siphon phenomenon do not generate spontaneously. When the water is passing through the nozzle of the jet, the high-speed flow forms a zone of negative pressure in the pipeline, there by leads to promote the formation of siphon, and the pressure gradient on jet and outlet maintain the stability of siphon.



(a) Variation of pressure (b) Variation of velocity

Fig. 4 Variation of the pressure and velocity on the Jet



(a) Variation of pressure (b) Variation of velocity

Fig. 5 Variation of the pressure and velocity on the top of pipeline

5 Conclusions

Based on computational fluid dynamics, the gas-liquid two-phase flow in the siphon pipeline was numerically analyzed by using the commercial CFD software FLUENT. The contents of this paper mainly include the following aspects.

(1) The siphon action is simulated effectively by the gas-liquid two-phase flow model and demonstrated by the flow field of the air and water.
(2) Through analyzing the state of gas-liquid two-phase flow in the siphon pipeline, it is found that the siphon phenomenon is related to negative pressure and flow velocity in the pipeline, which causing more obvious siphon action.
(3) Through the data analysis, the high speed flow from the jet creates negative pressure conditions to promote the formation of siphon, and then the pressure gradient caused by jet and outlet confirmed the stability of the siphon. While the siphon phenomenon occurs earlier, the siphon action is stronger and the water saving is more obvious. Additionally, the results of analysis can provide the theory foundation for the development of high-quality water-saving toilet.

Acknowledgements This project is supported by National Natural Science Foundation of China (Grant No. 51505087) and The Education Department of Fujian Province (Grant No. JA15054).

References

1. Y. Ma, J. Bai, New national standard of building sanitary ceramics on the implementation of strategic significance [J]. Journal of Ceramics **5**, 45–47 (2016)
2. Y. Li, 66% water-saving toilet waste 21 tons/year of water [J]. China Quality Journey **2**, 18–21 (2016)

3. Z. Liu, M. Xiong, Z. Yu, Unsteady siphon flow controlling and optimizing based on dynamic flow condition [J]. Journal of Hunan University: Natural Sciences. **10**, 31–35 (2012)
4. Y. Zhou, G. Yan, Numerical simulation for flow field of mixed elbow based on fluent [J]. Machinery Design and Manufacture **9**, 211–213 (2011)
5. T. Yan, L. Chen, M. Xu et al., Siphon Pipeline Resistance Characteristic Research [J]. Procedia Engineering **28**(8), 99–104 (2012)
6. P. Rollet-Miet, D. Laurence, J. Ferziger, LES and RANS of turbulent flow in tube bundles [J]. Int. J. Heat Fluid Flow **20**(3), 241–254 (1999)
7. Z. Han, J. Wang, X. Lan, *Fluid simulation calculation on engineering examples and application [M]* (Beijing Institute of Technology Press, Beijing, 2008)
8. D. C. Wilcox. Turbulence modeling for CFD [M]. DCW Industries, 2006
9. Z. Peng, *Optimization design in siphon pipeline based on the computational fluid mechanics [D]* (Hunan university, Changsha, 2009)
10. E. Aulisa, S. Manservisi, R. Scardovelli, A mixed markers and volume-of-fluid method for the reconstruction and advection of interfaces in two-phase and free-boundary flows. J. Comput. Phys. **188**(2), 611–639 (2003)
11. Z. Zhou, *Basic theory and practical application in computational fluid mechanics [M]* (Southeast University Press, Nanjing, 2008)

Molecular Dynamics Simulations on Nanoindentation Experiment of Single-Layer MoS_2 Circular Nanosheets

Jianyue Hu, Minglin Li, Weidong Wang and Longlong Li

Abstract The single-layer molybdenum disulfide ($SLMoS_2$) belongs to a graphene-like layered two-dimensional (2D) nanocrystal, consisting of S-Mo-S sandwich structure. The pristine $SLMoS_2$ holds unique properties due to its intrinsic direct bandgap of $\sim$1.8 eV and high elastic modulus of $\sim$0.2 TPa, which attracts extensive applications on 2D nanodevices. Since Lee's pioneering work on graphene, the nanoindentation experiment based on the elastic membrane theory is routinized for obtaining the in-plane elastic modulus of other 2D nanomaterials, including $SLMoS_2$. Compared with the widely investigation of graphene nanoindentation with molecular dynamics simulations (MDs), only few studies focus on nanoindentation MDs of $SLMoS_2$. In the present work, we perform MDs for the $SLMoS_2$ nanoindentation to investigate the effects of indenter radius, loading speeds, and single S vacancies point defects on its mechanical response. Typical indentation load–displacement curves are obtained and fitted to deduce the elastic modulus of $SLMoS_2$ and the pre-tension. It is found that the extracted elastic modulus and pre-tension of defect-free $SLMoS_2$ increases with the indenter radius. Faster the indentation loading is extremely higher the elastic modulus of $SLMoS_2$ presents. The elastic modulus of $SLMoS_2$ with S point defects has the tendency to increase as the defect concentration varying from 0.1% to 1.0%, which is more obvious as the indenter radius becomes small.

J. Hu · M. Li (✉)
School of Mechanical Engineering and Automation, Fuzhou University, 2, Xueyuan Road, Minhou, Fuzhou 350116, Fujian, China
e-mail: liminglin@fzu.edu.cn

J. Hu
e-mail: 634180160@qq.com

W. Wang (✉) · L. Li
School of Mechanical-Electronic Engineering, Xidian University, Xi'an 710071, China
e-mail: wangwd@mail.xidian.edu.cn

J. Hu
Center of New Technology Development, Fujian Province Equipment Inspection Institute, 370 Lubin Road, Cangshan, Fuzhou 350008, Fujian, China

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9_34

Keywords Molecular dynamics simulation · Nanoindentation · Young's modulus

1 Introduction

The single-layer molybdenum disulfide ($SLMoS_2$) belongs to graphene-like layered two-dimensional (2D) nanocrystals, consisting of the S-Mo-S sandwich structure. The pristine $SLMoS_2$ holds unique properties due to its intrinsic direct bandgap of $\sim$1.8 eV [1] and high elastic modulus of $\sim$0.2 TPa [2–4], which attracts extensive applications on 2D nanodevices. Since Lee's pioneering work on graphene [5], the nanoindentation experiment based on the elastic membrane theory is routinized for obtaining the in-plane elastic modulus of other 2D nanomaterials, including $SLMoS_2$ [2–4]. Compared with the widely investigation of graphene nanoindentation with molecular dynamics simulations (MDs), only few studies focus on nanoindentation MDs of $SLMoS_2$ [6, 7]. Zhao et al. [6] use MD nanoindentation simulations to obtained the temperature-dependent mechanical properties of $SLMoS_2$ with the Stillinger-Weber (SW) potential. Stewart et al. [7] has predicted the mechanical properties of MoS_2 at a limited thermal conditions. Although the above researches have been carried out for mechanical properties of $SLMoS_2$, the values of elastic modulus of the $SLMoS_2$ have not yet reached a consensus. In addition, few papers systematically describe how various factors, such as indenter radii, loading speed S point defects, influence the mechanical property during the nanoindentation process for the $SLMoS_2$. In the work, we perform MDs for the $SLMoS_2$ nanoindentation to investigate the effects of indenter radius, loading speeds and S point defects on its mechanical response.

2 Models and Methods

We carry out a set of classical MD simulations on nanoindentation of the circular $SLMoS_2$. In Fig. 1, the model is sketched as a circular ring with an inner radius r and an outer radius R. In our works, these radii of r and R are specified as 10 nm and 11 nm, respectively, and the model has total atom number of 10864, with the middle layer of Mo atoms (colored in red) sandwiched between the top layer of S atoms (colored in green) and the bottom layer of S atoms (colored in blue). The process of the nanoindentation experiment is shown in the Fig. 1b–d. The atoms colored in purple are fixed after the system is relaxed to mimic a nanosheet suspended over a substract with a circular hole. A gray diamond sphere is initially layed above the circular nanosheet with the height enough to avoid the initial interaction between them. When the indentation begins, the diamond sphere gradually moves down until penetrating through the circular nanosheet.

All simulations are carried out by use of open source large-scale atomic/molecular massively parallel simulator (LAMMPS) [8]. The open source software, Visual Molecular Dynamics (VMD), is used to visualize the nanoindentation

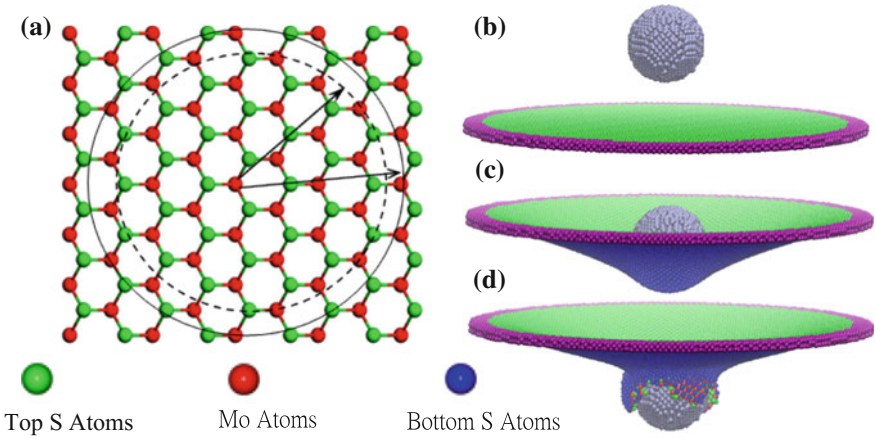


Fig. 1 Atomic configuration of the $SLMoS_2$ for the nanoindentation experiment. **a** the circular model, **b–d** the process of a diamond sphere penetrate through the circular nanosheet

process of $SLMoS_2$ sheets. The key of MD simulation is the selection of atom interaction potential function, which directly determines the accuracy of the simulation. The interatomic interaction in $SLMoS_2$ and the C-C interaction in diamond indenter are described by the REBO potential [9] and the AIREBO potential [10], respectively. The Lennard-Jones (L-J) potential [11] is used to model the interaction between the diamond indenter and the $SLMoS_2$. Before the indentation, the initial atomic structure of the circular nanosheet is relaxed by using of the energy minimization with the conjugate gradient method, followed by a 50 ps MD run with the NPT ensemble based on the Nose/Hoover thermostat and barstat method. The ambient temperature and pressure are controlled to be 0.1 K and 0.1 bar, respectively. In addtion, the indenter is regarded as a rigid object so that the atomic configuration of the diamond indenter do not change during the MD simulations. After the system is relaxed, the boundary is fixed and the circular $SLMoS_2$ is transferred to be controlled with the NVT ensemble. The velocity-verlet algorithm is adopted to solve the Newtonian equation with a timestep of 1 fs. During the simulation, the indenter force and the deflection of the nanosheet are recorded to be used to extract the Young's modulus and the pretension stress, based on the following equation [5, 12]

$$F = t\sigma_0 \pi R_{MoS_2} \left(\frac{\delta}{R_{MoS_2}}\right) + Etq^3 R_{MoS_2} \left(\frac{\delta}{R_{MoS_2}}\right)^3 \tag{2.1}$$

where F is the indenter force, δ is the maximum deflection of the nanosheet subjected to the indenter. t is the thickness of the $SLMoS_2$ given as 6.15 Å, and $q = 1/(1.05 - 0.15v - 0.16v^2)$ is a dimensionless constant, where v is the Poisson's ratio of bulk MoS_2, v = 0.27 [5]. E and σ_0 are the Young's modulus and pretension stress, respectively.

3 Results and Discussion

In the experiments of indentation with AFM, the radius of the indenter is much less than the suspended samples. In MD simulations, we change the radius ratio of the indenter to the circular nanosheet to investigate the effect of indenter radius on the mechanical response of the $SLMoS_2$, as shown in Fig. 2. It is clearly observed that increasing the radius ratio produces larger indenter forces and maximum deflections. Unlike MD simulations for graphenes, the indenter force smoothly increases until a sudden drop present, followed by a rough curve due to the atomic fluctuation, a plastic behavior, during the indentation. Therefore, to extract the in-plane elastic modulus of the nanosheet, curves AC (all data before a plastic behavior) and AB (small deflection range) can be used to the fitting operation, as shown in Fig. 3a. With the AC and AB curves, the Young's modulus and pretension stress can be deduced, as shown in Fig. 3b, c. It is worth noting that the Young's moduli obtained from the AB curves are larger than those from the AC curves, whereas the pretension stress obtained from the AB curves are less than those from the AC curves. However, both of them are in agreement with the experimental results ($\sim$200 GPa). In addition, both Young's modulus and pretension stress increase with the indenter radius. To mimic the indentation experiment with AFM, the radius ratio of 0.2 is used in the following simulations.

Compared with the experimental indentation with AFM, the loading spead of the indenter in MD simulaitons is much faster. To see the effect of the loading spead on the indentation, the loading spead of the indenter is changed from 0.1 Å/ps to 1.0 Å/ps. Figure 4 shows the relation curves between the indenter force and the deflection of the circular nanosheet with different loading spead. At an identical deflection, the indenter force is found to increase with the loading spead. However, the curve for the spead of 0.2 Å/ps is closed to that for the spead of 0.1 Å/ps. In such cases, the speed of 0.20 Å/ps can be regarded as the critical loading speed. When the loading speed decreases from 1.0 Å/ps to 0.5 Å/ps, with the decline of

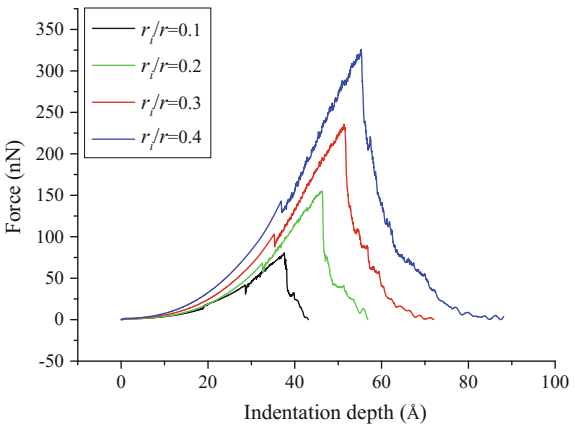


Fig. 2 The indenter force versus the deflection of the circular nanosheet with different radius ratio of the indenter to the circular nanosheet, ranging from 0.1 to 0.4

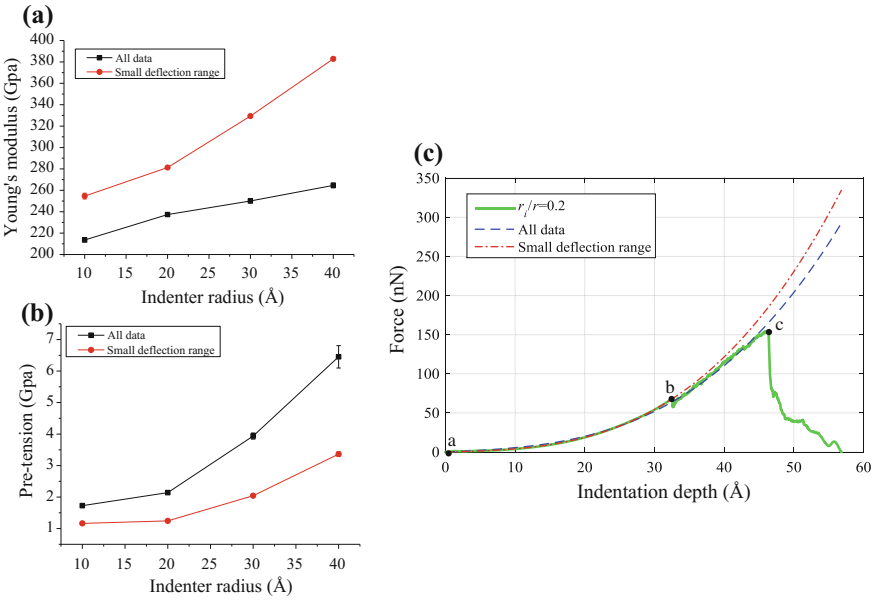


Fig. 3 Based on fitting the AC and AB curves (**a**), the Young's modulus (**b**) and pretension stress (**c**) are shown varying with the indenter radius

50%, the value of the maximum deflection will increase by 37%, and the value of the maximum indenter force will decrease by 31.1% correspondingly. Nevertheless, when the loading speed decreases from 0.2 Å/ps to 0.1 Å/ps, with the decline of 50%, the maximum indenter force decreases only 0.904%, and the deflection just increases 2.312%. Such results can be contributed to the fact that higher the loading speed of the indenter is, less time it will take to penetrate through the $SLMoS_2$, leading to the higher indenter force and lower deflection. In one word, the loading

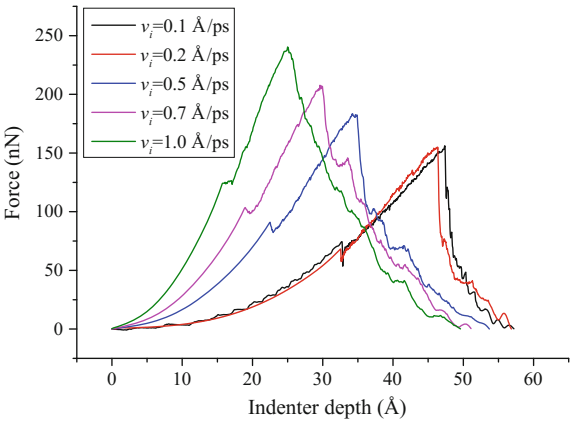


Fig. 4 The indenter force related to the deflection with different loading speeds ranging from 0.1 Å/ps to 1.0 Å/ps

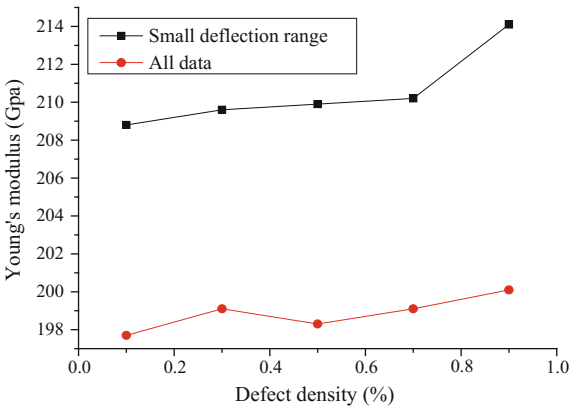


Fig. 5 The Young's modulus of the $SLMoS_2$ with VS point defects versus the defect concentration

spead of 0.2 Å/ps is advisable to simulate the indentation experiment of the $SLMoS_2$ nanosheet.

During the preparation, characterization and mainipulation of the $SLMoS_2$, several types and various concentrations of intrinsic point defects and grain boundaries [13–15] will be introduced. The performance of the material may benefit or suffer from such defects. The monosulfur vacancy (VS) is the most common point defect, frequently observed in experiments for its lowest formation energy (1.1 eV). Therefore, we study the Young's modulus of the $SLMoS_2$ with VS point defects versus the defect concentration, are shown in Fig. 5. The elastic modulus of $SLMoS_2$ with VS point defects has the tendency to increase with the defect concentration varying from 0.1% to 1.0%, but the increasing amplitude is not very significant.

4 Conclusions

In order to investigate the mechanical properties of $SLMoS_2$ with nanoindentation experiments, we carry out a set of MD simulations. The effect of the indenter radius, loading spead, and VS point defect concentration on the mechanical properties of the $SLMoS_2$ circular nanosheet are discussed in this work. Results show that increasing the indenter radius and loading speed produce increasing the indenter force at the same deflection, as well as the Young's modlus and the pretension stress. Moreover, it is surprising that increasing the concentration of VS point defect leads to slightly increasing in the Young's modulus of $SLMoS_2$.

Acknowledgements The authors gratefully acknowledge the financial support from the National Natural Science Foundation of China (Grant No. 50903017 and 51205302), and M. Li thanks the financial support from the Fujian Collaborative Innovation Center of High-End Manufacturing Equipment.

References

1. K.F. Mak, C. Lee, J. Hone, J. Shan, T.F. Heinz, Phys. Rev. Lett. **105**(13), 136805 (2010)
2. S. Bertolazzi, J. Brivio, A. Kis, ACS Nano **5**(12), 9703–9709 (2011)
3. A. Castellanos-Gomez, M. Poot, G.A. Steele, H.S. van der Zant, N. Agraït, G. Rubio-Bollinger, Nanoscale Res. Lett. **7**(1), 1–4 (2012)
4. R.C. Cooper, C. Lee, C.A. Marianetti, X. Wei, J. Hone, J.W. Kysar, Physical Review B **87**(3), 035423 (2013)
5. C. Lee, X. Wei, J.W. Kysar, J. Hone, Science **321**(5887), 385–388 (2008)
6. j. z. j.-w. j. a. T. Rabczuk, Applied Physics Letters 2013 (2013)
7. J.A. Stewart, D.E. Spearot, Modell. Simul. Mater. Sci. Eng. **21**(4), 045003 (2013)
8. S. Plimpton, J. Comput. Phys. **117**(1), 1–19 (1995)
9. S. Xiong, G. Cao, Nanotechnology **26**(18), 185705 (2015)
10. W. Wang, S. Li, J. Min, C. Yi, Y. Zhan, M. Li, Nanoscale Res. Lett. **9**(1), 1–8 (2014)
11. J.E. Jones, Proceedings of the Royal Society of London **106**(2), 303–313 (1924)
12. U. Komaragiri, M. Begley, J. Simmonds, J. Appl. Mech. **72**(2), 203–212 (2005)
13. J. Hong, Z. Hu, M. Probert, K. Li, D. Lv, X. Yang, L. Gu, N. Mao, Q. Feng and L. Xie, Nature communications 6 (2015)
14. W. Zhou, X. Zou, S. Najmaei, Z. Liu, Y. Shi, J. Kong, J. Lou, P.M. Ajayan, B.I. Yakobson, J.-C. Idrobo, Nano Lett. **13**(6), 2615–2622 (2013)
15. H.-P. Komsa, J. Kotakoski, S. Kurasch, O. Lehtinen, U. Kaiser, A.V. Krasheninnikov, Phys. Rev. Lett. **109**(3), 035503 (2012)

Author Index

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9

Subject Index

L. Yao et al. (eds.), *Advanced Mechanical Science and Technology for the Industrial Revolution 4.0*, https://doi.org/10.1007/978-981-10-4109-9